Probabilités et Statistiques

ISEN Nantes

Cours A3

2023 – 2024

Table des matières

CHAPITRE 1

Probabilités et combinatoire

Sommaire

1.1 Notion de probabilité

Une **expérience aléatoire** est une expérience dont on ne peut pas prévoir le résultat par avance et/ou qui donne des résultats différents lorsqu'on la reproduit dans des conditions identiques.

1.1.1 Événement aléatoire

Définition 1.1.1. L'ensemble de tous les résultats possibles d'une expérience aléatoire est appelé l'**univers** de cette expérience. On le note $\boldsymbol{\Omega}$.

Remarque 1.1.2. Un résultat possible de l'expérience est alors un élément de l'univers. On utilise souvent ω pour désigner un résultat dans $\boldsymbol{\Omega}$.

Exemple 1.1.3.

- Lancer d'une pièce : $\boldsymbol{\Omega} =$ . . . .
- Lancer d'un dé : $\boldsymbol{\Omega} =$ . . . .
- Lancer de deux dés : $\boldsymbol{\Omega} =$ . . . .
- Durée de vie d'une bactérie : $\boldsymbol{\Omega} =$ . . . .

- Jeu de fléchette sur une cible circulaire de rayon 15 cm : $\Omega =$ ....
- Cours d'un actif financier sur un intervalle de temps $[t_1; t_2]$: $\Omega =$ ....

Définition 1.1.4. Un **événement aléatoire** est un sous-ensemble de l'univers : $A \subset \Omega$, dont on veut calculer la probabilité de réalisation.

Exemple 1.1.5.

- Lancer d'une pièce : la partie $\{F\}$ correspond à l'événement " ...."
- Lancer d'un dé : la partie .... correspond à l'événement "Le dé tombe sur un résultat impair."
- Lancer de n pièces (0 ou 1) : $\Omega = \{0; 1\}^n$ et la partie $\left\{\omega \in \Omega \mid \sum_{k=1}^{n} \omega_k \geq \frac{n}{2}\right\}$ correspond à l'événement " ...."
- Lancer de deux pièces : la partie .... correspond à l'événement "La somme des deux lancers est inférieure à 4."
- Cours d'un actif : la partie $\left\{f \in C^0\left([t_1; t_2], \mathbb{R}_+^*\right) \mid \forall t \in [t_1; t_2],\ f(t) \leq 1000\right\}$ correspond à l'événement " ...."

Remarque 1.1.6 (Rappels : opérations ensemblistes)**.** Si A et B sont deux événements liés à une expérience aléatoire :

- A^c ou $\overline{A}$ ou $\Omega \setminus A$ est l'événement contraire de A. Il est réalisé ssi A n'est pas réalisé.
- $A \cap B$ est réalisé ssi A et B sont réalisés simultanément.
- $A \cup B$ est réalisé ssi A ou B est réalisé (ou les deux).
- $A \subset B$ signifie que si A est réalisé alors B est aussi réalisé.
- $A \cap B = \emptyset$ signifie que A et B sont incompatibles (ils ne peuvent pas être réalisés en même temps).
- Ω est l'événement certain, $\emptyset$ est l'événement impossible.

1.1.2 Tribu

Définition 1.1.7 (Tribu)**.** Une **tribu** (ou σ-algèbre) sur Ω est une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ qui vérifie :

- $\emptyset \in \mathcal{A}$;
- $A \in \mathcal{A} \Longrightarrow \overline{A} \in \mathcal{A}$;

- Si pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{A}$ alors $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Remarque 1.1.8. La tribu sera en fait l'ensemble des événements observables, c'est-à-dire ceux dont on pourra calculer la probabilité de réalisation.
A priori, la tribu n'est pas $\mathcal{P}(\Omega)$ tout entier.

Exemple 1.1.9.

- $\{\emptyset; \Omega\}$ est la tribu grossière.
- Si $A \subset \Omega$, $\{\emptyset; A; \overline{A}; \Omega\}$ est la tribu de Bernoulli.
- $\mathcal{P}(\Omega)$ est une tribu.

Proposition 1.1.10 (Propriétés d'une tribu). *Soit $\mathcal{A}$ une tribu sur Ω,*

- $\emptyset \in \mathcal{A}$ *et* $\Omega \in \mathcal{A}$ *;*
- *Si pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{A}$ alors $\displaystyle\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

A retenir sur les tribus : La définition et les propriétés ne sont pas bien importantes pour nous. Dans notre cours, il y aura deux cas :

- Ω est fini ou dénombrable, auquel cas la tribu est $\mathcal{P}(\Omega)$.
- Ω est un intervalle de $\mathbb{R}$ ou une partie de $\mathbb{R}^n$, auquel cas la tribu sera $\mathcal{B}(\Omega)$ que l'on appelle tribu borélienne. Tout ce qu'il faut retenir, c'est que cette tribu contient tous les intervalles (ou pavés en dimension supérieure). On pourra donc calculer la probabilité de n'importe quelle réunion d'intervalles.

1.1.3 Probabilité

Définition 1.1.11 (Probabilité). Soit $(\Omega, \mathcal{A})$ un univers muni d'une tribu. Une **probabilité** sur $(\Omega, \mathcal{A})$ est une fonction $\mathbb{P} : \mathcal{A} \longrightarrow [0; 1]$ qui vérifie

- $\mathbb{P}(\Omega) = 1$;
- Si pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{A}$ et les A_n sont deux à deux disjoints, alors

$$\mathbb{P}\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

$\mathbb{P}(A)$ est alors la probabilité de l'événement A.

Remarque 1.1.12. La probabilité d'un événement est toujours un réel entre 0 et 1.

Proposition 1.1.13 (Propriétés). *Soit $\mathbb{P}$ une probabilité sur $(\Omega, \mathcal{A})$:*

- $\mathbb{P}(\emptyset) = 0$ *;*
- *Si $A \in \mathcal{A}$ alors*

$$\mathbb{P}\left(\overline{A}\right) = 1 - \mathbb{P}(A).$$

- *Si* $A \subset B$ *alors* $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- *Si* A *et* B *sont dans* $\mathcal{A}$ *avec* $A \cap B = \emptyset$, *alors*

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

- *Si* A *et* B *sont dans* $\mathcal{A}$ *alors*

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

Si $\Omega = [\![1;6]\!]$ et pour tout $i \in [\![1;6]\!]$, $\mathbb{P}(\{i\}) = \frac{1}{6}$, on définit $A = \{1;3;5\}$ et $B = \{4;5;6\}$. On a alors

$$\mathbb{P}(A \cup B) = \frac{5}{6} \quad \text{et} \quad \mathbb{P}(A) + \mathbb{P}(B) = 1.$$

1.2 Combinatoire

1.2.1 Probabilité uniforme

Voyons ici un cas particulier très utilisé en pratique : supposons que Ω est fini de cardinal $n = \text{card}(\Omega) = |\Omega|$ et que chaque résultat a la même probabilité de se produire. On dit que les résultats possibles sont **équiprobables**.

Exemple 1.2.1.

- Pile ou Face : $\Omega = \{0;1\}$ avec $\mathbb{P}(\{0\}) = \mathbb{P}(\{1\}) = \ \ldots$.
- Lancer d'un dé : $\Omega = [\![1;6]\!]$ avec $\mathbb{P}(\{i\}) = \ \ldots$ pour tout $i \in [\![1;6]\!]$.
- Choix d'une carte dans un jeu mélangé de 32 cartes : $\Omega = [\![1;32]\!]$ avec $\mathbb{P}(\{i\}) = \ \ldots$. pour tout i.
- Tirage d'une boule dans une urne de n boules : $\Omega = [\![1;n]\!]$ avec $\mathbb{P}(\{i\}) = \ \ldots$ pour tout i.

Définition 1.2.2. Si Ω est fini, la **probabilité uniforme** sur Ω est la probabilité $\mathbb{P}$ telle que

$$\forall \omega \in \Omega,\ \mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}.$$

Dans ce cas, pour tout événement $A \subset \Omega$, on a

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{nombre de résultats favorables}}{\text{nombre de résultats total}}.$$

Remarque 1.2.3.

- C'est de cette situation qu'on parle quand on utilise l'expression "au hasard".
- Le calcul d'une probabilité se ramène ici à du dénombrement, de la combinatoire : combien y a-t-il de résultats favorables ?

1.2.2 Rappels de dénombrement

Principes additif et multiplicatif

- Soient E et F des ensembles disjoints, alors le cardinal de $E \cup F$ est égal à $\mathrm{card}(E) + \mathrm{card}(F)$.
  Dit autrement : si l'on est dans la situation 1 ou la situation 2 (et pas les deux), on additionne les résultats.

- Soient E et F deux ensembles alors le cardinal de $E \times F$ est $\mathrm{card}(E) \times \mathrm{card}(F)$.
  Dit autrement : quand on doit choisir un premier terme puis choisir un deuxième terme (indépendamment du premier choix), on multiplie les résultats.

Choisir p éléments parmi n éléments Soit E un ensemble contenant n éléments. Soit $p \in \mathbb{N}$.

- **Listes :** Une p-liste (ou un p-uplet) de E est un élément de E^p, c'est-à-dire une énumération ordonnée de p éléments de E.
  Le nombre de p-listes d'un ensemble à n éléments est n^p.

- **Arrangements :** Un p-arrangement de E est une p-liste de E dont les éléments sont deux à deux distincts. Ainsi, si $p > n$, il ne peut exister de p-arrangement de E.
  Si $p \leq n$, le nombre de p-arrangements d'un ensemble à n éléments est $n(n-1)\cdots(n-p+1)$.

- **Permutations :** Une permutation est un n-arrangement d'un ensemble à n éléments.
  Il y a donc $n!$ permutations d'un ensemble de cardinal n.

- **Combinaisons :** Une p-combinaison de E est une partie à p éléments de E.
  Il y a $\binom{n}{p}$ p-combinaisons d'un ensemble à n éléments.

Résumé :

	L'ordre compte	L'ordre ne compte pas
Les éléments peuvent se répéter	Liste	-
Les éléments sont distincts	Arrangement	Combinaison

Coefficients binomiaux

- On a $\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon} \end{cases}$.

- Si $1 \leq k \leq n$, on a
$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

- Pour tout $0 \leq k \leq n$,
$$\binom{n}{n-k} = \binom{n}{k}.$$

- Pour tout $0 \leq k \leq n$,
$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$
Cette relation permet de construire le triangle de Pascal.

- Cas particuliers à connaître : pour tout n,
$$\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n \quad \text{et} \quad \binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}.$$

1.2.3 Probabilité non uniforme

Ce n'est pas parce qu'il n'y a que deux possibilités qu'il y a une chance sur deux !

Définition 1.2.4 (Expérience de Bernoulli)**.** Une expérience de Bernoulli est une expérience aléatoire avec seulement deux résultats possibles : succès ou échec. La probabilité de succès est notée p, la probabilité d'échec est alors $1 - p$.

Exemple 1.2.5. On lance un dé et on regarde si on obtient un 6 (succès) ou pas (échec). Ceci est une expérience de Bernoulli, de probabilité de succès . . ..

Proposition 1.2.6.
- *On effectue n fois la même expérience de Bernoulli, la probabilité d'avoir k succès (et donc $n - k$ échecs) est* . . . .
- *On effectue la même expérience de Bernoulli jusqu'à obtenir un succès, la probabilité que ce premier succès soit au $k^{\text{ème}}$ essai est* . . . .

Exemple 1.2.7. Autre expérience avec une mesure de probabilité non uniforme : On lance un dé pipé tel que $\mathbb{P}(\{1\}) = \frac{1}{3}$, $\mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \frac{1}{6}$ et $\mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{1}{9}$.
La probabilité d'avoir un résultat impair est . . ..

1.3 Probabilités conditionnelles

Une information supplémentaire sur une expérience aléatoire modifie les probabilités attachées : cela change la vraisemblance de certains événements.

Exemple 1.3.1. On lance un dé. A priori, la probabilité d'obtenir la valeur i est $\frac{1}{6}$.
Si on nous dit que le résultat est pair, il est clair que la probabilité d'avoir 1 tombe à 0, ainsi que celle d'avoir 3 ou 5.
Que vaut la nouvelle probabilité d'obtenir 2 ? . . . .
. . . .
. . . .

Définition 1.3.2. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Soient $A, B \in \mathcal{A}$ avec $\mathbb{P}(B) > 0$.
La **probabilité de A sachant B** est
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

On note parfois cette probabilité $\mathbb{P}_B(A)$.

On a toujours $\mathbb{P}\left(\overline{A} \mid B\right) = 1 - \mathbb{P}(A \mid B)$ mais en général $\mathbb{P}\left(A \mid \overline{B}\right) \neq 1 - \mathbb{P}(A \mid B)$.

Définition 1.3.3. Un **système complet d'événements** de $(\Omega, \mathcal{A}, \mathbb{P})$ est une famille $A_1, \dots, A_n$ d'événements de $\mathcal{A}$ telle que :

- pour tout i, $\mathbb{P}(A_i) > 0$;
- pour tous $i \neq j$, $\mathbb{P}(A_i \cap A_j) = 0$;
- $\displaystyle\bigcup_{i=0}^{n} A_i = \Omega$.

Théorème 1.3.1 (Formule des probabilités totales). *Soit $A \in \mathcal{A}$ et $B_1, \dots, B_n$ un système complet d'événements. Alors*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i).$$

Théorème 1.3.2 (Formule de Bayes). *Soient A, $B \in \mathcal{A}$ avec $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Alors*

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

Remarque 1.3.4. Souvent, on combine ces deux derniers résultats. Si on connaît $\mathbb{P}(A)$, $\mathbb{P}(B \mid A)$ et $\mathbb{P}(B \mid \overline{A})$, on peut calculer $\mathbb{P}(A \mid B)$:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A)\mathbb{P}(B \mid A)}{\mathbb{P}(A)\mathbb{P}(B \mid A) + \mathbb{P}(\overline{A})\mathbb{P}(B \mid \overline{A})}.$$

1.4 Indépendance d'événements

Deux événements sont dits indépendants si le fait de savoir que l'un s'est réalisé ne donne aucune information sur les chances de réalisation de l'autre.

Définition 1.4.1. Soient A et B deux événements (avec $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$). On dit que A et B sont **indépendants** si

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

(ou de façon équivalente si $\mathbb{P}(B \mid A) = \mathbb{P}(B)$).

Proposition 1.4.2. *A et B sont indépendants si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

L'indépendance et l'incompatibilité ne sont pas du tout la même chose ! Au contraire : si deux événements sont incompatibles, ils ne peuvent pas être indépendants puisque le fait de savoir que l'un s'est réalisé assure que l'autre ne se réalise pas.

Remarque 1.4.3. Lorsqu'on a calculé des probabilités avec le dénombrement, on a utilisé une hypothèse d'indépendance sans le dire !

Exemple 1.4.4.

- On lance un même dé deux fois de suite : le résultat du deuxième lancer est indépendant du premier lancer.

- On tire une carte dans un jeu de 52 cartes, on note A l'événement "La carte est une Dame" et B l'événement "La carte est un Coeur".
  On a alors
  $$\mathbb{P}(A \cap B) = \frac{1}{52} = \mathbb{P}(A)\mathbb{P}(B).$$
  Les événements A et B sont indépendants.

- On lance deux dés, on note A l'événement "La somme est paire" et B l'événement "Le premier dé donne 1".
  On a alors $\mathbb{P}(A) = \ldots$ et $\mathbb{P}(B) = \ldots$. Ainsi,
  $$\mathbb{P}(A \cap B) = \mathbb{P}\left(\{(1,1),(1,3),(1,5)\}\right) = \frac{1}{12} = \mathbb{P}(A)\mathbb{P}(B).$$
  Les événements A et B sont donc indépendants.

- On tire deux boules dans une urne .... remise : les deux numéros obtenus sont indépendants.

- On tire deux boules dans une urne .... remise : les deux numéros obtenus ne sont pas indépendants.

CHAPITRE 2

Variables aléatoires discrètes

Sommaire

Dans toute la suite de ce cours, nous allons parler de variable aléatoire. Une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans F est une fonction $X : \Omega \longrightarrow F$ qui vérifie une certaine condition de mesurabilité.
Lors d'une expérience aléatoire, une variable aléatoire va représenter une mesure que l'on fait pendant cette expérience. Par exemple, si on lance n dés, on peut vouloir mesurer la somme totale obtenue. La variable aléatoire représentant cela sera

$$X : \begin{array}{ccc} [\![1,6]\!]^n & \longrightarrow & [\![n,6n]\!] \\ (i_1,\ldots,i_n) & \longmapsto & i_1+\ldots+i_n \end{array}.$$

2.1 Définition et notions associées

Définition 2.1.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Une **variable aléatoire discrète** sur cet espace est une fonction $X : \Omega \longrightarrow E$ avec <u>E fini ou dénombrable</u>, telle que pour tout $x \in E$, $X^{-1}(\{x\}) \in \mathcal{A}$.

Remarque 2.1.2.

- Chaque réalisation de l'expérience aléatoire donne un nouveau résultat $\omega \in \Omega$ et donc une nouvelle valeur $X(\omega)$. Le but est de voir avec quelle probabilité X prend certaines valeurs.

- La condition $\forall x \in E,\ X^{-1}(\{x\}) \in \mathcal{A}$ est une condition de mesurabilité que l'on ne vérifiera pas dans les exemples traités.

Exemple 2.1.3. On lance deux dés équilibrés. L'univers est $\Omega = [\![1,6]\!]^2$, la tribu est $\mathcal{A} = \mathcal{P}(\Omega)$ et la probabilité $\mathbb{P}$ est la probabilité uniforme sur $\mathcal{A}$.

Soit $X : \begin{array}{ccc} \Omega & \longrightarrow & [\![2,12]\!] \\ (i,j) & \longmapsto & i+j \end{array}$.

On a $\mathbb{P}(X=2) =$ .... De la même façon, on trouve $\mathbb{P}(X=3) =$ ....., $\mathbb{P}(X=4) =$ ....., $\mathbb{P}(X=5) =$ ..., $\mathbb{P}(X=6) =$ ....., $\mathbb{P}(X=7) =$ ..., $\mathbb{P}(X=8) =$ ....., $\mathbb{P}(X=9) =$ ..., $\mathbb{P}(X=10) =$ ....., $\mathbb{P}(X=11) =$ ..... et $\mathbb{P}(X=12) =$ ......

Remarque 2.1.4. On a vu que $\mathbb{P}$ est une fonction prenant en entrée un sous-ensemble de Ω. Or on a écrit au-dessus "$\mathbb{P}(X=2)$". Ceci est une notation : $\mathbb{P}(X=2)$ est en fait $\mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) = 2\}\big)$.

2.1.1 Loi

Définition 2.1.5. La **loi** d'une variable aléatoire discrète X est la donnée de toutes les probabilités $\mathbb{P}(X=x)$ pour tout $x \in X(\Omega)$.

Remarque 2.1.6. On peut représenter la loi d'une variable aléatoire discrète sous forme d'un tableau. Avec notre exemple précédent :

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X=x)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

On peut remarquer que la somme des valeurs dans la deuxième ligne donne 1.

Proposition 2.1.7. *Si X est une variable aléatoire discrète sur Ω, on a toujours*

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

2.1.2 Fonction de répartition

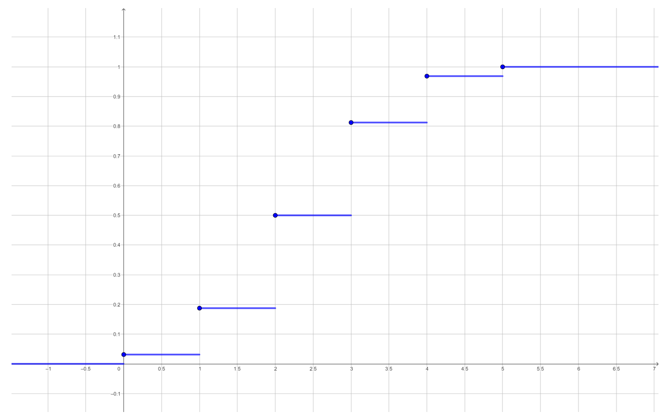
Définition 2.1.8. Soit X une variable aléatoire discrète à valeurs réelles. La **fonction de répartition** de X est la fonction $F_X : \begin{array}{ccc} \mathbb{R} & \longrightarrow & [0,1] \\ t & \longmapsto & \mathbb{P}(X \le t) \end{array}$.

Proposition 2.1.9. *Si X est une variable aléatoire discrète à valeurs réelles, on peut noter $X(\Omega) = \{\ldots,,x_{-2},x_{-1},x_0,x_1,x_2,\ldots\}$ avec $x_i < x_{i+1}$. La fonction de répartition de X est alors constante sur chaque intervalle $[x_i,x_{i+1}[$ et présente des discontinuités en chaque x_i. La hauteur du saut en x_i est alors* ....

Exemple 2.1.10. On lance une pièce équilibrée 5 fois, $\Omega = \{P, F\}^5$. Soit X la variable aléatoire donnant le nombre de Pile obtenus. X est discrète car $X(\Omega) =$ . . . .
On a ainsi $\mathbb{P}(X = 0) =$ . . . ., $\mathbb{P}(X = 1) =$ . . . . ., $\mathbb{P}(X = 2) =$ . . . ., $\mathbb{P}(X = 3) =$ . . . . ., $\mathbb{P}(X = 4) =$ . . . . . et $\mathbb{P}(X = 5) =$ . . . . ..
Soit F_X la fonction de répartition de X, calculons par exemple $F_X(2,3)$:

$F_X(2,3) =$ . . . .

$$F_X : t \longmapsto \begin{cases} 0 & \text{si } t < 0 \\ \frac{1}{32} & \text{si } 0 \leq t < 1 \\ \frac{3}{16} & \text{si } 1 \leq t < 2 \\ \frac{1}{2} & \text{si } 2 \leq t < 3 \\ \frac{13}{16} & \text{si } 3 \leq t < 4 \\ \frac{31}{32} & \text{si } 4 \leq t < 5 \\ 1 & \text{si } t \geq 5 \end{cases} .$$



Proposition 2.1.11. *Soit X une variable aléatoire discrète à valeurs réelles, soient $a > b$ deux réels. Alors*

$$\mathbb{P}(a \leq X < b) = F_X(b) - F_X(a).$$

Les inégalités strictes ou larges dans le terme de gauche sont importantes !

2.1.3 Espérance

L'espérance d'une variable aléatoire est un concept qui correspond à la notion intuitive de moyenne.

Définition 2.1.12. Soit X une variable aléatoire discrète à valeurs dans $\{x_1, x_2, x_3, \ldots\}$. On définit l'espérance de X par

$$\mathbb{E}[X] = \sum_{k=1}^{+\infty} x_k \mathbb{P}(X = x_k),$$

si cette série converge. Si cette série ne converge pas, X n'admet pas d'espérance.

Remarque 2.1.13.

- En particulier, si $X : \Omega \longrightarrow \mathbb{N}$ alors $\mathbb{E}[X] = \sum_{k=1}^{+\infty} k\mathbb{P}(X = k)$.

- Si $X(\Omega) = \{x_1, \ldots, x_n\}$ est fini, alors X admet une espérance et $\mathbb{E}[X] = \sum_{k=1}^{n} x_k \mathbb{P}(X = x_k)$.

Exemple 2.1.14.

- Reprenons l'exemple du lancer de deux dés avec X qui calcule la somme. On a

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

  En moyenne, la somme de deux dés vaut . ..

- Soit X une variable aléatoire telle que $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$. Alors

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

Remarque 2.1.15. Il existe des variables aléatoires discrètes qui n'admettent pas d'espérance. Par exemple, soit X telle que pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = \frac{\dots}{\dots . k^2}$. Alors la série $\sum_{k \geq 1} k\mathbb{P}(X = k)$ diverge donc X n'a pas d'espérance.

L'espérance n'est pas la même chose que la valeur médiane ! Une médiane m de X est une valeur telle que $\mathbb{P}(X \leq m) \geq \frac{1}{2}$ et $\mathbb{P}(X \geq m) \geq \frac{1}{2}$. Par exemple, si $\mathbb{P}(X = -4) = \mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \frac{1}{3}$ alors 1 est une médiane de X mais $\mathbb{E}[X] = -\frac{1}{3}$.

Définition 2.1.16. On dit que X est **centrée** si $\mathbb{E}[X] = 0$.

Proposition 2.1.17. *Soit X une variable aléatoire discrète avec $X(\Omega) = \{x_1, x_2, \dots\}$, soit $f : \mathbb{R} \longrightarrow \mathbb{R}$.*
Alors $f(X)$ est une variable aléatoire discrète et

$$\mathbb{E}\big[f(X)\big] = \sum_{k=1}^{+\infty} f(x_k)\mathbb{P}(X = x_k).$$

Exemple 2.1.18. Soit X une variable aléatoire de loi :

x_k	-2	-1	0	2
$\mathbb{P}(X = x_k)$	0.25	0.1	0.2	0.55

Alors $\mathbb{E}\big[X^2\big] = \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$
Il est aussi possible de déterminer d'abord la loi de X^2 puis de calculer son espérance.

Comme on vient de le voir, en général, $\mathbb{E}\big[X^2\big] \neq \mathbb{E}[X]^2$!

Proposition 2.1.19. *L'espérance est linéaire : si X et Y sont deux variables aléatoires et a, b sont des réels alors*

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

De plus, si $a \in \mathbb{R}$ alors $\mathbb{E}[a] = a$.

2.1.4 Variance

La variance mesure la dispersion d'une variable aléatoire autour de son espérance.

Définition 2.1.20. Soit X une variable aléatoire discrète à valeurs réelles. La **variance** de X est définie par

$$\mathrm{Var}(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big],$$

si cette quantité existe et est finie.

Proposition 2.1.21. *Si la variance de X existe et si X^2 admet une espérance, on a la formule*

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

Définition 2.1.22. Si X admet une variance, l'**écart-type** de X est défini par

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

Exemple 2.1.23. Reprenons l'exemple du lancer de deux dés et de la variable aléatoire X qui calcule la somme obtenue. On a :

$$\mathbb{E}\big[X^2\big] = \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..$$

On en déduit que $\mathrm{Var}(X) = \dots.$ et l'écart-type est $\sigma(X) = \dots\dots\dots\dots\dots\dots..$

Proposition 2.1.24. *Soit X une variable aléatoire admettant une variance, soient a, b deux réels. Alors :*

- $\mathrm{Var}(a) = 0$ *et* $\mathrm{Var}(X + b) = \mathrm{Var}(X)$ *;*
- $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$ *et* $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

Définition 2.1.25. Une variable aléatoire discrète est dite **centrée réduite** si $\mathbb{E}[X] = 0$ et $\mathrm{Var}(X) = 1$.
Si X admet une variance, la **centrée réduite** de X est la variable aléatoire $\dfrac{X - \mathbb{E}[X]}{\sigma(X)}$.

2.1.5 Moments

Les moments sont une généralisation des notions d'espérance et de variance.

Définition 2.1.26. Soit X une variable aléatoire discrète à valeurs réelles, soit $k \in \mathbb{N}^*$.
Le **moment d'ordre** k de X est (s'il existe)

$$M_k(X) = \mathbb{E}\big[X^k\big].$$

Le **moment centré d'ordre** k de X est (s'il existe)

$$M_k^{\mathrm{c}}(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^k\big].$$

Remarque 2.1.27.

- Le moment d'ordre 1 est donc l'espérance et le moment centré d'ordre 2 est la variance.
- Pour la culture : le coefficient d'asymétrie d'une variable aléatoire X est $\gamma = \dfrac{M_3^{\mathrm{c}}(X)}{\sigma(X)^3}$ et le coefficient d'aplanissement de X est $\kappa = \dfrac{M_4^{\mathrm{c}}(X)}{\sigma(X)^4}$.

2.1.6 Fonction génératrice des moments et fonction caractéristique

Définition 2.1.28. Soit X une variable aléatoire discrète à valeurs réelles. La **fonction génératrice des moments** de X est la fonction

$$G_X : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \mathbb{E}\left[e^{tX}\right] \end{array}.$$

Proposition 2.1.29. *Si G_X est k fois dérivable en 0, on a $G_X^{(k)}(0) = M_k(X)$.*

Définition 2.1.30. Soit X une variable aléatoire discrète à valeurs réelles. La **fonction caractéristique** de X est la fonction

$$\varphi_X : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{C} \\ t & \longmapsto & \mathbb{E}\left[e^{itX}\right] \end{array}.$$

Proposition 2.1.31. *Ces deux fonctions caractérisent la loi de X, c'est-à-dire que si deux variables alétoires dicrètes ont la même fonction caractéristique ou la même fonction génératrice des moments alors elles ont la même loi.*

2.2 Lois discrètes usuelles

2.2.1 Loi de Bernoulli $b(p)$

Une variable aléatoire X suit la loi de Bernoulli $b(p)$ (ce qu'on note $X \sim b(p)$) avec $p \in [0,1]$ si :

$$\mathbb{P}(X = 1) = p \quad \text{et} \quad \mathbb{P}(X = 0) = 1 - p.$$

Ceci représente une épreuve binaire : seuls deux résultats sont possibles. Souvent, 1 représente un succès et 0 un échec.

Proposition 2.2.1. *Supposons que $X \sim b(p)$. Alors $\mathbb{E}[X] = p$ et $\mathrm{Var}(X) = p(1-p)$.*
On a aussi $G_X(t) = pe^t + (1-p)$ et $\varphi_X(t) = pe^{it} + (1-p)$.

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

Une variable aléatoire X suit la loi binomiale $\mathcal{B}(n,p)$ avec $n \in \mathbb{N}$ et $p \in [0,1]$ si :

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

X compte alors le nombre de succès lors de n épreuves de Bernoulli $b(p)$ indépendantes.

Proposition 2.2.2. *Supposons que $X \sim \mathcal{B}(n,p)$. Alors $\mathbb{E}[X] = np$ et $\mathrm{Var}(X) = np(1-p)$.*
On a aussi $G_X(t) = \left(pe^t + (1-p)\right)^n$ et $\varphi_X(t) = \left(pe^{it} + (1-p)\right)$.

2.2.3 Loi uniforme discrète

Une variable aléatoire suit une loi uniforme sur $[\![1, n]\!]$ (on note parfois $X \sim \mathcal{U}([\![1, n]\!])$ si :

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

La loi uniforme peut avoir pour support un autre ensemble, par exemple $[\![-n, n]\!]$.

2.2.4 Loi géométrique $\mathcal{G}(p)$

Une variable aléatoire X suit la loi géométrique $\mathcal{G}(p)$ avec $p \in]0, 1[$ si :

$$\forall\, k \in \mathbb{N}^*,\ \mathbb{P}(X = k) = p(1-p)^{k-1}.$$

X donne alors le rang du premier succès lors de répétitions d'épreuves de Bernoulli $b(p)$ successives et indépendantes. Elle correspond au temps d'attente avant le premier succès.

Proposition 2.2.3. *Supposons que* $X \sim \mathcal{G}(p)$*. Alors* $\mathbb{E}[X] = \frac{1}{p}$ *et* $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

Remarque 2.2.4. La loi géométrique n'a pas de mémoire, c'est-à-dire que pour tous k, $l \in \mathbb{N}^*$, on a

$$\mathbb{P}\big(X > k + l \mid X > k\big) = \mathbb{P}(X > l).$$

C'est la seule discrète qui vérifie cela.

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

Une variable aléatoire X suit la loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda \geq 0$ si :

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

La loi de Poisson est aussi appelée loi des événements rares.

Proposition 2.2.5. *Supposons que* $X \sim \mathcal{P}(\lambda)$*. Alors* $\mathbb{E}[X] = \lambda$ *et* $\mathrm{Var}(X) = \lambda$.

2.2.6 Loi hypergéométrique $\mathcal{HG}(N, m, n)$

Une variable aléatoire X suit la loi hypergéométrique $\mathcal{HG}(N, m, n)$ avec $N \in \mathbb{N}^*$, $m \in [\![0, N]\!]$ et $n \in \mathbb{N}$ si :

$$\forall\, k \in [\![0, n]\!],\ \mathbb{P}(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}.$$

Exemple 2.2.6. Une urne contient N boules, dont m blanches. On tire n boules, successivement et sans remise. On note X le nombre de boules blanches obtenues. La variable aléatoire X suit alors la loi hypergéométrique $\mathcal{HG}(N, m, n)$.

CHAPITRE 3

Variables aléatoires à densité

Sommaire

Avant de parler de densité, précisons la notion de variable aléatoire **réelle** : Une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ est une fonction $X : \Omega \longrightarrow \mathbb{R}$ qui vérifie :

$$\forall t \in \mathbb{R},\ X^{-1}(]-\infty; t]) \in \mathcal{A}.$$

Cette dernière condition est une condition de mesurabilité que l'on ne vérifiera jamais en pratique.

3.1 Définition et notions associées

Définition 3.1.1 (Variable aléatoire continue)**.** Une variable aléatoire réelle X définie sur Ω est dite **continue** si $X(\Omega)$ est infini non dénombrable.

Remarque 3.1.2. Très souvent, l'image de X sera un intervalle de $\mathbb{R}$.

Exemple 3.1.3. On voudrait une variable aléatoire X à valeurs dans $[0;1]$ qui donne la même probabilité à chaque point de $[0;1]$. Si cette probabilité est $p > 0$, alors on a $\mathbb{P}(X \in [0;1]) = +\infty$ ce qui est absurde.
Il faut donc que pour tout $x \in [0;1]$, $\mathbb{P}(X = x) = 0$. On ne peut donc pas définir la loi de X en donnant la probabilité de chaque résultat.
Par contre, on peut utiliser la fonction de répartition F_X : si la probabilité est uniforme sur $[0;1]$, alors pour $0 \leq a < b \leq 1$, la probabilité de tomber dans $[a;b]$ ne doit dépendre que de la longueur $b - a$. Ainsi, on dira que pour tous $0 \leq a < b \leq 1$, $\mathbb{P}(a \leq X \leq b) = b - a$. On obtient alors bien $\mathbb{P}(0 \leq X \leq 1) = 1$.
Rappelons que dans le cas discret, $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$. Ici, puisque $\mathbb{P}(X = a) = 0$, on obtient $F_X(b) = b - a + F_X(a)$. Si $a = 0$, on obtient $F_X(b) = b$ si $0 \leq b \leq 1$. Ainsi notre variable aléatoire uniforme sur $[0;1]$ a pour fonction de répartition

$$F_X : x \longmapsto \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}.$$

Remarquons que F_X est la primitive de limite 1 en $+\infty$ de la fonction

$$f_X : x \longmapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{si } x > 1 \end{cases}.$$

Cette fonction sera la densité de X.

3.1.1 Densité

Définition 3.1.4 (Densité)**.** On dit qu'une variable aléatoire réelle est **à densité** s'il existe une fonction $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ positive et intégrable telle que

$$\forall\, a \leq b \in \mathbb{R},\ \mathbb{P}\left(a \leq X \leq b\right) = \int_a^b f_X(t)\,\mathrm{d}t.$$

La fonction f_X est alors la **densité** de X.

Proposition 3.1.5. *Si X admet pour densité f_X, alors sa fonction de répartition $t \longmapsto \mathbb{P}(X \leq t)$ est donnée par*

$$F_X(x) = \int_{-\infty}^x f_X(t)\,\mathrm{d}t.$$

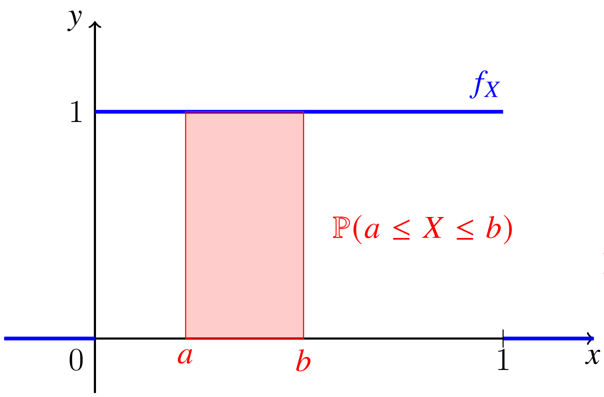
En particulier, si la densité f_X est continue alors F_X est dérivable et $F_X' = f_X$.

Remarque 3.1.6.

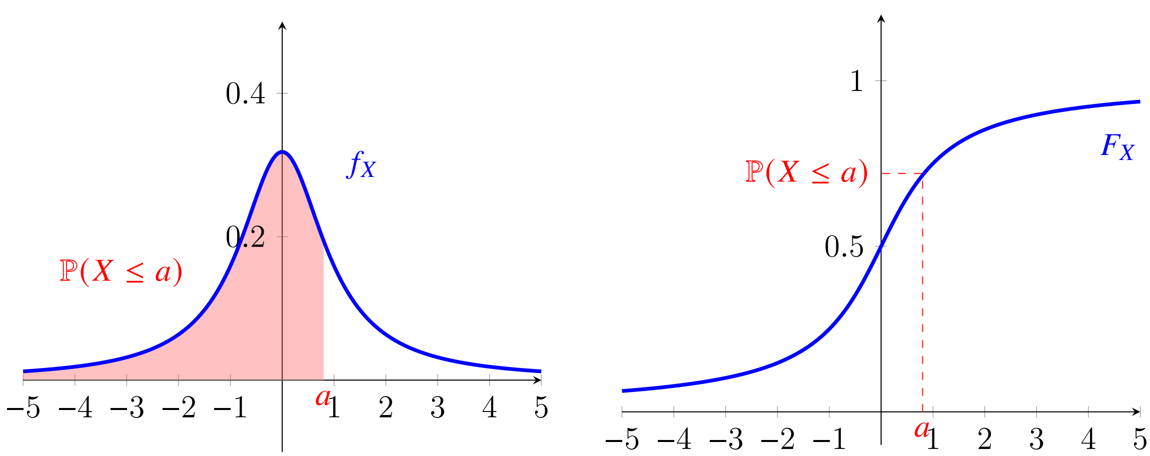
- Si la densité est seulement continue par morceaux, F_X est dérivable par morceaux et en dérivant sur chaque sous-intervalle, on retrouve f_X.
- Une fonction de répartition de variable aléatoire à densité est toujours continue.
- Si f est la densité d'une variable aléatoire réelle, on doit avoir
$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$$

Exemple 3.1.7.

- On a vu l'exemple de la loi uniforme sur $[0;1]$. Sa densité est f_X définie par $f_X(t) = 1$ si $0 \le t \le 1$ et $f_X(t) = 0$ sinon. La fonction de répartition correspondante est à droite.



- On considère X dont la densité f_X est donnée ci-dessous. La fonction de répartition F_X est alors la primitive de f_X qui tend vers 1 en $+\infty$.



La densité ne donne pas directement la probabilité de prendre telle ou telle valeur : $f_X(x)$ n'est PAS la probabilité $\mathbb{P}(X = x)$, il faut se souvenir que cette probabilité est rigoureusement nulle. Par contre, de façon pratique, les domaines où la densité est élevée sont les domaines où X a le plus de chances de tomber.

Proposition 3.1.8. *Soit X une variable aléatoire réelle à densité. Sa fonction de répartition F_X est continue, croissante et vérifie*
$$\lim_{x\to-\infty} F_X(x) = 0 \quad \textit{et} \quad \lim_{x\to+\infty} F_X(x) = 1.$$

Remarque 3.1.9. Si X est à densité, on a pour tout $a \in \mathbb{R}$,

$$\mathbb{P}(X = a) = 0, \quad \mathbb{P}(X < a) = \mathbb{P}(X \leq a) \quad \text{et} \quad \mathbb{P}(X > a) = \mathbb{P}(X \geq a).$$

Définition 3.1.10. La **loi** d'une variable aléatoire réelle à densité est la *mesure* $\mathbb{P}_X$ sur $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ donnée par

$$\forall\, B \in \mathcal{B}(\mathbb{R}),\ \mathbb{P}_X(B) = \int_B f_X(t)\,\mathrm{d}t.$$

$\mathcal{B}(\mathbb{R})$ est ici la tribu *borélienne* de $\mathbb{R}$, c'est la plus petite tribu sur $\mathbb{R}$ qui contienne tous les intervalles.

En pratique : déterminer la loi, c'est déterminer la densité ou la fonction de répartition.

3.1.2 Espérance

Définition 3.1.11. Si X est une variable aléatoire réelle à densité f_X alors l'**espérance** de X est

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t,$$

lorsque cette intégrale existe.

Remarque 3.1.12. Il faut ici faire le lien avec la formule $\displaystyle\sum_{k=-\infty}^{+\infty} k\mathbb{P}(X = k)$ pour une variable aléatoire réelle discrète.

Définition 3.1.13. Une variable aléatoire X est dite **centrée** si $\mathbb{E}[X] = 0$.

Exemple 3.1.14.

- Soit X suivant la loi uniforme sur $[0; 1]$, son espérance est

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

- Supposons que X ait pour densité $f_X(t) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ (on rappelle la valeur de l'intégrale de Gauss : $\displaystyle\int_{-\infty}^{+\infty} e^{-t^2}\,\mathrm{d}t = \sqrt{\pi}$). Son espérance est

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

  Cette variable aléatoire est donc centrée.

- Supposons que X ait pour densité $f_X(t) = \dfrac{1}{\pi(1+t^2)}$ sur $\mathbb{R}$. La fonction $t \longmapsto \dfrac{t}{1+t^2}$ est positive et équivalente à $\dfrac{1}{t}$ en $+\infty$, l'intégrale définissant l'espérance n'est donc pas convergente. Cette variable aléatoire n'admet pas d'espérance.

Proposition 3.1.15 (Lemme de transfert)**.** *Soit X une variable aléatoire réelle à densité, soit $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$. On suppose que X admet une espérance. Alors la variable aléatoire $\varphi(X)$ admet aussi une espérance et*

$$\mathbb{E}\left[\varphi(X)\right] = \int_{-\infty}^{+\infty} \varphi(t) f_X(t)\,\mathrm{d}t.$$

Remarque 3.1.16. Le cas le plus utilisé en pratique est $\varphi(t) = t^2$. On a ainsi

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{+\infty} t^2 f_X(t)\,\mathrm{d}t.$$

Exemple 3.1.17.

- Reprenons la loi uniforme sur $[0;1]$, l'espérance de X^2 est :

  $\mathbb{E}\left[X^2\right] = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.$

- Avec la loi uniforme sur $[0;1]$, on peut calculer par exemple l'espérance de e^X :

  $\mathbb{E}\left[e^X\right] = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.$

- Soit X de densité $f_X(t) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$. L'espérance de X^2 est :

  $\mathbb{E}\left[X^2\right] = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

Proposition 3.1.18. *L'espérance est linéaire :* $\mathbb{E}\left[aX + bY\right] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.
L'espérance est croissante : si $X \leq Y$ *alors* $\mathbb{E}[X] \leq \mathbb{E}[Y]$. *En particulier si* $X \geq 0$, *alors* $\mathbb{E}[X] \geq 0$.

3.1.3 Variance

Définition 3.1.19. Soit X une variable aléatoire réelle à densité admettant une espérance. La **variance** de X est définie par

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \int_{-\infty}^{+\infty} (t - \mathbb{E}[X])^2 f_X(t)\,\mathrm{d}t,\end{aligned}$$

si cette espérance existe.

Définition 3.1.20. On dit qu'une variable aléatoire X est **centrée réduite** si $\mathbb{E}[X] = 0$ et $\mathrm{Var}(X) = 1$.

Proposition 3.1.21. *Supposons que* X *admet une variance alors*

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \\ &= \int_{-\infty}^{+\infty} t^2 f_X(t)\,\mathrm{d}t - \left(\int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t\right)^2.\end{aligned}$$

Exemple 3.1.22.

- Loi uniforme sur $[0;1]$: on a vu que $\mathbb{E}\left[X^2\right] = \dfrac{1}{3}$ et donc

  $$\mathrm{Var}(X) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

- Densité $f_X(t) = e^{-\frac{t^2}{2}}$: on obtient

  $$\mathrm{Var}(X) = 1 - 0^2 = 1.$$

  Cette loi est donc centrée réduite.

Proposition 3.1.23. *Soit* X *une variable aléatoire réelle admettant une variance. Alors pour tous* $a, b \in \mathbb{R}$, $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

3.1.4 Quantiles

La notion de quantile correspond à une inversion de la fonction de répartition : étant donnée une probabilité $\alpha \in [0;1]$, on cherche le réel q_α tel que $F_X(q_\alpha) = \alpha$.

Définition 3.1.24. Soit X une variable aléatoire, supposons que sa fonction de répartition F_X est strictement croissante.
Soit $\alpha \in [0;1]$, le **quantile d'ordre** α de X est alors le réel q_α tel que $\mathbb{P}(X \leq q_\alpha) = \alpha$.

Définition 3.1.25. Le quantile d'ordre $\frac{1}{2}$ est appelé **médiane**.
Le quantile d'ordre $\frac{1}{4}$ est appelé **premier quartile** et le quantile d'ordre $\frac{3}{4}$ est appelé **troisième quartile**.
Le quantile d'ordre $\frac{k}{10}$ est appelé $k^{\text{ème}}$ **décile**.

L'idée derrière les quartiles, déciles ou centiles est de découper l'effectif des résultats possibles en sections d'égale probabilité.

La médiane n'a a priori rien à voir avec l'espérance ! En règle générale, l'espérance est beaucoup plus sensible aux valeurs extrêmes que la médiane.

3.1.5 Moments

Définition 3.1.26. Soit X une variable aléatoire réelle à densité, soit $k \in \mathbb{N}^*$.
Le **moment d'ordre** k de X est (s'il existe)

$$M_k(X) = \mathbb{E}\left[X^k\right].$$

Le **moment centré d'ordre** k de X est (s'il existe)

$$M_k^{\mathrm{c}}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^k\right].$$

3.1.6 Fonction génératrice des moments et fonction caractéristique

Définition 3.1.27. Soit X une variable aléatoire réelle à densité, la **fonction génératrice des moments** de X est

$$G_X : t \longmapsto \mathbb{E}\left[e^{tX}\right] = \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,\mathrm{d}x,$$

pour tout t où cette espérance existe.

Proposition 3.1.28. *Si G_X est k fois dérivable en 0 alors $G_X^{(k)}(0) = M_k(X)$.*

Définition 3.1.29. Soit X une variable aléatoire réelle à densité, la **fonction caractéristique** de X est

$$\varphi_X : \begin{array}{ccl} \mathbb{R} & \longrightarrow & \mathbb{C} \\ t & \longmapsto & \mathbb{E}\left[e^{itX}\right] = \displaystyle\int_{-\infty}^{+\infty} e^{itx} f_X(x)\,\mathrm{d}x \end{array}.$$

Proposition 3.1.30. *La fonction génératrice des moments et la fonction caractéristique caractérisent la loi de X. Dit autrement, si deux variables aléatoires ont la même fonction génératrice ou la même fonction caractéristique, alors elles ont la même loi.*

3.2 Lois à densité usuelles

3.2.1 Loi uniforme sur un intervalle

Soient $a < b \in \mathbb{R}$. Une variable aléatoire suit une loi uniforme sur $[a; b]$ (on note parfois $X \sim \mathcal{U}([a; b])$) si sa densité est

$$f_X : t \longmapsto \frac{1}{b-a}\mathbf{1}_{[\mathbf{a};\mathbf{b}]}(\mathbf{t}).$$

Sa fonction de répartition est alors $F_X : t \longmapsto \begin{cases} 0 & \text{si } t \leq a \\ \dfrac{t-a}{b-a} & \text{si } a < t < b \\ 1 & \text{si } t \geq b \end{cases}$.

Proposition 3.2.1. *Supposons que $X \sim \mathcal{U}([a; b])$. Alors $\mathbb{E}[X] = \dfrac{a+b}{2}$ et* $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$.

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

Une variable aléatoire suit la loi exponentielle $\mathcal{E}(\lambda)$ avec $\lambda > 0$ si sa densité est

$$f_X : t \longmapsto \begin{cases} 0 & \text{si } t < 0 \\ \lambda e^{-\lambda t} & \text{si } t \geq 0 \end{cases}.$$

Sa fonction de répartition est alors $F_X : t \longmapsto 1 - e^{-\lambda t}$.

Proposition 3.2.2. *Supposons que $X \sim \mathcal{E}(\lambda)$. Alors $\mathbb{E}[X] = \dfrac{1}{\lambda}$ et* $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$.

Remarque 3.2.3. La loi exponentielle modélise habituellement la durée de vie d'un phénomène continu et sans mémoire (ou sans usure). La loi exponentielle est en effet sans mémoire, dans le sens où, si $X \sim \mathcal{E}(\lambda)$ et si $t, s \in \mathbb{R}_+^*$, alors

$$\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s).$$

C'est en fait l'unique loi à densité sans mémoire.

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

Voir section suivante.

3.2.4 Loi de Cauchy $\mathcal{C}(a, b)$

Une variable aléatoire X suit la loi de Cauchy $\mathcal{C}(a, b)$ avec $a > 0$ et $b \in \mathbb{R}$, si sa densité est

$$f_X : t \longmapsto \frac{a}{\pi(a^2 + (t-b)^2)}.$$

Sa fonction de répartition est alors $F_X : t \longmapsto \dfrac{1}{2} + \dfrac{1}{\pi}\arctan\left(\dfrac{t-b}{a}\right)$.

Proposition 3.2.4. *Supposons que $X \sim \mathcal{C}(a, b)$. Alors $\mathbb{E}[X]$ n'existe pas et* $\mathrm{Var}(X)$ *non plus.*

3.3 La loi normale

La loi normale, ou loi de Laplace-Gauss, est sans doute la loi de probabilité la plus connue. Sa densité présente une forme caractéristique de "courbe en cloche".

3.3.1 Loi normale standard

Définition 3.3.1. Une variable aléatoire X suit la **loi normale centrée réduite** ou **loi normale standard**, notée $\mathcal{N}(0,1)$, si sa densité est

$$f_X \,:\, t \longmapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Proposition 3.3.2. *Si $X \sim \mathcal{N}(0,1)$ alors $\mathbb{E}[X] = 0$ et $\mathrm{Var}(X) = 1$.*

Remarque 3.3.3. Il n'existe pas d'expression pour la fonction de répartition de la loi normale centrée réduite autre que

$$F_X \,:\, t \longmapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{s^2}{2}} \,\mathrm{d}s.$$

Proposition 3.3.4. *Soit $X \sim \mathcal{N}(0,1)$. Soit $\alpha > 0$, on a*

$$\mathbb{P}(-\alpha \le X \le \alpha) = 2F_X(\alpha) - 1.$$

Ceci découle de la parité de la densité de X.

Exemple 3.3.5. Deux cas particuliers qui seront très utiles en statistiques :

- $\mathbb{P}(-1,96 \le X \le 1,96) \simeq 0,95$.
- $\mathbb{P}(-2,6 \le X \le 2,6) \simeq 0,99$.

Proposition 3.3.6. *Supposons que $X \sim \mathcal{N}(0,1)$. Alors*

$$G_X \,:\, t \longmapsto e^{\frac{t^2}{2}} \quad \text{et} \quad \varphi_X \,:\, t \longmapsto e^{-\frac{t^2}{2}}.$$

3.3.2 Loi normale généralisée

Définition 3.3.7. Une variable aléatoire X suit la **loi normale** $\mathcal{N}(\mu,\sigma^2)$, avec $\mu \in \mathbb{R}$ et $\sigma > 0$ si sa densité est

$$f_X \,:\, t \longmapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right).$$

Proposition 3.3.8. *Supposons que $X \sim \mathcal{N}(\mu,\sigma^2)$. Alors $\mathbb{E}[X] = \mu$ et $\mathrm{Var}(X) = \sigma^2$.*

Remarque 3.3.9. Il est très important de remarquer que si $X \sim \mathcal{N}(\mu,\sigma^2)$ alors $\dfrac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.

On trouve parfois une autre convention de notation pour la loi normale : $\mathcal{N}(\mu,\sigma)$ où μ est toujours la moyenne mais σ est l'écart-type.
Dans tout ce cours et en évaluation, quand on écrit $\mathcal{N}(a,b)$, b est la variance.
Sur `R`, l'autre convention est utilisée : pour la densité de $\mathcal{N}(2,25)$, on écrira `dnorm(x,2,5)`.

Remarque 3.3.10. La densité de $\mathcal{N}(\mu,\sigma^2)$ présente toujours une courbe en cloche, mais centrée en μ et aplatie ou resserrée selon la valeur de σ^2.

3.4 Autres lois à densité dérivées de la loi normale

Les trois lois à densité présentées dans cette section nous seront utiles en statistiques.

3.4.1 Loi du χ^2

Définition 3.4.1. On considère n variables aléatoires $X_1, \ldots, X_n$ indépendantes et suivant toutes la loi normale standard. Alors la variable aléatoire $S^2 = X_1^2 + \ldots + X_n^2$ suit une loi appelée **loi du χ^2 à n degrés de liberté** et notée $\chi^2(n)$.

Remarque 3.4.2. Si $X \sim \chi^2(n)$, sa densité est donnée par $f_X(t) = \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} t^{\frac{n}{2}-1} e^{-\frac{t^2}{2}}$ si $t \geq 0$ et $f_X(t) = 0$ si $t < 0$.
On rappelle que la fonction Gamma d'Euler est définie par $\Gamma(x) = \displaystyle\int_0^{+\infty} t^{x-1} e^{-t}\,\mathrm{d}t$.

Proposition 3.4.3. *Si $X \sim \chi^2(n)$, alors $\mathbb{E}[X] = n$ et $\mathrm{Var}(X) = 2n$.*

3.4.2 Loi de Student

Définition 3.4.4. Si $X \sim \mathcal{N}(0,1)$ et si $Y \sim \chi^2(n)$ alors la variable aléatoire $T = \sqrt{n}\dfrac{X}{\sqrt{Y}}$ suit une loi appelée **loi de Student à n degrés de liberté** et notée $\mathcal{T}(n)$.

Remarque 3.4.5. Si $X \sim \mathcal{T}(n)$ alors sa densité est $f_X : t \longmapsto \dfrac{1}{\sqrt{n\pi}} \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \dfrac{t^2}{n}\right)^{-\frac{n+1}{2}}$.

Proposition 3.4.6. *Si $X \sim \mathcal{T}(n)$ alors $\mathbb{E}[X] = 0$ dès que $n \geq 2$ et $\mathrm{Var}(X) = \dfrac{n}{n-2}$ dès que $n \geq 3$.*

3.4.3 Loi de Fisher-Snedecor

Définition 3.4.7. Si $X_1 \sim \chi^2(n_1)$ et $X_2 \sim \chi^2(n_2)$ avec $n_1, n_2 \in \mathbb{N}$, alors la variable aléatoire $F = \dfrac{n_2 X_1}{n_1 X_2}$ suit une loi appelée **loi de Fisher-Snedecor de paramètres n_1 et n_2** et notée $\mathcal{F}(n_1, n_2)$.