

## Notions de base de probabilité

La première partie de ce poly explique brièvement pourquoi on modélise les probabilités de cette façon. La deuxième partie rappelle les définitions de base des probabilités : espace probabilisé, variable aléatoire, loi, espérance, variance. Son objectif est d'éviter de perdre du temps en présentant des définitions soporifiques en cours, tout en s'assurant que tout le monde sait de quoi on parle. En troisième partie sont présentées brièvement les lois "classiques". On reviendra dessus en cours.

### 1 Modélisation des probabilités

#### 1.1 Idée intuitive des probabilités et ses limites

Les premières formalisations de la notion de hasard, au XVIIème siècle, répondaient pour l'essentiel à diverses questions issues de la théorie des jeux.

**Exemples.**

Si on lance un dé, on a une chance sur 6 de tomber sur 1, autrement dit la probabilité d'obtenir 1 est  $1/6$ . Si on tire 5 cartes dans un jeu de 52 cartes, quelle est la probabilité d'avoir les 4 rois ? Le nombre de tirages possibles est  $\binom{52}{5}$ , le nombre de tirages favorable est  $\binom{48}{1} = 48$  (il faut choisir une 5ème carte parmi les  $52 - 4$  cartes qui ne sont pas des rois). Donc la probabilité cherchée est  $48/\binom{52}{5}$ .

Quand le résultat de l'expérience ne peut prendre qu'un nombre fini de valeurs, on est essentiellement ramené à faire du dénombrement. Il en va autrement quand le résultat de l'expérience peut prendre un nombre non dénombrable de valeurs.

**Exemple.** Soit  $S$  une sphère représentant la surface de la terre. On choisit un point  $x$  au hasard sur  $S$  (l'impact d'une météorite, par exemple). Quelle est la probabilité que ce point soit dans l'hémisphère nord ? Intuitivement, on a autant de chance de tomber dans l'hémisphère nord que dans l'hémisphère sud puisqu'ils ont la même surface, donc la probabilité est  $1/2$ . En suivant la même idée, la probabilité que  $x$  soit sur la terre ferme est

$$\frac{\text{surface des terres émergées}}{\text{surface totale de la planète}}$$

La surface étant une mesure, on se trouve donc dans le cadre de la théorie de la mesure. Si on note  $P(A) = \frac{\text{surface de } A}{\text{surface totale}}$ ,  $P$  est une mesure de probabilité, c'est-à-dire que la mesure totale  $P(S)$  vaut 1. La probabilité que  $x$  tombe dans  $A$  est  $P(A)$ .

Vous avez vu en théorie de la mesure qu'une mesure n'est définie que pour les ensembles appartenant à une tribu donnée. Le paradoxe suivant montre que l'idée intuitive de probabilité peut être mise en défaut si on sort de ce cadre.

**Le paradoxe de Banach-Tarski**

Soit  $S$  une sphère. Il est possible de construire un ensemble  $A \subset S$  qui a les propriétés suivantes. D'une part, il existe 3 rotations  $R_1, R_2, R_3$  telles que  $S$  soit union disjointe des images de  $A$  par ces 3 rotations :

$$S = R_1(A) \cup R_2(A) \cup R_3(A) \quad \text{et} \quad R_i(A) \cap R_j(A) = \emptyset \text{ si } i \neq j.$$

Intuitivement, les ensembles obtenus par rotation de  $A$  ont la même surface que  $A$ , donc si on choisit un point au hasard sur  $S$ , la probabilité que  $x \in A$  est  $1/3$ .

D'autre part, il existe 4 rotations  $R'_1, R'_2, R'_3, R'_4$  telles que

$$S = R'_1(A) \cup R'_2(A) \cup R'_3(A) \cup R'_4(A) \quad \text{et} \quad R'_i(A) \cap R'_j(A) = \emptyset \text{ si } i \neq j.$$

Si on raisonne comme précédemment, on conclut cette fois que la probabilité que  $x \in A$  est  $1/4$  ! La raison de cette contradiction est que l'ensemble  $A$  n'est pas mesurable et qu'on ne peut pas définir sa surface. Dans les années 1930, Kolmogorov a développé une théorie des probabilités rigoureuse s'appuyant sur la théorie de la mesure et l'intégration de Lebesgue. C'est l'approche communément utilisée de nos jours. Elle permet d'unifier dans un même cadre toutes les situations, ce qui la rend particulièrement efficace.

## 1.2 La boîte noire $\Omega$

Comment modéliser le résultat d'une expérience aléatoire? Le résultat  $X$  est généralement un nombre réel, mais il peut prendre plusieurs valeurs si on renouvelle l'expérience. Si on note  $\omega$  le paramètre (ou l'ensemble de paramètres) qui conditionne le résultat de l'expérience, la valeur de  $X$  est déterminée par  $\omega$ . Si on note  $\Omega$  l'ensemble des paramètres  $\omega$ ,  $X$  est alors une fonction définie sur  $\Omega$ . Le côté aléatoire de l'expérience n'apparaît pas dans la définition de  $X$ , il est rejeté sur le paramètre  $\omega$ . L'expérience aléatoire consiste à choisir  $\omega$  au hasard, et  $\omega$  détermine  $X$ . La façon de choisir  $\omega$  au hasard est donnée par une probabilité  $P$  sur  $\Omega$ .

Reprenons l'exemple de la météorite, avec  $X$  qui vaut 1 si la météorite tombe dans l'hémisphère nord et 0 sinon; le paramètre  $\omega$  est ici le point d'impact de la météorite. La probabilité que l'impact soit dans l'ensemble  $A$  est proportionnelle à la surface de  $A$ .

En fait, "l'ensemble des paramètres qui conditionnent le résultat de l'expérience" peut avoir plusieurs significations, de qui donne des ensembles  $\Omega$  et des probabilités  $P$  différents.

**Exemple.** Considérons une roue avec 5 secteurs circulaires de même taille, numérotés de 1 à 5. On fait tourner la roue et  $X$  est le secteur dans lequel la roue s'immobilise.

$\omega$  peut être le secteur où la roue s'arrête, d'où  $\Omega = \{1, 2, 3, 4, 5\}$ . On prend  $P$  la probabilité uniforme sur  $\Omega$  ( $P(1) = P(2) = \dots = P(5) = \frac{1}{5}$ ), et  $X(\omega) = \omega$ . La loi de  $X$  est donnée par

$$P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = \frac{1}{5}.$$

Une autre possibilité est de considérer l'angle  $\omega$  au moment de l'arrêt, et  $\Omega = [0, 2\pi[$ .  $P$  est la probabilité uniforme sur  $\Omega$ :  $P = \frac{\lambda}{2\pi}$ , où  $\lambda$  est la mesure de Lebesgue, et

$$X(\omega) = 1 \text{ si } 0 \leq \omega < \frac{2\pi}{5}, \quad X(\omega) = 2 \text{ si } \frac{2\pi}{5} \leq \omega < \frac{4\pi}{5}, \dots \quad X(\omega) = 5 \text{ si } \frac{8\pi}{5} \leq \omega < 2\pi.$$

On trouve la même loi pour  $X$ .

Une troisième possibilité est de prendre pour  $\Omega$  l'ensemble de tous les paramètres physiques de la roue au moment du lancer (position de la roue, vitesse de lancer, ...). Cela détermine entièrement la position au moment de l'arrêt puisque, selon les lois de la mécanique, c'est un système déterministe. Malgré la complexité considérable de  $\Omega$ , cela ne change pas l'expérience, ni la loi de  $X$ .

L'ensemble  $\Omega$  est parfois appelé "l'univers", en général on se contente de savoir que  $\Omega$  existe sans chercher à le déterminer. Ce qui compte vraiment est en fait la loi de  $X$ , c'est-à-dire ce qu'on peut observer. Quand on considère une unique variable aléatoire à valeurs dans  $E$ , on peut se contenter de prendre  $\Omega = E$ , comme dans le premier cas de l'exemple ci-dessus. Mais quand on veut considérer plusieurs variables aléatoires, il est nécessaire de les définir sur le même ensemble  $\Omega$ , et  $\Omega$  est alors une espèce de "boîte noire" qui permet de donner du sens aux calculs. On en reparlera en cours.

## 2 Vocabulaire

### 2.1 Espace probabilisé

**Définition.** Soit  $\Omega$  un ensemble. Une **tribu** (on dit aussi une  $\sigma$ -algèbre) est un ensemble  $\mathcal{F}$  de parties de  $\Omega$  vérifiant les propriétés suivantes :

- $\Omega \in \mathcal{F}$ ,
- si  $A \in \mathcal{F}$  alors  $A^c \in \mathcal{F}$  (où  $A^c$  désigne le complémentaire de  $A$  dans  $\Omega$ ),
- Si  $A_n \in \mathcal{F}$  pour tout  $n \in \mathbb{N}$  alors  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$ .

L'ensemble des parties de  $\Omega$ , noté  $\mathcal{P}(\Omega)$ , est une tribu. C'est la tribu généralement utilisée quand  $\Omega$  est fini ou dénombrable.

Quand  $\Omega$  est un intervalle ou une partie de  $\mathbb{R}^n$ , la tribu généralement utilisée est la tribu des boréliens.

**Définition.** Soit  $\Omega$  un ensemble et  $\mathcal{F}$  une tribu de  $\Omega$ . Une **mesure** (définie sur  $\mathcal{F}$ ) est une application  $\mu$  de  $\mathcal{F}$  dans  $\mathbb{R}^+ \cup \{+\infty\}$  telle que  $\mu(\emptyset) = 0$ , et si  $(A_n)_{n \in \mathbb{N}}$  est une famille d'ensembles disjoints de  $\mathcal{F}$ , alors  $\mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$  ( $\sigma$ -additivité).

Si de plus  $\mu(\Omega) = 1$ , on dit que  $\mu$  est une **mesure de probabilité** ou une **probabilité**.

**Définition.** Soit  $\Omega$  un ensemble,  $\mathcal{F}$  une tribu sur  $\Omega$  et  $P$  une mesure de probabilité définie sur  $\mathcal{F}$ . On dit que  $(\Omega, \mathcal{F}, P)$  est un **espace probabilisé** (ou **espace de probabilité**).

Un ensemble  $A \in \mathcal{F}$  est appelé un **événement**. Si  $P(A) = 1$ , on dit que l'événement  $A$  a lieu  **$P$ -presque sûrement** ( $P$ -p.s.), ou simplement **presque sûrement** (p.s.).

Dans toute la suite du cours,  $(\Omega, \mathcal{F}, P)$  désignera un espace probabilisé. Cet espace sera souvent sous-entendu.

### Exemples.

- On note  $\#A$  le cardinal de l'ensemble  $A$  (c'est-à-dire le nombre d'éléments dans  $A$ ). Si  $\Omega$  est fini, la probabilité uniforme sur  $\Omega$  est donnée par  $P(A) = \frac{\#A}{\#\Omega}$  pour tout  $A \subset \Omega$ . Cette probabilité est définie sur la tribu  $\mathcal{P}(\Omega)$ .
- Soit  $\lambda$  la mesure de Lebesgue sur  $\mathbb{R}$  et  $a, b \in \mathbb{R}$ ,  $a < b$ . La probabilité uniforme sur  $[a, b]$  est  $P = \frac{1}{b-a}\lambda$ . Elle est définie sur les boréliens de  $[a, b]$ . Pour un sous-intervalle  $I$  de longueur  $L$ , on a  $P(I) = \frac{L}{b-a}$  (probabilité proportionnelle à la longueur de l'intervalle et vérifiant  $P([a, b]) = 1$ ).

## 2.2 Variable aléatoire, loi

**Définition.** Soit  $\mathcal{F}$  une tribu sur  $\Omega$  et  $\mathcal{G}$  une tribu sur  $E$ . Une **variable aléatoire** (v.a.) à valeurs dans  $E$  est une fonction  $X: \Omega \rightarrow E$  mesurable pour  $\mathcal{F}$  et  $\mathcal{G}$ .

Dans ce cours, on considérera essentiellement des **variables aléatoires réelles**, c'est-à-dire  $E = \mathbb{R}$  et  $\mathcal{G}$  est la tribu des boréliens de  $\mathbb{R}$ . Si  $X$  est une fonction mesurable à valeurs dans  $\mathbb{C}$  ou  $\mathbb{R}^d$ , on parle de variable aléatoire complexe ou vectorielle.

Les variables aléatoires sont généralement notées en lettres capitales :  $X, Y, Z, \dots$

### Notation.

$P(X \in B)$  est une notation probabiliste abrégée signifiant  $P(\{\omega \in \Omega \mid X(\omega) \in B\}) = P(X^{-1}(B))$ .  
De même :  $P(X = 0) = P(\{\omega \in \Omega \mid X(\omega) = 0\})$ ,  $P(X > 1) = P(\{\omega \in \Omega \mid X(\omega) > 1\})$ .

**Définition.** Soit  $X$  une variable aléatoire définie sur l'espace probabilisé  $(\Omega, \mathcal{F}, P)$  à valeurs dans  $(E, \mathcal{G})$ . La **loi de  $X$**  est la mesure image de  $P$  par  $X$  : c'est la mesure de probabilité  $P_X$  sur  $E$  définie par

$$\forall B \in \mathcal{G}, P_X(B) \stackrel{\text{def}}{=} P(X \in B).$$

(La notation de la loi de  $X$  varie selon les livres :  $P_X, P^X, \mathcal{L}(X), \dots$ )

On note souvent  $X \sim Q$  pour dire que  $X$  est de loi  $Q$  (c'est-à-dire  $P_X = Q$ ).

**Exemple.** On lance 2 fois une pièce équilibrée (c'est-à-dire qu'on a autant de chance de tomber sur pile ou sur face). On note 0 pour pile et 1 pour face.

Soit  $\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ ; c'est l'ensemble des tirages possibles.

On prend la probabilité uniforme sur  $\Omega$  : si  $A \subset \Omega$ ,  $P(A) = \frac{\#A}{\#\Omega}$ .

On définit la variable aléatoire  $X: \Omega \rightarrow \mathbb{R}$  comme étant le nombre de 1 (face) obtenu au cours des 2 lancers.  $X$  prend ses valeurs dans  $\{0, 1, 2\}$ . Sa loi est donnée par :

$$P(X = 0) = P(\{(0, 0)\}) = \frac{1}{4}, P(X = 1) = P(\{(0, 1), (1, 0)\}) = \frac{1}{2}, P(X = 2) = P(\{(1, 1)\}) = \frac{1}{4}.$$

De façon équivalente, on peut écrire  $P_X = \frac{1}{4}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{4}\delta_2$ , où  $\delta_a$  est la masse de Dirac en  $a$  (cf section 2.1).

Soit  $Y$  le nombre de 0.  $X$  et  $Y$  ont la même loi bien que  $X$  et  $Y$  ne soient pas les mêmes variables aléatoires.

## 2.3 Espérance, variance

Soit  $X$  une variable aléatoire réelle définie sur  $\Omega$ .

On dit que  $X \in L^p(\Omega, \mathcal{F}, P)$ , ou simplement que  $X \in L^p$ , si  $\int_{\Omega} |X|^p dP < +\infty$ .

$X \in L^1$  signifie que  $X$  est intégrable.

**Remarque.** Comme  $P(\Omega) = 1$ , si  $X$  est bornée alors  $X \in L^1$ . De plus, si  $X \in L^p$  et  $p \geq q \geq 1$  alors  $X \in L^q$ .

**Définition.** Soit  $X$  une variable aléatoire réelle définie sur  $\Omega$ . Si  $X \geq 0$  ou si  $X \in L^1$ , son **espérance** est

$$E(X) = \int_{\Omega} X dP = \int_{\Omega} X(\omega) dP(\omega).$$

L'espérance est notée avec une double barre dans certains livres :  $\mathbb{E}(X)$ .

Si  $E(X) = 0$ , on dit que  $X$  est **centrée**.

**Propriété.**  $E(aX + b) = aE(X) + b$ . En particulier,  $X - E(X)$  est une variable aléatoire centrée.

**Définition.** Soit  $X$  une variable aléatoire réelle définie sur  $\Omega$ . Si  $X \in L^p$ , avec  $p \in \mathbb{N}^*$ , le **moment d'ordre  $p$**  de  $X$  est  $E(X^p)$ .

Si  $X \in L^2$  et  $m = E(X)$ , sa **variance** est  $\text{Var}(X) \stackrel{\text{def}}{=} E((X - m)^2) = E(X^2) - m^2$ . C'est le moment centré d'ordre 2. L'**écart-type** de  $X$  est  $\sigma(X) = \sqrt{\text{Var}(X)}$ . On pose souvent  $\text{Var}(X) = \sigma^2$ .

Si  $\text{Var}(X) = 1$ , on dit que  $X$  est **réduite**.

**Propriétés.**

$$\text{Var}(X + b) = \text{Var}(X).$$

$$\text{Var}(aX) = a^2 \text{Var}(X).$$

En particulier, si  $\sigma = \sigma(X)$  et  $m = E(X)$ , alors  $\frac{X-m}{\sigma}$  est une variable aléatoire centrée réduite.

**Définition.**

Soit  $X$  et  $Y$  deux variables aléatoires définies sur  $\Omega$ . Si  $X$  et  $Y$  sont  $L^2$ , leur **covariance**  $\text{Cov}(X, Y)$  est

$$\text{Cov}(X, Y) \stackrel{\text{def}}{=} E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y).$$

On a :  $\text{Cov}(X, X) = \text{Var}(X)$ .

## 2.4 Traduction entre langage probabiliste et théorie de la mesure

Si vous n'êtes pas à l'aise avec le langage probabiliste, n'hésitez pas à faire la traduction avec le langage de la théorie de la mesure et de l'intégration, avec lequel vous êtes sûrement plus familiers. Le langage probabiliste véhicule mieux l'intuition probabiliste mais est souvent elliptique (ce qui est parfois dangereux). Le langage de la théorie de la mesure est plus précis sur la nature des objets mathématiques qu'on manipule. Voici un petit tableau de traduction.

probabilités	théorie de la mesure
événement	ensemble mesurable
variable aléatoire réelle $X$	fonction mesurable $X: \Omega \rightarrow \mathbb{R}$
$P(X \in B)$	$P(\{\omega \in \Omega \mid X(\omega) \in B\})$
probabilité que $X$ soit dans $B$	mesure de l'ensemble des points $\omega$ tels que $X(\omega) \in B$
espérance	intégrale
$E(X), E(\varphi(X))$	$\int_{\Omega} X(\omega) dP(\omega), \int_{\Omega} \varphi(X(\omega)) dP(\omega)$
fonction caractéristique	transformée de Fourier (à un coefficient près)

## 3 Les lois classiques

### 3.1 Lois discrètes

**Notation.** Soit  $\Omega$  un ensemble et  $x \in \Omega$ . La **masse de Dirac en  $x$** , notée  $\delta_x$ , est la mesure définie par :

$$\forall A \subset \Omega, \delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

C'est une mesure de probabilité définie sur la tribu  $\mathcal{P}(\Omega)$ . Elle est également définie sur toute tribu  $\mathcal{F}$  de  $\Omega$  puisque  $\mathcal{F} \subset \mathcal{P}(\Omega)$ .

**Définition.** Une probabilité  $Q$  est dite **discrète** si elle s'écrit comme combinaison linéaire finie ou dénombrable de masses de Dirac :

$$Q = \sum_{i \in I} p_i \delta_{x_i}, \text{ où } I \text{ est un ensemble dénombrable.}$$

Comme  $Q$  est une probabilité, on a nécessairement  $p_i \geq 0$  et  $\sum_{i \in I} p_i = 1$ .

**Définition.** Une variable aléatoire est dite **discrète** si sa loi est une probabilité discrète. Autrement dit, une variable aléatoire  $X$  est discrète si et seulement si  $X$  prend (presque sûrement) ses valeurs dans un ensemble fini ou dénombrable  $E$ .

La plupart du temps,  $X$  prend des valeurs entières.

Notons  $E = \{x_i; i \in I\}$ . Si  $\forall i \in I, P(X = x_i) = p_i$ , la loi de  $X$  est  $P_X = \sum_{i \in I} p_i \delta_{x_i}$ .

Pour donner la loi d'une variable aléatoire discrète  $X$ , il est équivalent de donner  $\forall i \in I, P(X = x_i) = p_i$  ou d'écrire  $P_X$  sous forme de combinaison linéaire de masses de Dirac.

### 3.2 Exemples de lois discrètes

#### a) Loi uniforme sur un ensemble fini $\Omega$ .

C'est la loi qui consiste à donner la même probabilité  $\frac{1}{N}$  à chaque élément d'un ensemble de cardinal  $N$ .

Exemple : tirage d'une carte dans un jeu de 52 cartes. La probabilité de tirer le roi de coeur est  $1/52$ .

#### b) Loi de Bernoulli $b(p) = \mathcal{B}(1, p)$ , avec $0 < p < 1$ .

La variable aléatoire  $X$  suit une loi de Bernoulli de paramètre  $p$  si  $X$  prend ses valeurs (p.s.) dans  $\{0, 1\}$ , avec  $P(X = 1) = p$  et  $P(X = 0) = 1 - p$ . Cette loi est notée  $b(p)$  ou  $\mathcal{B}(1, p)$ . Ceci s'écrit également :

$$b(p) = (1 - p)\delta_0 + p\delta_1.$$

Exemple : jeu de pile ou face. On lance une pièce de monnaie, 1 représente un succès, 0 un échec. La pièce est équilibrée si  $p = 1/2$ , sinon on dit que la pièce est biaisée ou pipée.

Plus généralement, quand une v.a. ne prend p.s. que 2 valeurs, on dit qu'elle suit une loi de Bernoulli.

#### c) Loi binomiale $\mathcal{B}(n, p)$ , avec $n \geq 1$ et $0 < p < 1$ .

La variable aléatoire  $X$  suit une loi binomiale de taille  $n$  et de paramètre  $p$  si  $X$  prend ses valeurs dans  $\{0, 1, \dots, n\}$  et  $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$  pour  $0 \leq k \leq n$ . Cette loi est notée  $\mathcal{B}(n, p)$ , ce qui s'écrit :

$$\mathcal{B}(n, p) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} \delta_k.$$

Exemple : loi du nombre de faces obtenus si on fait  $n$  lancers successifs d'une pièce ayant la probabilité  $p$  de tomber sur face.

#### d) Loi de Poisson $\mathcal{P}(\lambda)$ , avec $\lambda > 0$ .

La variable aléatoire  $X$  suit une loi de Poisson de paramètre réel  $\lambda > 0$  si  $X$  prend ses valeurs dans  $\mathbb{N}$  et  $P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$  pour tout entier  $k \in \mathbb{N}$ . Cette loi est notée  $\mathcal{P}(\lambda)$ , ce qui s'écrit :

$$\mathcal{P}(\lambda) = \sum_{k \in \mathbb{N}} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \delta_k.$$

Contrairement aux lois précédentes, l'interprétation de la loi de Poisson n'est pas immédiate. La loi de Poisson est la "loi des événements rares" : nombre de pannes pendant une durée donnée, nombre de clients entrant dans une boutique, ... On verra en cours la justification de cette modélisation.

### 3.3 Lois continues

**Définition.** Une probabilité  $Q$  définie sur  $\mathbb{R}$  est dite **continue** (ou à densité) s'il existe une fonction mesurable  $f: \mathbb{R} \rightarrow \mathbb{R}$  à valeurs positives telle que, pour tout ensemble borélien  $B \in \mathcal{B}(\mathbb{R})$ , on a

$$Q(B) \stackrel{\text{def}}{=} \int_B f(x) dx = \int_{\Omega} \mathbf{1}_B(x) f(x) dx.$$

$f$  est appelée la **densité** de  $Q$ . On note  $dQ = f(x)dx$  (on verra pourquoi avec la formule de transfert).

Comme  $Q$  est une probabilité, on a nécessairement  $\int_{\mathbb{R}} f(x) dx = 1$ .

On peut aussi parler de loi continue sur  $\mathbb{R}^d$  avec une densité  $f: \mathbb{R}^d \rightarrow \mathbb{R}^+$ .

**Remarque.** La terminologie est trompeuse : la fonction  $f$  n'est pas nécessairement continue !  
 On dit aussi que  $Q$  est absolument continue par rapport à la mesure de Lebesgue (ce qui signifie qu'un ensemble de mesure nulle pour la mesure de Lebesgue est également de mesure nulle pour  $Q$ ).

**Remarque.** Si  $X$  est une variable aléatoire de loi continue  $Q$ , alors  $P(X = a) = 0$  pour tout  $a \in \mathbb{R}$  car  $P(X = a) = Q(\{a\}) = \int_a^a f(x) dx = 0$ . On dit que  $X$  n'a pas d'atome.

**Remarque.** Il existe des lois qui ne sont ni discrètes ni continues (ni un mélange des deux).

### 3.4 Exemples lois continues

#### a) Loi uniforme $\mathcal{U}([a, b])$ avec $a < b$ .

C'est la loi qui donne à chaque sous-intervalle de  $[a, b]$  une probabilité proportionnelle à sa longueur : si  $c, d \in [a, b]$  avec  $c \leq d$ ,  $P([c, d]) = \frac{d-c}{b-a}$  (le dénominateur  $b - a$  est juste là pour avoir  $P([a, b]) = 1$ ).

Exprimée en terme de densité : la densité est  $\frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$ .

Intuitivement, cette loi donne "le même poids" à chacun des points de l'ensemble. L'ensemble étant non dénombrable, c'est la densité qui est égale partout sur l'ensemble (un point isolé a une probabilité nulle).

On peut généraliser la loi uniforme à n'importe quel ensemble mesurable  $E \subset \mathbb{R}^n$  dont la mesure de Lebesgue est finie et non nulle (par exemple un pavé). La loi uniforme est alors la restriction à  $E$  de la mesure de Lebesgue, renormalisée en divisant par la mesure de Lebesgue de  $E$ .

#### b) Loi exponentielle $\mathcal{E}(\lambda) = \exp(\lambda)$ , avec $\lambda > 0$ .

Une variable aléatoire réelle  $X$  est de loi exponentielle de paramètre réel  $\lambda > 0$  si sa densité est

$$\mathbb{1}_{\mathbb{R}_+}(x)\lambda e^{-\lambda x} = \begin{cases} 0 & \text{si } x < 0 \\ \lambda e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

(l'indicatrice  $\mathbb{1}_{\mathbb{R}_+}$  indique que  $X$  prend en fait ses valeurs dans  $\mathbb{R}^+$  (p.s.) car  $P(X < 0) = 0$ ).  
 Cette loi est notée  $\mathcal{E}(\lambda)$  ou  $\exp(\lambda)$ .

Exemple : durée de vie d'un atome radioactif.

C'est une loi sans mémoire : ce qui se passe après le temps  $t$  est indépendant de ce qui s'est passé avant (voir l'exercice 7).

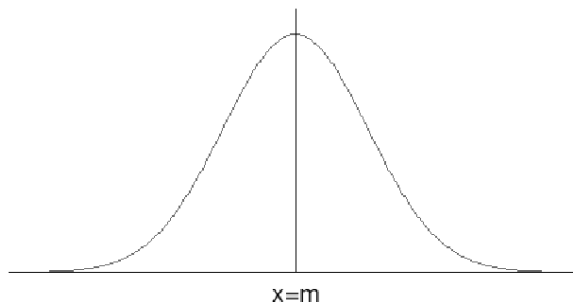
#### b) Loi de Gauss, loi normale $\mathcal{N}(m, \sigma^2)$ , avec $m \in \mathbb{R}$ et $\sigma^2 > 0$ .

La densité de la loi normale réduite  $\mathcal{N}(0, 1)$  est  $\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ .

La densité de la loi normale  $\mathcal{N}(m, \sigma^2)$  est  $\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$ .

Cette loi est également appelée loi de Gauss (surtout pour  $\mathcal{N}(0, 1)$ ) ou loi gaussienne. La courbe de la densité de cette loi est appelée une gaussienne.

Dessin d'une gaussienne  
 (courbe "en cloche")



$x = m$  est l'axe de symétrie de la densité de  $\mathcal{N}(m, \sigma^2)$ , et la courbe s'écrase extrêmement vite sur 0 quand on s'éloigne. Le paramètre  $\sigma^2$  détermine l'allure de la cloche : plus  $\sigma^2$  est petit, plus la cloche est étroite et monte haut, autrement dit plus la densité est concentrée autour de la valeur  $m$ . Les paramètres  $m$  et  $\sigma^2$  sont en fait égaux respectivement à l'espérance (valeur moyenne) et à la variance de cette loi.

La loi normale est la "loi des erreurs". Très utilisée en statistiques, elle modélise à peu près tout ce qui varie autour d'une valeur moyenne : poids des Français, durée d'ensoleillement à Paris au mois d'août, etc. Cette modélisation est en partie justifiée par le théorème central limite, que nous verrons en cours.