Najmeddine Attia
Meriem Ben Hadj Khlifa

L1
L2
L3

# L'indispensable en statistiques et en probabilités pour la Licence

## Cours, exemples pratiques et exercices corrigés

ellipses

# Chapitre 1

# Rappels de probabilités

## 1.1 Espace probabilisé

### Définition et propriétés

Lorsque l'on modélise une expérience aléatoire, on ne peut pas en général définir la probabilité de toutes les parties de l'univers $\Omega$ (l'ensemble de toutes les éventualités possibles) toutes les déterminations du hasard dans l'expérience considérée). On est obligé de se restreindre à certaines catégories de parties, qui forment une tribu. Les éléments de la tribu sont les événements de la modélisation : ce sont les parties pour lesquelles la notion de probabilité aura un sens. Dans ce cas $(\Omega, \mathscr{A})$ est appelé espace probabilisable et les éléments de $\mathscr{A}$ sont appelés les événements. Les singletons de $\mathscr{A}$ (s'ils existent) s'appellent les événements élémentaires.

**R** Soit $A$ un événement, et supposons que l'on ne s'intéresse qu'aux événements qui y sont associés. On modélise alors l'expérience aléatoire par la tribu, notée $\sigma(A)$, engendrée par cet événement : c'est la plus petite (au sens de l'inclusion) tribu qui contient $A$. Plus précisément,

$$\sigma(A) = \bigcap_{\mathscr{A}_i \text{ tribu}, \ A \subset \mathscr{A}_i} \mathscr{A}_i.$$

■ **Exemple 1.1**
1. $\sigma(\mathscr{A}) = \mathscr{A}$.
2. Soit $A \in \mathscr{A}$ alors $\sigma(A) = \{\emptyset, A, \overline{A}, \Omega\}$.
3. La tribu engendrée par la classe des ouverts de $\mathbb{R}$ est $\mathscr{B}(\mathbb{R})$ : tribu borélienne.

■

**R** On peut construire plusieurs tribus sur un ensemble $\Omega$.

**Définition 1.1.1** Une famille $(A_n)_n$, au plus dénombrable d'éléments de $\mathscr{A}$, est un système complet d'événements si

- les événements sont deux à deux incompatibles c'est-à-dire,

$$\forall i \neq j, \ A_i \cap A_j = \emptyset,$$

- $\bigcup_n A_n = \Omega$.

**Définition 1.1.2 — Espace probabilisé.** Soit $(\Omega, \mathscr{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega, \mathscr{A})$ toute application $\mathbb{P} : \mathscr{A} \to [0,1]$ telle que

1. $\mathbb{P}(\Omega) = 1$ ;
2. pour toute famille $(A_i)_{i\in I}$ au plus-dénombrable d'événements de $\mathscr{A}$ deux à deux incompatibles, la série $\sum \mathbb{P}(A_i)$ converge et

$$\mathbb{P}\Big(\bigcup_{i\in I} A_i\Big) = \sum_{i\in I} \mathbb{P}(A_i). \qquad (\ \sigma\text{-additivité})$$

Si $\mathbb{P}$ est une probabilité sur $\Omega$, le triplet $(\Omega, \mathscr{A}, \mathbb{P})$ est appelé espace probabilisé.

**R** Un espace de probabilité est un cas particulier d'espace mesuré, pour lequel la masse totale de la mesure est égale à 1. Pour $A \in \mathscr{A}, \mathbb{P}(A)$ représente la probabilité d'occurrence de l'événement $A$.

**Proposition 1.1.1 — Propriétés des probabilités.** Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et soient $A$ et $B$ deux événements. Alors

1. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$;
2. $\mathbb{P}(\emptyset) = 0$ ;
3. si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ ; (croissance)
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ ;
5. pour toute famille $(A_n)_{n\geq 1}$ d'événements au plus dénombrable de $\Omega$, on a

$$\mathbb{P}\Big(\bigcup_n A_n\Big) \leq \sum_n \mathbb{P}(A_i).$$

**R** Si $A \cap B = \emptyset$ alors

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

**Proposition 1.1.2 — continuité des probabilités.** Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.

1. Pour toute suite $(C_n)$ croissante $(C_n \subset C_{n+1})$ d'événements de $\mathscr{A}$, on a
$$\mathbb{P}\Big(\bigcup_n C_n\Big) = \lim_{n\to+\infty} \mathbb{P}(C_n). \qquad \text{(continuité croissante)}$$
2. Pour toute suite $(D_n)$ décroissante $(D_{n+1} \subset D_n)$ d'événements de $\mathscr{A}$, on a
$$\mathbb{P}\Big(\bigcap_n D_n\Big) = \lim_{n\to+\infty} \mathbb{P}(D_n). \qquad \text{(continuité décroissante)}$$

R

1. Si $\Omega$ est au plus-dénombrable on peut prendre toujours $\mathscr{A} = \mathscr{P}(\Omega)$. Dans ce cas, une probabilité $\mathbb{P}$ est entièrement déterminée par sa valeur sur les événements élémentaires.
2. Si $\Omega$ est non dénombrable, il est nécessaire en général de restreindre la classe des événements, afin de pouvoir construire un espace probabilisé.

■ **Exemple 1.2 — Univers fini.** On lance un dé deux fois et on s'intéresse aux résultats affichés. Il est naturel de prendre

$$\Omega = [|1,6|]^2, \quad \mathscr{A} = \mathscr{P}(\Omega) \quad et \quad \forall (i,j) \in \Omega, \quad \mathbb{P}\big\{(i,j)\big\} = \frac{1}{36}.$$

$\mathbb{P}$ est la probabilité uniforme sur $\Omega$. Le choix de $\mathbb{P}$ correspond à l'idée que tous les résultats possibles pour les deux lancers sont équiprobables. En particulier,

$$\forall A \in \mathscr{A}, \qquad \mathbb{P}(A) = \frac{Card\, A}{Card\, \Omega} = \frac{Card\, A}{36}.$$

■

■ **Exemple 1.3 — Univers dénombrable.** On jette une pièce de monnaie équilibrée jusqu'à obtenir « Face ». On peut choisir

$$\Omega = \mathbb{N}^* \cup \{\infty\} \quad \text{et} \quad \mathscr{A} = \mathscr{P}(\Omega)$$

avec l'événement $\{n\}$ représente « obtenir le premier Face au $n$-ième lancer » et $\{\infty\}$ représente « Face ne sorte jamais ». Il est naturel de prendre

$$\mathbb{P}(\{n\}) = 2^{-n}, \qquad n = 1, 2 \ldots,$$

en effet, si on lance $n$ fois une pièce de monnaie équilibrée alors tous les issues sont équiprobables. En particulier,

$$\mathbb{P}(\{\infty\}) = 1 - \sum_{n=1}^{+\infty} 2^{-n} = 0,$$

ce qui est conforme à l'intuition.
La probabilité que le premier Face sorte après un nombre impair de lancers est de

$$\mathbb{P}(\{1,3,5,\ldots\}) = \sum_{n=0}^{+\infty} \mathbb{P}(2n+1) = 2^{-1} \sum_{n=0}^{+\infty} 2^{-2n} = \frac{2}{3}.$$

La probabilité que le premier Face sorte après un nombre pair de lancers est de

$$\mathbb{P}(\{2,4,6,\ldots\}) = \sum_{n=1}^{+\infty} \mathbb{P}(2n) = \sum_{n=1}^{+\infty} 2^{-2n} = \sum_{n=0}^{+\infty} 2^{-2n} - 1 = \frac{1}{3}.$$

■

**■ Exemple 1.4 — Univers non-dénombrable.**

1. Soit $\Omega = \mathbb{R}$ et $\mathscr{A} = \mathscr{B}(\mathbb{R})$. On vérifie facilement que, pour $a < b \in \mathbb{R}$, l'application $\mathbb{P}$ définie par

$$\mathbb{P}(A) = \frac{1}{b-a} \int_A \mathbf{1}_{[a,b]}(x)\, dx, \qquad \forall A \in \mathscr{A}$$

est une probabilité sur $\mathbb{R}$.

2. $\Omega = \mathbb{R}$ et $\mathscr{A} = \mathscr{B}(\mathbb{R})$. Soit $a \in \mathbb{R}$, alors, pour $A \in \mathscr{A}$,

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ \\ 0 & \text{si } a \notin A \end{cases}$$

est une probabilité sur $\mathbb{R}$ appelée probabilité de Dirac.

■

**Définition 1.1.3** Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $A \in \mathscr{A}$.
• Si $\mathbb{P}(A) = 0$ on dit que $A$ est négligeable.
• Si $\mathbb{P}(A) = 1$ on dit que $A$ est presque sûr.
Une propriété définie sur $\Omega$ est vraie presque sûrement (on note p.s.) si elle est vraie en dehors d'un ensemble négligeable.

**R** Ces notions dépendent de la probabilité choisie. En effet, soit la tribu $\mathscr{A} = \{\emptyset, A, \bar{A}, \mathbb{R}\}$, avec $A = \{0\}$. Considérons $\mathbb{P}_1$ la mesure de Lebesque sur $[0,1]$ et $\mathbb{P}_2 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Alors $A$ est négligeable par rapport à la mesure $\mathbb{P}_1$ mais on a $\mathbb{P}_2(A) = \frac{1}{2}$.

Dans la suite, lorsque nous considérerons $\mathbb{R}$ comme espace probabilisé, nous le supposerons toujours muni de sa tribu borélienne, sauf mention du contraire.

## Probabilité conditionnelle

**Définition et Théorème. 1.1.3 — Probabilité conditionnelle.** Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $B \in \mathscr{A}$ tel que $\mathbb{P}(B) \neq 0$. Alors la fonction $\mathbb{P}_B$ définie sur $\mathscr{A}$ par

$$\forall A \in \mathscr{A}, \quad \mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

est une probabilité sur $\Omega$. On l'appelle probabilité conditionnelle sachant $B$. On la note aussi par $\mathbb{P}(\cdot \mid B)$.

**R**

1. Pour tout événement $A$ on a $\mathbb{P}_B(\overline{A}) = 1 - \mathbb{P}_B(A)$.
2. $\mathbb{P}_B(\overline{B}) = 0$, on dit que $\mathbb{P}_B$ est portée par $B$.
3. $\mathscr{A}_B = \{A \cap B, A \in \mathscr{A}\}$ est une tribu et $(\Omega, \mathscr{A}_B, \mathbb{P}_B)$ est un espace probabilisé.

## Indépendance des événements

Intuitivement deux événements sont indépendants si la réalisation de l'un de ces événements n'influe pas sur la probabilité de l'autre.

**Définition 1.1.4 — Événements indépendants.** Deux événements A et B de l'espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ sont dit indépendants si

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

**R**

1. Bien que la définition utilisant les probabilités conditionnelles soit plus intuitive : $\mathbb{P}_B(A) = \mathbb{P}(A)$, elle a l'inconvénient d'être moins générale, et de ne pas faire jouer un rôle symétrique aux deux événements $A$ et $B$. En particulier, si $\mathbb{P}(B) > 0$. Alors A et B sont indépendants si, et seulement si, $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

2. Un événement $A$ de probabilité nulle est indépendant de tout événement $B$. En effet,
$$0 = I\!P(A \cap B) \leq I\!P(A) = 0 = I\!P(A)\, I\!P(B).$$
3. Un événement A est indépendant de lui-même à la condition que $I\!P(A) = 0$ ou $I\!P(A) = 1$. En effet, dans ce cas
$$I\!P(A) = I\!P(A \cap A) = I\!P(A)\; I\!P(A).$$
4. Si $A$ et $B$ sont indépendants alors il en est de même pour $\overline{A}$ et $B$, $\overline{A}$ et $\overline{B}$. En effet
$$\begin{aligned} I\!P(\overline{A} \cap B) &= I\!P(B) - I\!P(A \cap B) = I\!P(B) - I\!P(A)\, I\!P(B) \\ &= I\!P(B)\big(1 - I\!P(A)\big) = I\!P(B)\, I\!P(\overline{A}). \end{aligned}$$
5. Si $I\!P(A) = 1$ alors $A$ est indépendant de tout événement. En effet, $\overline{A}$ est de probabilité nulle qui est indépendant de tout événement.

**Définition 1.1.5 — Indépendance mutuelle.** Soit $(\Omega, \mathscr{A}, I\!P)$ un espace probabilisé. On dit que les événements $(A_i)_{i \in I}$ sont mutuellement indépendants (ou indépendants) si
$$\forall J \subset I \ \text{ finie} \qquad I\!P\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} I\!P(A_j).$$

**R** Si $(A_i)_i$ est une famille d'événements mutuellement indépendants alors les événements $(A_i)_i$ sont indépendants deux à deux. Cependant, la réciproque est fausse.

## 1.2 Variable aléatoire, loi

**Définition 1.2.1** Soit $(\Omega, \mathscr{A}, I\!P)$ un espace probabilisé et $(E, \mathscr{E})$ un espace mesurable. On appelle variable aléatoire sur $\Omega$ toute application mesurable $X : (\Omega, \mathscr{A}) \to (E, \mathscr{E})$.
- Si $E \subset \mathbb{R}$ on parle de variable aléatoire réelle (v.a.r.).
- Si $E \subset \mathbb{Z}$ on parle de variable aléatoire entière.

■ **Exemple 1.5** Exemples de variables aléatoires :
1. Variable $T$ correspondant à la taille d'un étudiant.
2. Variable $L$ correspondant à la longueur d'un train.
3. Variable $A$ correspondant au temps d'attente à une caisse.
4. Variable $B$ correspondant au poids d'un bébé à la naissance.

■

**Définition 1.2.2** Soit $X$ une variable aléatoire réelle. On appelle loi de $X$ la mesure-image de $\mathbb{P}$ par $X$. C'est donc la mesure de probabilité sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, notée $\mathbb{P}_X$, définie par

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)), \qquad B \in \mathscr{B}(\mathbb{R}).$$

En pratique on écrit plutôt :

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) \qquad \Big( = \mathbb{P}\big(\{\omega \in \Omega : X(\omega) \in B\}\big)\Big).$$

**R** Deux variables aléatoires peuvent suivre la même loi sans être égales.

**Théorème. 1.2.1 — Théorème de transfert.** Soit $X$ une v.a.r. définie sur l'espace $(\Omega, \mathscr{A}, \mathbb{P})$. Alors, une fonction $f : \mathbb{R} \to \mathbb{R}$ est $\mathbb{P}_X$-intégrable si $f \circ X$ est $\mathbb{P}$-intégrable et dans ce cas on a

$$\int_\Omega f \circ X(\omega)\, d\mathbb{P}(\omega) = \int_\mathbb{R} f(x) d\mathbb{P}_X(x).$$

### Fonction de répartition d'une variable aléatoire réelle

**Définition 1.2.3** Soit $X$ est une v.a.r., la fonction de répartition de $X$ est la fonction $F_X : \mathbb{R} \to [0,1]$ définie par

$$F_X(t) = \mathbb{P}(X \leq t) = \mathbb{P}_X\big(]-\infty, t]\big), \quad \forall t \in \mathbb{R}.$$

**Proposition 1.2.2** Soit $X$ une variable aléatoire de fonction de répartition $F_X$. Alors,

1. $F_X$ est croissante ;
2. $F_X$ est continue à droite en tout point ;
3. $F_X$ admet une limite à gauche en tout point.
4. $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.

**R**

1. $\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F_X(b) - F(a)$.
2. $\mathbb{P}(X > b) = 1 - F_X(b)$.
3. Pour tout $x \in \mathbb{R}$, $0 \leq F_X(x) \leq 1$

4. $I\!P(X < x)$ peut être distinct de $I\!P(X \leq x)$ en effet
$$I\!P(X \leq x) - I\!P(X < x) = I\!P(X = x).$$
En particulier $F_X$ est continue en $x \iff I\!P(X = x) = 0$.
5. $F_X$ caractérise la loi $I\!P_X$ de $X$. C'est-à-dire
$$F_X = F_Y \iff I\!P_X = I\!P_Y.$$

## Variables aléatoires discrètes

**Définition 1.2.4** Soit $X : \Omega \to E$ une v.a.r.. $X$ est dite variable aléatoire discrète si $E$ est au plus-dénombrable. La loi de $X$ est alors
$$I\!P_X = \sum_{x \in E} p_x \, \delta_x$$
où $p_x = I\!P(X = x)$ et $\delta_x$ désigne la mesure de Dirac en $x$.

■ **Exemple 1.6**

1. On dit que $X$ suit la loi uniforme sur un ensemble fini $E = \{x_1, \ldots, x_n\}$, et on note $X \sim \mathscr{U}(E)$, si $X$ est à valeurs dans $E$ et
$$I\!P(X = x) = \frac{1}{Card(E)}, \qquad \forall x \in E.$$
2. On dit $X$ suit la loi de Bernoulli de paramètre $p \in [0,1]$ et on note $X \sim \mathscr{B}(p)$ si $X$ est à valeurs dans $\{0,1\}$ avec
$$I\!P(X = 1) = p \quad \text{et} \quad I\!P(X = 0) = 1 - p.$$
3. Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$. On dit $X$ suit la loi binomiale de paramètre $n$ et $p$ et on note $X \sim \mathscr{B}(n,p)$ si $X$ à valeurs dans $[|0,n|]$ et, pour tout $k \in [|0,n|]$, on a
$$I\!P(X = k) = \mathbf{C}_n^k p^k (1-p)^{n-k}.$$
avec $\mathbf{C}_n^k = \dfrac{n!}{k!(n-k)!}$ est le coefficient binomial.
4. On dit que $X$ suit la loi géométrique de paramètre $p \in ]0,1[$ et on note $X \sim \mathscr{G}(p)$ si $X$ à valeurs dans $\mathbb{N}^*$ et
$$I\!P(X = k) = (1-p)^{k-1} p.$$
5. On dit que $X$ suit la loi de Poisson de paramètre $\lambda > 0$ et on note $X \sim \mathscr{P}(\lambda)$ si $X$ est à valeurs dans $\mathbb{N}$, et
$$I\!P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Intuitivement, elle correspond au nombre d'événements rares qui se sont produits durant une période longue.

■

R Si X suit une loi géométrique, alors X est « sans mémoire » à savoir :

$$\forall\,(k,l)\in\mathbb{N}^{*2},\qquad \mathbb{P}_{(\mathsf{X}>k)}(\mathsf{X}>k+l)=\mathbb{P}(\mathsf{X}>l). \tag{1.1}$$

Cette propriété dit, par exemple, que même si le numéro 1 n'est pas sorti pendant 10 semaines consécutives à la loterie, cela ne rend pas sa prochaine apparition plus probable.

**Exercice 1.1** Soit $n\in\mathbb{N}^*$. Un tireur vise une cible $n$ fois de suite avec une probabilité $p\in]0,1[$ de la toucher. On suppose que les tirs sont indépendants, on note X le nombre de cibles atteintes.

1. Déterminer la loi de la variable X.
2. Combien doit-il tirer de coups, en fonction de $p$, afin que la cible soit atteinte avec une probabilité supérieure ou égale à $0,95$ ?

**Solution 1.1** 1. Il est clair que $\mathsf{X}\sim\mathscr{B}(n,p)$. Donc $\forall\,k\in\{0,\cdots,n\}$, on a

$$\mathbb{P}(\mathsf{X}=k)=\mathbf{C}_n^k p^k\,(1-p)^{n-k}.$$

2. La cible est atteinte signifie que $\mathsf{X}\geq 1$. Ainsi, On cherche l'entier $n$ tel que $\mathbb{P}(\mathsf{X}\geq 1)\geq 0,95$.

$$\begin{aligned}
\mathbb{P}(\mathsf{X}\geq 1)\geq 0,95 &\iff 1-\mathbb{P}(\mathsf{X}=0)\geq 0,95\\
&\iff 1-(1-p)^n\geq 0,95\\
&\iff (1-p)^n\leq 0,05\\
&\iff n\ln(1-p)\leq \ln(0,05)\\
&\iff n\geq \frac{\ln(0,05)}{\ln(1-p)}.
\end{aligned}$$

**Exercice 1.2** Une source souhaite envoyer un message à une destination. Pour cela, elle se connecte à un réseau de communication. Or, ce réseau est caractérisé par une probabilité de perte de message $q$. On a la probabilité de transmission du message est $p=(1-q)$.

Comme la source souhaite nécessairement que le message soit délivré à la destination, elle sera obligée de retransmettre le message autant de fois qu'il est perdu jusqu'à ce qu'une transmission réussisse. On suppose que les pertes de messages sont indépendantes.

Soit $X$ une variable aléatoire qui représente le nombre de transmissions nécessaires pour recevoir le message.

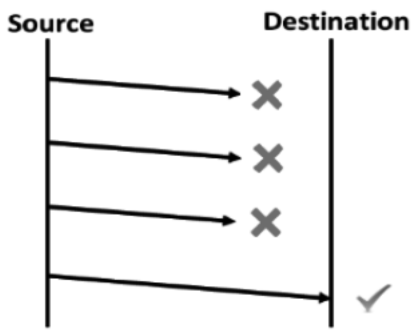


FIGURE 1.1 – shéma illustratif

Soient les événements
**E** : le message est perdu.
**S** : le message est transmis.

1. Déterminer $X(\Omega)$ l'ensemble des valeurs prise par la v.a. $X$.
2. Exprimer $\mathbb{P}(X=4)$ en fonction de **E** et **S**. Puis l'écrire en fonction de $p$.
3. Déterminer alors $\mathbb{P}(X=k), \forall k \in X(\Omega)$.

   Comme le nombre de transmissions peut parfois devenir important, on veut borner celui-là et considérer qu'une transmission fiable ne réussit pas au-delà d'une limite $l$ et que la connexion vers la destination sera perdue en dépassant $l$. Dans ce cas, on veut calculer le nombre de transmissions nécessaires pour que la connexion ne soit pas perdue.
   On a également $\mathbb{E}(X) = \frac{1}{p}$ et la fonction de répartition de $X$ est donnée par $F(x) = \mathbb{P}(X \leq x) = 1-(1-p)^k$.
4. Exprimer la probabilité de perte de la connexion en fonction de $p$ et $l$.
   Selon des expériences ultérieures, le nombre de transmissions moyen nécessaire à la réception du message est égal à 4 et la probabilité que la connexion soit perdue est égale à $0,1$.
5. Déterminer le nombre de transmissions $l$ à partir duquel la connexion serait perdue.

**Solution 1.2** 1. Il est clair qu'une transmission fiable d'un message est une répétition de l'épreuve de Bernoulli de façon indépendante jusqu'à l'obtention d'un succès.
$X$ : le nombre de transmissions nécessaires à la réception d'un message.
$X(\Omega) = \{1,2,\dots\} = \mathbb{N}^*$.

$X$ se comporte comme étant une variable aléatoire discrète.

2. On a
$$\begin{array}{rcl} I\!P(X=4) & = & I\!P(E\cap E\cap E\cap S) \\ & = & I\!P(E).I\!P(E).I\!P(E).I\!P(S) \ \ (Les\ preuves\ sont\ indpendantes) \\ & = & I\!P(E)^3.I\!P(S) \\ & = & q^3.p \\ & = & (1-p)^{4-1}.p \end{array}$$
3. On généralise,
$$\begin{array}{rcl} I\!P(X=k) & = & I\!P(E)^{k-1}.I\!P(S) \\ & = & q^{k-1}.p \\ & = & (1-p)^{k-1}p. \end{array}$$
Donc c'est la loi géométrique.
4. La connexion à la destination est supposée perdue pour $X > l$, donc on doit déterminer $I\!P(X > l)$.
$$\begin{array}{rcl} I\!P(X>l) & = & 1-I\!P(X\leq l) \\ & = & 1-F(l). \end{array}$$
Donc il suffit de calculer $F(l)$.
$$\begin{array}{rcl} F(l) & = & \sum_{k=1}^{l} I\!P(X=k) = \sum_{k=1}^{l}(1-p)^{k-1}p \\ & = & p\sum_{k=1}^{l}(1-p)^{k-1} = p\sum_{k=0}^{l-1}(1-p)^{k} \\ & = & p\frac{1-(1-p)^l}{1-(1-p)} = p\frac{1-(1-p)^l}{p} \\ & = & 1-(1-p)^l \end{array}$$
d'où,
$$I\!P(X>l) = 1-(1-(1-p)^l) = (1-p)^l.$$
5. Le nombre de transmissions moyen nécessaire à la réception du message est égal à 4 donc $X$ suit la loi géométrique de paramètre $\frac{1}{4}$, et comme la probabilité que la connexion soit perdue est $I\!P(X > l) = 0,1$, on a donc,
$$\begin{array}{rcl} I\!P(X>l) & = & (1-p)^l = (1-0,25)^l = 0,1 \\ & \Rightarrow & \ln((1-0,25)^l) = \ln(0,1) \\ & \Rightarrow & l.\ln(0,75) = \ln(0,1) \\ & \Rightarrow l = & \frac{\ln(0,1)}{\ln(0,75)} = 8. \end{array}$$

## Variables aléatoires à densité

**Définition 1.2.5** Soit $f$ une fonction continue sauf en un nombre fini de points et $X$ une v.a.r.. On dit que $X$ suit une loi diffuse (absolument continue) de densité $f$ si sa loi $\mathbb{P}_X$ est de densité $f$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$ ($\mathbb{P}_X = f dx$), c'est-à-dire

$$\mathbb{P}(X \in A) = \int_A f(x)dx, \qquad \forall A \in \mathscr{B}(\mathbb{R}).$$

**R** $\int_{\mathbb{R}} f(x)dx = \mathbb{P}(X \in \mathbb{R}) = 1$ et pour tous $\alpha \leq \beta$, on a

$$\begin{aligned} \mathbb{P}(\alpha \leq X \leq \beta) &= \mathbb{P}(\alpha \leq X < \beta) = \mathbb{P}(\alpha < X \leq \beta) \\ &= \mathbb{P}(\alpha < X < \beta) = \int_\alpha^\beta f(x)dx. \end{aligned}$$

En particulier, pour $\alpha = \beta$, on a $\mathbb{P}(X = \alpha) = 0$. Par conséquent,

- $\mathbb{P}(X \leq \alpha) = \mathbb{P}(X < \alpha) = \int_{-\infty}^{\alpha} f(x)dx.$
- $\mathbb{P}(X \geq \alpha) = \mathbb{P}(X > \alpha) = \int_{\alpha}^{+\infty} f(x)dx.$

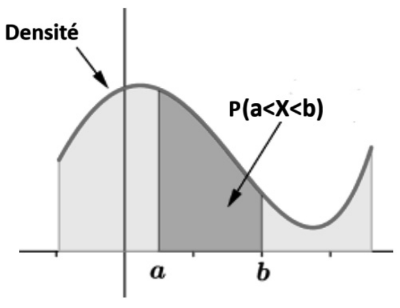


FIGURE 1.2 – La probabilité $\mathbb{P}(a \leq X \leq b)$

**Théorème. 1.2.3** $f : \mathbb{R} \to \mathbb{R}$ est une densité de probabilité si, et seulement si,

1. $f$ est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini des points.
2. $f(x) \geq 0, \quad \forall x \in \mathbb{R}.$
3. $\int_{-\infty}^{+\infty} f(x)dx = 1.$

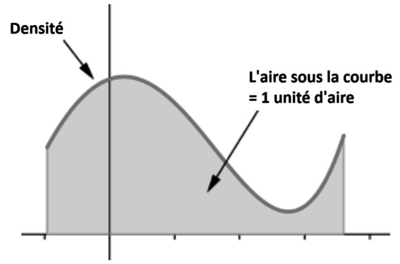


FIGURE 1.3 – Densité de probabilité

Si $f$ est la densité d'une variable $X$ alors elle est reliée à $F_X$, la fonction de répartition de $X$, de la façon suivante.

**Proposition 1.2.4** Soit $X$ une variable aléatoire de densité $f$, alors sa fonction de répartition $F_X$ vérifie :

1. pour tout $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f(t)dt$.
2. $F_X$ est continue sur $\mathbb{R}$.
3. Si $f$ est continue en $x_0$, alors $F_X$ est dérivable en $x_0$ et $F'_X(x_0) = f(x_0)$.

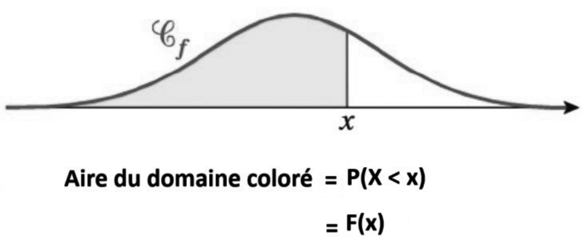


FIGURE 1.4 – Fonction de répartition

**Exercice 1.3** Vérifier que $f_1$ et $f_2$ sont des densités de probabilités et tracer leurs courbes représentatives :

$$f_1(x) = \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } 0 \leq x \leq 1 \\ 0 \text{ si } x > 1 \end{cases} \qquad f_2(x) = \begin{cases} 0 \text{ si } x < -1 \\ \text{x+1 si } -1 \leq x \leq 0 \\ \text{-x+1 si } 0 \leq x \leq 1 \\ 0 \text{ si } x > 1 \end{cases}$$

**Solution 1.3** 1. $f_1(x) \geq 0$, la fonction $f_1$ est continue par morceaux sur $\mathbb{R}$ et,

$$\int_{\mathbb{R}} f_1(x)dx = \int_{-\infty}^{0} f_1(x)dx + \int_0^1 f_1(x)dx + \int_1^{+\infty} f_1(x)dx = 0 + [x]_0^1 + 0 = 1.$$

2. $f_2(x) \geq 0$, la fonction $f_2$ est continue par morceaux sur $\mathbb{R}$ et,

$$\begin{aligned}\int_{\mathbb{R}} f_2(x)dx &= \int_{-\infty}^{-1} f_2(x)dx + \int_{-1}^{0} f_2(x)dx + \int_0^1 f_2(x)dx + \int_1^{+\infty} f_2(x)dx \\ &= 0 + [\frac{x^2}{2} + x]_{-1}^0 + [-\frac{x^2}{2} + x]_0^1 + 0 = 1.\end{aligned}$$

■ **Exemple 1.7** Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f(x) = \cos(x)\mathbf{1}_{]0,\frac{\pi}{2}]}(\mathbf{x})$. Il est clair que $f$ est positive sur $\left]0, \frac{\pi}{2}\right]$ et continue sauf en 0. De plus, on a

$$\begin{aligned}\int_{-\infty}^{+\infty} f(t)dt &= \int_{-\infty}^{0} f(t)dt + \int_0^{\frac{\pi}{2}} f(t)dt + \int_{\frac{\pi}{2}}^{+\infty} f(t)dt \\ &= \int_0^{\frac{\pi}{2}} \cos(t)dt = \sin\left(\frac{\pi}{2}\right) - \sin(0) = 1.\end{aligned}$$

Donc $f$ est bien une densité de probabilité. Cherchons maintenant la fonction de répertition de $X$. Pour $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f(t)dt$. On a

* Si $x < 0$ alors $F_X(x) = 0$.
* si $0 \leq x < \frac{\pi}{2}$ alors $F_X(x) = \int_0^x f(t)dt = \int_0^x \cos(t)dt = \sin(x)$.
* si $x \geq \frac{\pi}{2}$ alors $F(x) = \int_0^{\frac{\pi}{2}} \cos(t)dt = 1$.

■

**Exercice 1.4** On considère une variable aléatoire $X$ de densité

$$f(x) = \frac{c}{1+x^2}.$$

1. Pour quelle(s) valeur(s) de $c$ la fonction $f$ est-elle bien une densité ?
2. Calculer la fonction de répartition de $X$.
3. Déterminer $\mathbb{P}(X < 0)$ et $\mathbb{P}(-1 < X < 1)$.
   On rappelle que

$$\arctan(0) = 0, \arctan(1) = \frac{\pi}{4} \text{ et } \lim_{x \to +\infty} \arctan(x) = \frac{\pi}{2}.$$

**Solution 1.4** 1. Il faut que $c \geq 0$ et choisir la valeur de c de la manière suivante :

$$\begin{aligned}c\int_{-\infty}^{+\infty} \frac{dx}{(1+x^2)} = 1 &\iff c\Big[\arctan(x)\Big]_{-\infty}^{+\infty} = 1 \\ &\iff c\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = 1 \\ &\iff c = \frac{1}{\pi}\end{aligned}$$

La fonction $f$ est donc une densité si et seulement si $c = \frac{1}{\pi}$.

2. Calculons la fonction de répartition

$$\begin{aligned} F(x) &= \frac{1}{\pi}\int_{-\infty}^{x} \frac{dt}{1+t^2} = \frac{1}{\pi}\Big[\arctan(x)\Big]_{-\infty}^{x} \\ &= \frac{1}{\pi}\Big(\arctan(x) - \Big(-\frac{\pi}{2}\Big)\Big) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}. \end{aligned}$$

3. On a, $I\!\!P(X < 0) = I\!\!P(X \leq 0) = F(0) = \frac{1}{2}$

$$\text{et} \quad I\!\!P(-1 < X < 1) = F(1) - F(-1) = \frac{2}{\pi}\arctan(1) = \frac{1}{2}.$$

**Exercice 1.5** On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} cx^2 & \text{si } 0 \leq x \leq 3 \\ 0 & \text{sinon} \end{cases}$$

1. Evaluer la constante c pour que f soit une densité de probabilité.
2. Calculer et tracer la fonction de répartition de X.
3. Calculer $I\!\!P(1 < X < 2)$.
4. Déterminer l'espérance et la variance de X.
5. Pour tout $n \in \mathbb{N}^*$, déterminer le moment d'ordre n de X.

**Solution 1.5**

1. Il suffit de calculer $c\int_0^3 x^2 dx = 9c$. Donc $f$ est une densité de probabilité si et seulement si $c = \frac{1}{9}$.
2. La fonction de répartition est donnée par

$$F_X(x) = I\!\!P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{x^3}{27} & \text{si } 0 \leq x < 3 \\ 1 & \text{si } x \leq 3 \end{cases}$$

3. Comme X est absolument continue, on a

$$I\!\!P(1 < X < 2) = F_X(2) - F_X(1) = \frac{7}{27}.$$

4. On a, $I\!\!E[X] = \int_0^3 \frac{x^3}{9}dx = \frac{9}{4}$ et $I\!\!E[X^2] = \int_0^3 \frac{x^4}{9}dx = \frac{27}{5}$.
   Ainsi, $V(X) = \frac{27}{80}$.
5. Le moment d'ordre n est donné par $I\!\!E(X^n) = \int_0^3 \frac{x^{n+2}}{9}dx = \frac{3^{n+1}}{n+3}$.

■ **Exemple 1.8** 1. Soit $I$ un intervalle non vide d'extrémité $a,b$ $(a<b)$. On dit que $X$ suit la loi uniforme sur $I$ si $X$ est à densité

$$f(x)=\frac{1}{b-a}\mathbf{1}_I(x).$$

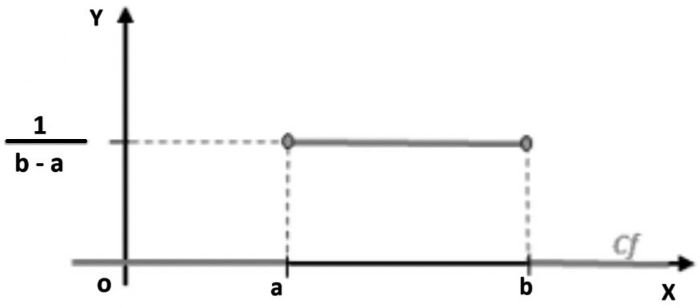


FIGURE 1.5 – Loi uniforme sur $[a,b]$

2. On dit que $X$ suit la loi exponentielle de paramètre $\lambda>0$, et note $X\sim\mathscr{E}(\lambda)$ si $X$ est à densité

$$f(x)=\lambda e^{-\lambda x}\mathbf{1}_{\mathbb{R}_+}(x).$$

La loi exponentielle possède la propriété d'absence de mémoire ou sans vieillissement.

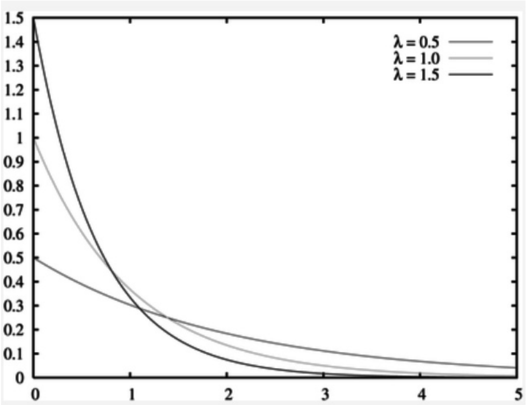


FIGURE 1.6 – Loi exponentielle

3. $X$ suit la loi de Cauchy de paramètre $c > 0$ si $X$ est à densité

$$f(x) = \frac{c}{\pi(x^2 + c^2)}.$$

4. On dit que $X$ suit la loi normale de paramètre $m \in \mathbb{R}$ et $\sigma^2$ ($\sigma > 0$) et on note $X \sim \mathcal{N}(m, \sigma^2)$, si $X$ est à densité

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

c'est la loi la plus importante en théorie des probabilités. Sa densité est la fameuse courbe en cloche. Voici des exemples de courbes pour quelques valeurs de $m$ et $\sigma$ :

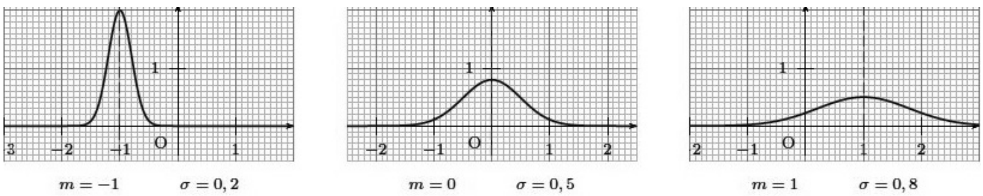


FIGURE 1.7 – Courbes de la loi normale

La loi normale est celle qui rend compte de diverses mesures d'une grandeur donnée, opérées à diverses reprises, chaque mesure étant sujette à des erreurs. La loi normale (ou de Laplace-Gauss, appelée « normale » par Pearson en 1893) est la loi de certains phénomènes continus qui fluctuent autour d'une valeur moyenne $\mu$, de manière aléatoire, résultante d'un grand nombre de causes indépendantes dont les effets s'ajoutent sans que l'un d'eux soit dominant : par exemple la taille d'un individu en cm, influencée par le sexe, la nourriture, l'environnement, l'hérédité, le lieu géographique . . .

■

**R**

1. Si $X \sim \mathcal{N}(m, \sigma^2)$ alors $Y = \dfrac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$ : loi normale centrée réduite.
2. Si $X \sim \mathcal{N}(0, 1)$ alors $\mathbb{P}_X = \mathbb{P}_{-X}$, on dit que $X$ est symétrique.
3. La courbe admet comme axe de symétrie la droite d'équation $x = m$.
4. Le maximum de la courbe est atteint en $m$, espérance de la variable $X$ (ce maximum valant $\frac{1}{\sigma\sqrt{2\pi}}$).
5. Plus $\sigma$ est grand, plus la courbe s'étale autour de la moyenne, en accord avec la signification de l'écart-type.

■ **Exemple 1.9** On note $Z \sim \mathcal{N}(0,1)$ de densité de probabilité $f$ et fonction de répartition notée $F$ alors

1. $F(z) = \mathbb{P}(Z < z) = \int_{-\infty}^{z} f(t)dt$.

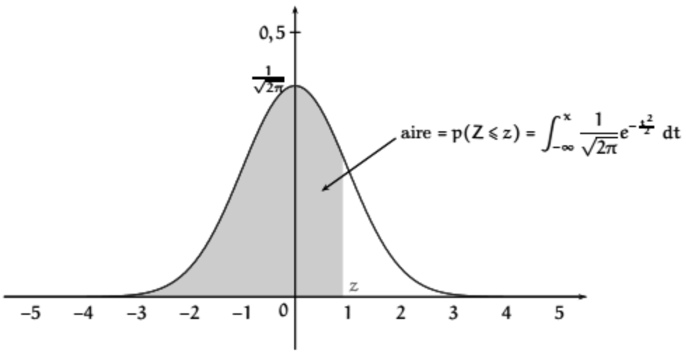


FIGURE 1.8 – La probabilité $\mathbb{P}(Z \leq z)$

2. $\mathbb{P}(Z \geq z) = 1 - \mathbb{P}(Z < z) = 1 - F(z)$.
3. $\mathbb{P}(a \leq Z \leq b) = F(b) - F(a)$.

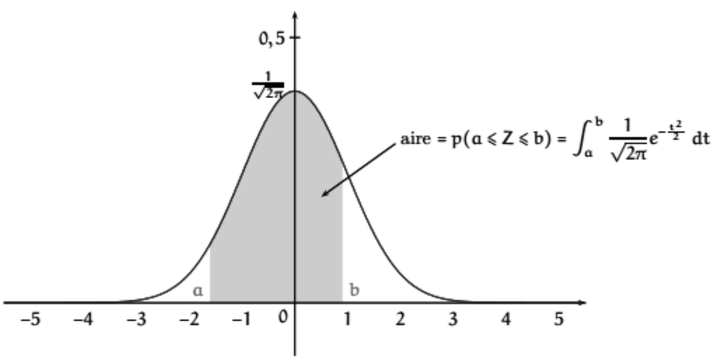


FIGURE 1.9 – La probabilité $\mathbb{P}(a \leq Z \leq b)$

4. $F(-z) = \mathbb{P}(Z < -z) = \mathbb{P}(Z > z) = 1 - \mathbb{P}(Z < z) = 1 - F(z)$.
5. $\mathbb{P}(-z \leq Z \leq z) = F(z) - F(-z) = F(z) - (1 - F(z)) = 2F(z) - 1$.

La table de la loi normale ne donne que les valeurs de la loi normale centrée réduite et pour des valeurs positives. En voici quelques exemples pour comprendre la méthode de lecture :

1. Calcul de $\mathbb{P}(Z \geq 1,25)$ : le nombre situé à l'intersection de la colonne $0,05$