

Chapitre 1

Rappels

L'objectif de ce chapitre est de rappeler les principales définitions et notations du calcul des probabilités. Les propriétés essentielles seront citées mais les preuves seront omises. Des références seront données à la fin de ce chapitre pour que le lecteur puisse trouver facilement le complément qui lui manque. Nous nous sommes restreints à des rappels "minimaux" afin d'éviter les développements certes intéressants mais qui ne seraient pas utilisés directement dans la suite du livre. Pour certaines notions, par exemple celle de convergence des variables aléatoires, il nous a paru plus judicieux de les faire figurer dans le chapitre où elles jouent un rôle majeur.

1.1 Espace de probabilité

On commence par introduire brièvement quelques notions de mesurabilité. Ω désigne un ensemble non vide et $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω .

Définition 1.1 Une **tribu** ou **σ -algèbre** est un sous-ensemble \mathcal{A} de $\mathcal{P}(\Omega)$ vérifiant :

1. $\Omega \in \mathcal{A}$;
2. \mathcal{A} est stable par passage au complémentaire : si $A \in \mathcal{A}$ alors $A^c = \bar{A} = \Omega \setminus A \in \mathcal{A}$;
3. \mathcal{A} est stable par réunion dénombrable : si $(A_n)_{n \geq 1}$ est une suite d'éléments de \mathcal{A} , alors $\bigcup_{n \geq 1} A_n \in \mathcal{A}$.

Les **événements** sont les éléments de \mathcal{A} .

Soit \mathcal{A} une tribu. Il est facile de montrer :

1. $\emptyset \in \mathcal{A}$;
2. \mathcal{A} est stable par intersection dénombrable : si $(A_n)_{n \geq 1}$ est une suite d'éléments de \mathcal{A} , alors $\bigcap_{n \geq 1} A_n \in \mathcal{A}$.
3. Soit $S \subset \mathcal{P}(\Omega)$, alors il existe une plus petite tribu (au sens de l'inclusion) notée $\sigma(S)$ qui contient S . $\sigma(S)$ est la **tribu engendrée** par S .

En particulier lorsque $\Omega = \mathbb{R}^d$ et S est l'ensemble des ouverts de \mathbb{R}^d la tribu $\sigma(S)$ est la **tribu borelienne** et est notée $\mathcal{B}(\mathbb{R}^d)$. Lorsque $d = 1$, alors $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I})$, où \mathcal{I} est l'ensemble des intervalles ouverts (resp. semi-ouverts, fermés) de \mathbb{R} .

Définition 1.2 1. Une application $X : (\Omega, \mathcal{A}) \rightarrow \mathbb{R}^d$ est une **variable aléatoire** (en abrégé **v.a.**) si X est **mesurable** :

$$X^{-1}(\Gamma) \in \mathcal{A} \quad \text{pour tout } \Gamma \in \mathcal{B}(\mathbb{R}^d),$$

où $X^{-1}(\Gamma)$ est l'image réciproque de Γ par X :

$$X^{-1}(\Gamma) = \{\omega \in \Omega; X(\omega) \in \Gamma\}.$$

Lorsque $d = 1$, on dit que X est une **variable aléatoire réelle** (v.a.r.) .

2. Une v.a. X est dite **étagée** ou **élémentaire** si :

$$X = \sum_{k=1}^n x_k 1_{A_k} \quad (A_k \in \mathcal{A}),$$

où 1_{A_k} représente la **fonction indicatrice** de A_k . En particulier si les A_k sont deux à deux disjoints, X prend la valeur x_k sur A_k .

3. Soit X une v.a. définie sur (Ω, \mathcal{A}) . La **tribu engendrée** par X est la σ -algèbre :

$$\sigma(X) = \{X^{-1}(A), A \in \mathcal{B}(\mathbb{R}^d)\}.$$

On montre sans difficulté :

1. Les sommes, différences, produits et quotients (lorsque c'est possible) de v.a. restent des v.a.
2. La composition d'une v.a. par une "bonne" fonction reste une v.a. Ainsi si X est une v.a. et f est continue, alors $f \circ X$ est une v.a.
3. les limites de suites de v.a. (lorsque ces limites existent) sont encore des v.a.
4. Par définition la tribu $\sigma(X)$ engendrée par une v.a. X définie sur (Ω, \mathcal{A}) est incluse dans \mathcal{A} . Soit Y une v.a. : $(\Omega, \mathcal{A}) \rightarrow \mathbb{R}^n$. Alors Y est $\sigma(X)$ -**mesurable** si et seulement s'il existe $f : \mathbb{R}^d \rightarrow \mathbb{R}^n$ (mesurable) telle que $Y = f(X)$.

Dans la suite on supposera que Ω est un ensemble non-vide et que \mathcal{A} est une tribu sur Ω .

Définition 1.3 1. Une **mesure positive** est une application $P : \mathcal{A} \rightarrow [0, \infty]$ vérifiant

$$(a) P(\emptyset) = 0;$$

(b) P est σ -**additive** : pour toute suite $(A_n)_{n \geq 1}$ d'éléments de \mathcal{A} deux à deux disjoints (i.e. $A_n \cap A_m = \emptyset$ si $n \neq m$) alors

$$P\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} P(A_n). \quad (1.1)$$

Dans la suite une mesure sera toujours sous-entendue être positive.

2. Une mesure P est **bornée** si $P(\Omega) < \infty$.

3. Une mesure P est σ -**finie** s'il existe une suite $(A_n)_{n \geq 1}$ d'éléments de \mathcal{A} telle

$$\Omega = \bigcup_{n \geq 1} A_n \quad \text{et } A_n \subset A_{n+1} \quad \text{et } P(A_n) < \infty \quad \text{pour tout } n.$$

Définition 1.4 1. Une **probabilité** est une mesure positive telle que $P(\Omega) = 1$.

2. Un **espace de probabilité** est un triplet (Ω, \mathcal{A}, P) où P est une probabilité sur (Ω, \mathcal{A}) .

Listons les principales propriétés du calcul des probabilités :

1. $P(\emptyset) = 0$.
2. $P(\bar{A}) = P(A^c) = 1 - P(A)$.
3. Si A_1, \dots, A_n sont n éléments de \mathcal{A} deux à deux disjoints alors

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k).$$

4. Si $A, B \in \mathcal{A}$, on a :

$$P(A) + P(B) = P(A \cup B) + P(A \cap B),$$

$$A \subset B \Rightarrow P(A) \leq P(B).$$

5. (continuité monotone) Soit $(A_n)_{n \geq 1}$ (resp. $(B_n)_{n \geq 1}$) une suite croissante (resp. décroissante) d'événements, i.e. $A_n \subset A_{n+1}$ pour tout n (resp. $B_{n+1} \subset B_n$ pour tout n) alors $n \mapsto P(A_n)$ est croissante (resp. $n \mapsto P(B_n)$ est décroissante) et :

$$P\left(\bigcup_{n \geq 1} A_n\right) = \lim_{n \rightarrow \infty} P(A_n), \quad P\left(\bigcap_{n \geq 1} B_n\right) = \lim_{n \rightarrow \infty} P(B_n).$$

6. Soit (A_n) une suite finie ou dénombrable d'éléments de \mathcal{A} alors :

$$P\left(\bigcup_n A_n\right) \leq \sum_n P(A_n).$$

Nous allons introduire les ensembles de probabilité nulle. Cette notion joue un rôle important par exemple dans la Section 2.4.

Définition 1.5 Soit (Ω, \mathcal{A}, P) un espace de probabilité.

1. Une partie N de Ω (attention N n'est pas nécessairement dans \mathcal{A}) est dite **négligeable** ou de **probabilité nulle** s'il existe $A \in \mathcal{A}$ tel que : $N \subset A$ et $P(A) = 0$.
2. Une propriété $\Pi(\omega)$ qui dépend de ω est dite **vraie presque sûrement** si l'ensemble des ω pour lesquels $\Pi(\omega)$ est fausse est négligeable.
3. L'espace de probabilité (Ω, \mathcal{A}, P) est **complet** si tous les ensembles négligeables appartiennent à \mathcal{A} .

Ainsi deux applications $X, Y : \Omega \rightarrow \mathbb{R}$ sont égales presque sûrement si $\{\omega; X(\omega) \neq Y(\omega)\}$ est négligeable.

Donnons quelques propriétés des ensembles négligeables :

1. \emptyset est négligeable.
2. Si $N \subset N'$ et N' est négligeable alors N est négligeable.
3. Soit $(N_k)_{k \in I}$ une famille finie ou dénombrable d'ensembles négligeables alors $\bigcup_{k \in I} N_k$ est négligeable.

On peut toujours supposer, sans perte de généralité, que l'espace de probabilité (Ω, \mathcal{A}, P) est complet. Cette hypothèse sera toujours supposée dans la suite.

Remarque 1.6 Soit P une mesure sur (Ω, \mathcal{A}) . Alors P vérifie les propriétés 1., 3., 4., 5. (avec les A_n) et 6. énoncées à la suite de la définition 1.4. On peut aussi définir la notion d'ensemble P négligeable. Les propriétés 1., 2. et 3. énoncées après la définition 1.5 restent vraies.

1.2 Loi d'une variable aléatoire

Définition 1.7 1. Soit X une v.a. définie sur l'espace de probabilité (Ω, \mathcal{A}, P) et à valeurs dans \mathbb{R}^d . La loi (ou la **distribution**) de X est la probabilité P_X sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$:

$$P_X(A) := P(X^{-1}(A)) = P(\{\omega; X(\omega) \in A\}), \quad \text{pour tout } A \in \mathcal{B}(\mathbb{R}^d). \quad (1.2)$$

On notera dans la suite $X \sim P_X$ pour signifier que X suit la loi P_X .

2. Deux v.a. X et X' à valeurs dans \mathbb{R}^d , définies sur (Ω, \mathcal{A}, P) resp. $(\Omega', \mathcal{A}', P')$, ont même loi (ou sont **identiques en loi**) si $P_X = P_{X'}$. Ce qui signifie :

$$P(X \in A) = P'(X' \in A), \quad \text{pour tout } A \in \mathcal{B}(\mathbb{R}^d). \quad (1.3)$$

Définition 1.8 La fonction de répartition F_X d'une v.a.r. X est l'application :

$$F_X(t) = P(X \leq t) = P_X([-\infty, t]), \quad \text{pour tout } t \in \mathbb{R},$$

où P_X désigne la loi de X .

Énonçons quelques propriétés des fonctions de répartition :

1. La fonction de répartition F_X d'une v.a.r. X est croissante, à valeurs dans $[0, 1]$, continue à droite et

$$\lim_{t \rightarrow -\infty} F_X(t) = 0, \quad \lim_{t \rightarrow \infty} F_X(t) = 1, \quad (1.4)$$

$$P_X([a, b]) = P(X \in]a, b]) = F_X(b) - F_X(a), \quad a < b.$$

2. La fonction de répartition caractérise la loi. Soient X et Y deux v.a.r.. Alors elles ont la même loi (i.e. $P_X = P_Y$) si et seulement si $F_X(t) = F_Y(t)$, pour tout $t \in \mathbb{R}$. Ce qui revient à prendre $A =]-\infty, t]$ dans l'identité (1.3).

3. Soit $F : \mathbb{R} \rightarrow [0, 1]$ une fonction croissante, continue à droite et telle que (1.4) ait lieu. On choisit $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}([0, 1])$ et P la mesure de Lebesgue sur $[0, 1]$. Alors il existe une v.a.r. X définie sur l'espace de probabilité $([0, 1], \mathcal{B}([0, 1]), P)$, dont la fonction de répartition F_X est égale à F .

4. La propriété 2. précédente se généralise au cas de plusieurs variables aléatoires. Soient X et Y deux v.a. à valeurs dans \mathbb{R}^n , X de composantes X_1, \dots, X_n et Y de composantes Y_1, \dots, Y_n . Alors X et Y ont même loi ssi

$$P(X_1 \leq t_1, \dots, X_n \leq t_n) = P(Y_1 \leq t_1, \dots, Y_n \leq t_n)$$

pour tout $t_1, \dots, t_n \in \mathbb{R}$.

Définition 1.9 Soit $\alpha \in]0, 1[$. Le **quantile** q_α d'ordre α est le plus petit réel x tel que $P(X \leq x) \geq \alpha$. La **médiane** est le quantile d'ordre $1/2$.

Il est facile de montrer que $F_X(q_\alpha -) = P(X < q_\alpha) \leq \alpha \leq F_X(q_\alpha) = P(X \leq q_\alpha)$. En pratique, lorsque la fonction de répartition F_X est continue et strictement croissante, alors le quantile d'ordre α est unique et est caractérisé par la relation : $F_X(q_\alpha) = \alpha$. Ce qui se produit dans le cas où X admet une densité strictement positive sur un intervalle (voir la Section 1.2.2). Dans le cas discret le calcul des quantiles est explicité dans la Section 1.2.1.

Dans le livre nous utiliserons essentiellement des v.a. qui sont discrètes ou à densité (voir plus bas). Toutefois, il existe des v.a. qui ne sont pas de ce type. Nous donnerons dans la Section 4.5 un exemple de v.a.r. dont la loi est singulière.

1.2.1 Lois discrètes

Définition 1.10 Une v.a. X est dite **discrète** si elle prend un nombre fini ou dénombrable de valeurs.

Soit X une v.a. discrète et $X(\Omega) = \{x_i; i \in I\}$ l'ensemble des valeurs prises par X . L'ensemble I est fini ou dénombrable. On suppose $x_i \neq x_j$ pour tout $i \neq j$. On peut écrire X sous la forme :

$$X = \sum_{i \in I} x_i 1_{A_i}, \quad \text{avec } A_i = X^{-1}(\{x_i\}) = \{X = x_i\}.$$

Notons que les événements A_i sont deux à deux disjoints et que leur réunion vaut Ω . Avec ces notations, la loi P_X de X est alors donnée par la formule :

$$P_X = \sum_{i \in I} p_i \delta_{x_i}, \quad (p_i := P(X = x_i)),$$

où δ_a désigne la **mesure de Dirac au point a** .

Donnons quelques propriétés des lois discrètes.

1. La loi de X est caractérisée par $\{x_i, i \in I\}$ et $(p_i)_{i \in I}$ où $p_i := P(X = x_i), i \in I$.
2. Réciproquement, soient $\{x_i, i \in I\}$ une famille de réels finie ou dénombrable, et $(p_i)_{i \in I}$ une famille de réels appartenant à $[0, 1]$ et telle que : $\sum_{i \in I} p_i = 1$. Alors il

existe une v.a.r. discrète X de loi $P_X = \sum_{i \in I} p_i \delta_{x_i}$.

Supposons que X prenne n valeurs distinctes ordonnées : $x_1 < \dots < x_n$. Alors :

$$F_X(t) = \begin{cases} 0 & \text{lorsque } t < x_1 \\ p_1 + \dots + p_i & \text{lorsque } x_i \leq t < x_{i+1} \text{ et } i < n \\ 1 & \text{lorsque } x_n \leq t. \end{cases}$$

De plus $F_X(x_i) - F_X(x_i-) = p_i$, où $F_X(x_i-)$ désigne la limite à gauche de F_X en x_i .

On vérifie facilement dans ce cas que F_X caractérise la loi de X : connaissant F_X on peut déterminer de manière unique x_1, \dots, x_n et p_1, \dots, p_n .

Le calcul des quantiles ne pose pas de difficulté, compte tenu de la Définition 1.9, on a :

$$q_\alpha = \begin{cases} x_1 & \text{lorsque } 0 < \alpha \leq p_1 \\ x_i & \text{lorsque } p_1 + \dots + p_{i-1} < \alpha \leq p_1 + \dots + p_i \text{ et } i < n \\ x_n & \text{lorsque } p_1 + \dots + p_{n-1} < \alpha < 1. \end{cases}$$

Nous allons à présent passer en revue les principales lois discrètes.

a) Loi de Bernoulli de paramètre $p \in [0, 1]$. Cette loi ne charge que 0 et 1 et s'écrit $p\delta_1 + (1-p)\delta_0$. Une **v.a. de Bernoulli** X a pour distribution la loi de Bernoulli : X ne prend que les valeurs 0 et 1 et

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

Lorsque $p = 1$ (resp. $p = 0$), alors X est (p.s.) constante : $X = 1$ (resp. $X = 0$).

On considère une expérience à deux issues possibles : succès ou échec. On suppose que la probabilité de succès vaut p (la probabilité d'échec vaut donc $1 - p$). Soit X la v.a. qui vaut 1 (resp. 0) lorsque l'expérience se solde par un succès (resp. échec). Alors X suit la loi de Bernoulli de paramètre p .

b) **Lois binomiales** de paramètres n et p . Soit n un entier plus grand que 1 et $0 < p < 1$. La loi binomiale de paramètres n et p est à support dans $\{0, 1, \dots, n\}$ et est égale à

$$\sum_{k=0}^n C_n^k p^k (1-p)^{n-k} \delta_k, \quad \text{avec } C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Cette loi est notée $\mathcal{B}(n, p)$. Par conséquent X suit la loi $\mathcal{B}(n, p)$ si X est à valeurs dans $\{0, 1, \dots, n\}$ et $P(X = k) = C_n^k p^k (1-p)^{n-k}$ pour tout entier k compris entre 0 et n . On dit aussi que X est une **v.a. binomiale**.

Il est facile de donner une réalisation probabiliste de la loi binomiale. On considère une expérience à deux issues possibles : succès et échec. On suppose que la probabilité de succès vaut p . On répète n fois cette expérience et de manière indépendante à chaque fois. On désigne par X_i la v.a. qui vaut 1 (resp. 0) si la i -ième expérience s'est soldée par un succès (resp. échec). Soit X le nombre de succès obtenus après avoir effectué n

expériences. Il est clair que : $X = \sum_{i=1}^n X_i$ et que les X_i sont indépendantes (cette notion sera vue plus bas dans la Section 1.3) et de même loi de Bernoulli de paramètre p . Alors X suit la loi $\mathcal{B}(n, p)$. Remarquons que la loi binomiale $\mathcal{B}(1, p)$ est la loi de Bernoulli de paramètre p .

c) **Lois géométriques** de paramètre $p \in]0, 1[$. La loi géométrique de paramètre p est à support dans les entiers plus grands que 1, et est égale à $\sum_{k \geq 1} p(1-p)^{k-1} \delta_k$. On

note $\mathcal{G}(p)$ cette distribution. Par définition une **v.a. géométrique** est une v.a. de loi géométrique : X est à valeurs dans $\{1, 2, 3, \dots\}$ et $P(X = k) = p(1-p)^{k-1}$, $k \geq 1$.

Comme dans le cas de la loi binomiale il est facile de construire une v.a. X de loi $\mathcal{G}(p)$. On reprend la modélisation stochastique de la loi binomiale. Soit X la première fois où un succès est obtenu. Alors X suit la loi $\mathcal{G}(p)$.

d) **Lois de Poisson**. La loi de Poisson de paramètre $\lambda > 0$ est à support dans les entiers et est égale à $\sum_{k \geq 0} \frac{\lambda^k}{k!} e^{-\lambda} \delta_k$. On note $\mathcal{P}(\lambda)$ cette distribution. Une **v.a. de Poisson** est

à valeurs dans \mathbb{N} et vérifie : $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$, $k \geq 0$.

1.2.2 Lois à densité

Définition 1.11 Soit X une v.a. à valeurs dans \mathbb{R}^d . On dit que X est une **v.a. à densité** s'il existe une fonction f borelienne, (presque sûrement) positive ou nulle telle que :

$$P_X(A) = P(X \in A) = \int_{\mathbb{R}^d} 1_A(x) f(x) dx, \quad \forall A \in \mathcal{B}(\mathbb{R}^d).$$

La fonction f est unique (à une égalité presque sûre près) et $\int_{\mathbb{R}^d} f(x) dx = 1$. On dit que f est la **densité** de X .

Par conséquent, si X est à densité, sa loi est la mesure de probabilité sur \mathbb{R}^d : $f(x) dx$. La loi de X est caractérisée par f . Réciproquement, soit $f : \mathbb{R}^d \rightarrow [0, \infty[$ une fonction (borelienne) telle que $\int_{\mathbb{R}^d} f(x) dx = 1$, alors il existe une v.a. X de densité f .

Supposons que X est une v.a. à densité f_X et à valeurs dans \mathbb{R} . Alors sa fonction de répartition F_X s'exprime à l'aide de f_X de la manière suivante :

$$F_X(t) = P(X \leq t) = \int_{-\infty}^t f_X(x) dx, \quad t \in \mathbb{R}.$$

Notons que dans ce cas F_X est continue : $F_X(t) = P(X < t)$, $\forall t \in \mathbb{R}$. Supposons que f_X soit continue sauf en un nombre fini de points t_1, \dots, t_n . Alors $f_X(t) = F'_X(t)$ pour tout $t \neq t_i$.

Nous listons plus bas les principales lois à densité que nous rencontrerons dans la suite de cet ouvrage.

a) Lois uniformes. On dit qu'une v.a.r. X suit la **loi uniforme sur l'intervalle** $[a, b]$ (avec $a < b$) si X admet pour densité la fonction $\frac{1}{b-a} 1_{[a,b]}(x)$. On note $\mathcal{U}([a, b])$ cette distribution. Sa fonction de répartition F_X est nulle sur $] -\infty, a]$ et vaut $F_X(x) = \frac{x-a}{b-a} 1_{\{a < x \leq b\}} + 1_{\{x > b\}}$ pour tout $x > a$.

Lorsque $a = 0$ et $b = 1$, on dit que X suit la **loi uniforme**. Si Y suit la loi uniforme, alors $a + (b-a)Y$ suit la loi uniforme sur l'intervalle $[a, b]$.

b) Lois exponentielles. Une v.a.r. X est dite de **loi exponentielle de paramètre** $\lambda > 0$, si X admet pour densité : $\frac{1}{\lambda} e^{-x/\lambda} 1_{\{x > 0\}}$. Cette loi sur \mathbb{R}_+ est notée $\mathcal{E}(\lambda)$. Il est facile de calculer la fonction de répartition d'une v.a. X de loi exponentielle de paramètre λ :

$$F_X(x) := P(X \leq x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - e^{-x/\lambda} & \text{si } x > 0. \end{cases}$$

Lorsque $\lambda = 1$, on dit que X suit la **loi exponentielle**. Il est aisé de vérifier que si $\lambda > 0$ et Y est de loi exponentielle, alors λY suit la loi exponentielle de paramètre λ .

c) Lois gaussiennes. i) Une v.a. réelle G suit la **loi gaussienne réduite et centrée** si G admet pour densité $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. On note $\mathcal{N}(0, 1)$ cette distribution de probabilité. Φ désigne la fonction de répartition de la loi gaussienne réduite et centrée :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt, \quad x \in \mathbb{R}.$$

Puisque la fonction $t \mapsto e^{-t^2/2}$ est une fonction paire, on en déduit facilement que Φ vérifie :

$$\Phi(-x) = 1 - \Phi(x), \quad \forall x \in \mathbb{R}.$$

Les valeurs de Φ sont tabulées. En particulier $\Phi(0) = 1/2$ et :

$$\Phi(1) = 0.841, \quad \Phi(2) = 0.977, \quad \Phi(3) = 0.999, \quad \Phi(1.64) = 0.950, \quad \Phi(1.96) = 0.975.$$

Les deux dernières valeurs intervenant en statistique (cf. le Chapitre 5).

De plus :

$$\Phi(x) = \frac{1}{2} \left(\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) + 1 \right), \quad \forall x \in \mathbb{R}. \quad (1.5)$$

L'intérêt de cette formule est que la fonction erf figure dans les logiciels de programmation et en particulier avec MAPLE. L'utilisation de tables n'est donc plus nécessaire.

Pour tout $\alpha \in]0, 1[$, on désigne par z_α le **quantile d'ordre** α de la loi gaussienne réduite et centrée, ce réel est défini de manière unique par la relation $\Phi(z_\alpha) = \alpha$. Puisque G a même loi que $-G$ on a :

$$P(-z_{(1+\alpha)/2} \leq G \leq z_{(1+\alpha)/2}) = \Phi(z_{(1+\alpha)/2}) - \Phi(-z_{(1+\alpha)/2}) = \alpha.$$

ii) Une v.a. X est **gaussienne de paramètres** m et σ , si $X = m + \sigma G$, avec G une v.a. de loi gaussienne réduite et centrée. On peut choisir $\sigma \geq 0$. De plus m (resp. σ^2) est la moyenne (resp. variance) de X . On note $\mathcal{N}(m, \sigma^2)$ la loi d'une v.a. gaussienne de moyenne m et de variance σ^2 . Lorsque $\sigma > 0$, X a pour densité :

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

iii) Soit X un vecteur aléatoire à valeurs dans \mathbb{R}^d , de coordonnées (X^1, \dots, X^d) .

X est un **vecteur gaussien de loi réduite et centrée** si les v.a.r. X^1, \dots, X^d sont indépendantes et suivent chacune une loi gaussienne réduite et centrée. X est un **vecteur gaussien** (ou une **v.a. gaussienne multidimensionnelle**) si pour tout vecteur u , le produit scalaire $\langle u, X \rangle$ est une v.a. réelle gaussienne. En particulier, pour tout i , la v.a.r. X^i est gaussienne. D'autres propriétés des vecteurs gaussiens seront données dans la Section 1.4.4.

d) **Lois gamma.** Une v.a. X suit la **loi gamma de paramètres** $a > 0$ et $b > 0$ si X a pour densité :

$$\frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b} \mathbf{1}_{\{x>0\}},$$

où $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, $x > 0$.

On note $\gamma(a, b)$ cette distribution sur \mathbb{R}_+ .

On appelle **loi gamma de paramètre** a la loi $\gamma(a) := \gamma(a, 1)$. Il est facile de se ramener à cette distribution ; en effet si X suit la loi $\gamma(a)$ alors bX suit la loi $\gamma(a, b)$.

Donnons deux exemples de lois gamma. Soit n est un entier strictement positif.

La loi $\gamma(n, a)$ est appelée **loi d'Erlang**. Sa densité est :

$$\frac{1}{a^n (n-1)!} x^{n-1} e^{-x/a} \mathbf{1}_{\{x>0\}}.$$

En particulier, si $n = 1$, la loi $\gamma(1, a)$ est égale à la loi exponentielle $\mathcal{E}(a)$.

La loi du **chi-deux** à n degrés de liberté, notée $\chi^2(n)$, est la loi $\gamma(n/2, 2)$. Pour tout $0 < \alpha < 1$, on note $\chi_\alpha^2(n)$ le quantile d'ordre α de la loi du $\chi^2(n)$. Ce réel positif est caractérisé par :

$$\frac{1}{2^{n/2} \Gamma(n/2)} \int_0^{\chi_\alpha^2(n)} x^{\frac{n}{2}-1} e^{-x/2} dx = \alpha.$$

Cette distribution intervient de manière essentielle en statistique (cf. le Chapitre 5).

1.3 Indépendance

Définition 1.12 1. Deux événements A et B sont **indépendants** si $P(A \cap B) = P(A)P(B)$.

2. Soient $\mathcal{A}_1, \dots, \mathcal{A}_n$ n tribus incluses dans \mathcal{A} . On dit que ces **tribus sont indépendantes** si pour tout $A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$ on a : $P(A_1 \cap \dots \cap A_n) = P(A_1) \times \dots \times P(A_n)$.