Chapitre 2

Les variables aléatoires et leurs caractéristiques

Dans ce chapitre, nous introduisons la notion fondamentale de variable aléatoire (réelle) qui jouera un rôle important dans la suite, aussi bien en théorie des probabilités qu'en statistique. Nous donnons en particulier des exemples classiques de variables discrètes et continues et nous introduisons certaines de leurs caractéristiques : fonction de répartition, densité, moyenne, variance, et autres moments.

2.1 Variables aléatoires

Nous commençons par donner ici les définitions d'une variable aléatoire et de la loi d'une variable aléatoire. Dans les prochains paragraphes, nous donnerons de nombreux exemples de telles variables dans les cas dicret et continu.

Définition 2.1.1. Une variable aléatoire X (réelle) est une application "mesurable" d'un espace de probabilités $(,\mathcal{F},\mathbb{P})$ dans l'ensemble $\mathbb{R}$ des nombres réels.

Remarque 2.1.2. Le terme "mesurable" indique que la fonction considérée doit "bien" se comporter vis à vis de la tribu $\mathcal{F}$. Précisément, l'image réciproque par X de tout intervalle doit être un élément de la tribu $\mathcal{F}$.

Exemple 2.1.3. Considérons un jeu de pile ou face avec une pièce équilibrée, que l'on modélise par un triplet $(,\mathcal{F},\mathbb{P})$ où $=\{\texttt{pile},\texttt{face}\}$, $\mathcal{F} \quad = \mathcal{P}()$, et $\mathbb{P}$ uniforme. Si on tombe sur `pile`, on gagne 10 euros, si on tombe sur `face` on perd 10 euros. Le gain G est une variable aléatoire. En effet, c'est une fonction définie sur l'ensemble Ω et à valeurs dans l'ensemble $\{-10, 10\} \subset \mathbb{R}$, avec

$$G(\texttt{pile}) = 10, \quad G(\texttt{face}) = -10.$$

Exemple 2.1.4. Considérons le jeu de deux dés, que l'on modélise par un triplet $(,\mathcal{F},\mathbb{P})$ où $\quad = \{1,2,3,4,5,6\}^2 = \{\omega = (\omega_1,\omega_2),\ \omega_1,\omega_2 \in \{1,2,3,4,5,6\}\}$, $\mathcal{F}$ est la tribu des parties $\mathcal{F} = \mathcal{P}(\Omega)$, et $\mathbb{P}$ uniforme. On note S la somme des deux dés.

Alors S est une variable aléatoire. C'est une fonction définie sur l'ensemble Ω et à valeurs dans l'ensemble $\{2, 3, \ldots, 12\} \subset \mathbb{R}$, avec

$$S(\omega) = S(\omega_1, \omega_2) := \omega_1 + \omega_2.$$

Si l'on dispose d'un espace de probabilités $(\Omega,\mathcal{F},\mathbb{P})$ et d'une variable aléatoire $X : \Omega \to X(\Omega) \subset \mathbb{R}$, on peut construire de façon naturelle une probabilité sur $X(\Omega)$, l'ensemble des valeurs prises par la fonction X.

Proposition 2.1.5. *Pour tout sous-ensemble "mesurable" B de $X()$, on définit :*

$$\mathbb{P}_X(B) = \mathbb{P}(\{\omega | X(\omega) \in B\}) = \mathbb{P}(\{X^{-1}(B)\}).$$

Ce faisant, on définit une probabilité sur $X(\Omega)$, appelée la loi de X.

Exemple 2.1.6. Considérons le jeu de pile ou face précédent où l'on gagne ou perd 10 euros selon que la pièce tombe sur `pile` ou `face`. Comme ci-dessus, on note G le gain après le lancer. La variable G définit une probabilité sur les gains possibles $G() = \{-10, 10\} \subset \mathbb{R}$:

$$\mathbb{P}_G(\{-10\}) := \mathbb{P}(G = -10) = \mathbb{P}(\{\omega,\ G(\omega) = -10\}) = \mathbb{P}(\texttt{face}) = 1/2,$$

$$\mathbb{P}_G(\{10\}) := \mathbb{P}(G = -10) = \mathbb{P}(\{\omega,\ G(\omega) = 10\}) = \mathbb{P}(\texttt{pile}) = 1/2.$$

Exemple 2.1.7. Considérons maintenant l'exemple précédent de la somme S de deux dés. On obtient alors une probabilité $\mathbb{P}_S$ sur l'ensemble $S() = \{2, 3, \ldots, 12\}$, avec

$$\begin{aligned}
&\mathbb{P}_S(\{2\}) := \mathbb{P}(\{\omega,\ S(\omega) = 2\}) = \mathbb{P}(\{(1,1)\}) = 1/36,\\
&\mathbb{P}_S(\{3\}) := \mathbb{P}(\{\omega,\ S(\omega) = 3\}) = \mathbb{P}(\{(1,2),(2,1)\}) = 2/36,\\
&\mathbb{P}_S(\{4\}) := \mathbb{P}(\{\omega,\ S(\omega) = 4\}) = \mathbb{P}(\{(1,3),(3,1),(2,2)\}) = 3/36,\\
&\mathbb{P}_S(\{5\}) := \mathbb{P}(\{\omega,\ S(\omega) = 5\}) = \mathbb{P}(\{(1,4),(4,1),(2,3),(3,2)\}) = 4/36,\\
&\mathbb{P}_S(\{6\}) := \mathbb{P}(\{\omega,\ S(\omega) = 6\}) = \mathbb{P}(\{(1,5),(5,1),(2,4),(4,2),(3,3)\}) = 5/36,\\
&\mathbb{P}_S(\{7\}) := \mathbb{P}(\{\omega,\ S(\omega) = 7\}) = \mathbb{P}(\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}) = 6/36,\\
&\mathbb{P}_S(\{8\}) := \mathbb{P}(\{\omega,\ S(\omega) = 8\}) = \mathbb{P}(\{(2,6),(6,2),(3,5),(5,3),(4,4)\}) = 5/36,\\
&\mathbb{P}_S(\{9\}) := \mathbb{P}(\{\omega,\ S(\omega) = 9\}) = \mathbb{P}(\{(3,6),(6,3),(4,5),(5,4)\}) = 4/36,\\
&\mathbb{P}_S(\{10\}) := \mathbb{P}(\{\omega,\ S(\omega) = 10\}) = \mathbb{P}(\{(4,6),(6,4),(5,5)\}) = 3/36,\\
&\mathbb{P}_S(\{11\}) := \mathbb{P}(\{\omega,\ S(\omega) = 11\}) = \mathbb{P}(\{(5,6),(6,5)\}) = 2/36,\\
&\mathbb{P}_S(\{12\}) := \mathbb{P}(\{\omega,\ S(\omega) = 12\}) = \mathbb{P}(\{(6,6)\}) = 1/36.
\end{aligned}$$

2.1.1 Variables aléatoires discrètes

On s'intéresse ici de plus près au cas de variables aléatoires discrètes qui constituent l'essentiel des variables que nous considérerons dans la suite.

Définition 2.1.8. On appelle variable aléatoire discrète une variable aléatoire $X : \to X()$ dont l'ensemble d'arrivée $X()$ est fini ou dénombrable.

Cela signifie de la fonction X ne peut prendre qu'un nombre fini ou dénombrable de valeurs $x_1, x_2, \ldots, x_n, \ldots$. Dans ce cas, la loi $\mathbb{P}_X$ de la variable X est caractérisée par la donnée des probabilités des singletons $\mathbb{P}_X(\{x_i\}) = \mathbb{P}(X = x_i) := p_i$, pour tout $i = 1, 2, \ldots$. Les nombres p_i verifient les propriétés du paragraphe 1.3.2 , à savoir : $p_i \geqslant 0$ pour tout i et $\sum_i p_i = 1$. La probabilité d'une partie mesurable A de $X(\Omega)$ est alors donnée par

$$\mathbb{P}(A) = \sum_{i, x_i \in A} p_i.$$

Exemples de variables discrètes

Voici quelques exemples classiques de loi discrètes.

Loi de Bernoulli : on dit qu'une variable aléatoire X suit une loi de Bernoulli de paramètre $p \in [0, 1]$ et on note $X \sim \mathcal{B}(p)$, si X est à valeurs dans l'ensemble $\{0, 1\}$ et

$$\mathbb{P}_X(\{0\}) = \mathbb{P}(X = 0) = 1 - p, \qquad \mathbb{P}_X(\{1\}) = \mathbb{P}(X = 1) = p.$$

C'est la loi d'un jet de pile ou face ou de n'importe quelle expérience aléatoire qui n'a que deux issues possibles.

Loi uniforme : on dit qu'une variable aléatoire X suit une loi uniforme sur un ensemble fini $E = \{x_1, \ldots, x_n\}$ et on note $X \sim U_E$, si X est à valeurs dans l'ensemble E et

$$\mathbb{P}(X = x_i) = \frac{1}{n}, \; i = 1, \ldots, n.$$

La loi uniforme est utilisée lorsque qu'aucun point de l'ensemble d'arrivée n'est privilégie : chacun a le même poids, ici $1/n$.

Loi binomiale : on dit qu'une variable aléatoire X suit une loi binomiale de paramètres (n, p) et on note $X \sim \mathcal{B}(n, p)$, si X est à valeurs dans l'ensemble $\{0, 1, \ldots, n\}$ et

$$\mathbb{P}_X(\{k\}) = \mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}, \; k \in \{0, 1, \ldots, n\}.$$

Si on joue a pile ou face n fois de suite, la loi binomiale est la loi du nombre de pile au cours des n lancers.

Loi géométrique : on dit qu'une variable aléatoire X suit une loi géométrique de paramètres p et on note $X \sim \mathcal{G}(p)$, si X est à valeurs dans l'ensemble $\{1, 2, \ldots\}$ et

$$\mathbb{P}_X(\{k\}) = \mathbb{P}(X = k) = p(1 - p)^{k-1}, \; k \in \{1, 2, \ldots\}.$$

Si l'on répète un jeu de pile ou face, la loi géométrique est la loi du temps d'apparition du premier pile.

Loi géométrique bis : on dit qu'une variable aléatoire X suit une loi géométrique de paramètres p et on note $X \sim \mathcal{G}(p)$, si X est à valeurs dans l'ensemble $\{0, 1, 2, \ldots\}$ et

$$\mathbb{P}_X(\{k\}) = \mathbb{P}(X = k) = p(1 - p)^{k}, \; k \in \{0, 1, 2, \ldots\}.$$

Loi de Poisson : on dit qu'une variable aléatoire X suit une loi de Poisson de paramètres λ et on note $X \sim \mathcal{P}(\lambda)$, si X est à valeurs dans l'ensemble $\{0, 1, 2, \dots\}$ et

$$\mathbb{P}_X(\{k\}) = \mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!},\ k \in \{0, 1, 2, \dots\}.$$

La loi de Poisson peut être vue comme un cas limite de loi binomiale. En effet, on montre qu'une loi de Poisson est la limite d'une $\mathcal{B}(n, p)$ pour laquelle on a $n \to \infty$ et $p \to 0$ et $np \to \lambda \neq \infty$.

2.1.2 Variables aléatoires continues

Nous introduisons à présent les variables aléatoires continues : ce sont des variables aléatoires qui peuvent prendre un nombre infini (non dénombrable) de valeurs, typiquement ce sont les variables à valeurs dans un intervalle de la droite réelle.

Définition 2.1.9. On dit qu'une variable aléatoire $X : \quad \to X()$ est continue si l'ensemble de ses valeurs $X()$ est un intervalle de $\mathbb{R}$. On dit qu'une variable continue admet une densité $f(x)$ si pour tout intervalle $[a, b] \subset X(\Omega)$:

$$\mathbb{P}_X([a, b]) = \mathbb{P}(X \in [a, b]) = \int_a^b f(x)dx,$$

où f est une fonction continue, positive sur $X(\Omega)$ telle que $\int_{X(\Omega)} f(x)dx = 1$.

Remarque 2.1.10. Si X est une variable continue et admet une densité f, alors pour tout $x_0 \in X()$, on a $\mathbb{P}_X(\{x_0\}) = \mathbb{P}(X = x_0) = 0$. Autrement, la variable X a une probabilité nulle de tomber sur un point donné de l'intervalle $X(\Omega)$. En revanche, on a une chance non nulle de tomber dans un petit intervalle autour de x_0 :

$$\mathbb{P}([x_0 - h, x_0 + h]) = \mathbb{P}(X = [x_0 - h, x_0 + h]) = \int_{x_0-h}^{x_0+h} f(x)dx > 0.$$

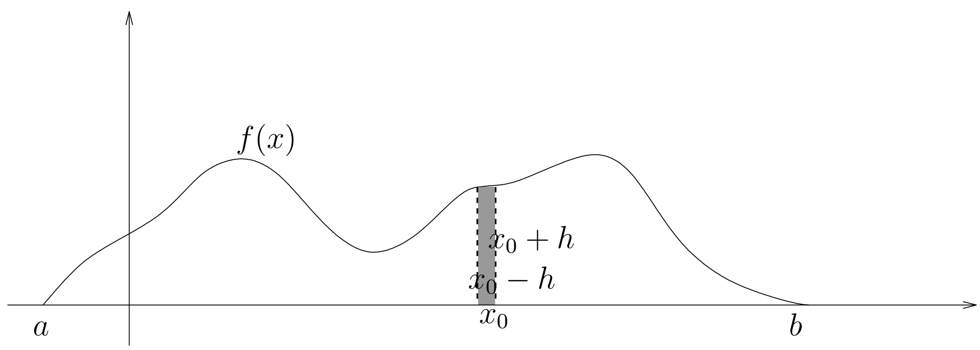


FIGURE 2.1 – Probabilité sur un intervalle via une densité.

Exemples de variables continues

Nous donnons à présent des exemples usuels de loi de probabilité sur des intervalles de $\mathbb{R}$. La loi gaussienne, appelée encore loi normale jouera en particulier un rôle fondamental dans la suite du cours.

Loi uniforme : on dit qu'une variable aléatoire X suit une loi uniforme sur un intervalle $[a, b]$ si X est à valeurs dans l'ensemble $[a, b]$ et pour tout$[c, d[\subset [a, b]$:

$$\mathbb{P}_X([c, d[) = \mathbb{P}(X \in [c, d[) = \frac{d-c}{b-a} = \frac{1}{b-a}\int_c^d 1dx,$$

autrement dit, X a la densité $f(x) \equiv 1/(b-a)$ sur l'intervalle $[a, b]$.

Loi normale ou gaussienne : on dit qu'une variable aléatoire X suit une loi normale de paramètres (μ, σ^2) et on note $X \sim \mathcal{N}(\mu, \sigma^2)$ si X est à valeurs dans $\mathbb{R}$ et pour tout intervalle $[a, b] \subset \mathbb{R}$:

$$\mathbb{P}_X([a, b]) = \mathbb{P}(X \in [a, b]) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,$$

autrement dit, X a pour densité $f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ sur $\mathbb{R}$.

Loi exponentielle : on dit qu'une variable aléatoire X suit une loi exponentielle de paramètre λ et on note $X \sim \mathcal{E}(\lambda)$ si X est à valeurs dans $[0, +\infty[$ et pour tout intervalle $[a, b] \subset \mathbb{R}$:

$$\mathbb{P}_X([a, b]) = \mathbb{P}(X \in [a, b]) = e^{-\lambda a} - e^{-\lambda b} = \int_a^b \lambda e^{-\lambda x}dx,$$

autrement dit, X a pour densité la fonction $f(x) = \lambda e^{-\lambda x}$ sur $\mathbb{R}^+$.

Loi gamma : on dit qu'une variable aléatoire X suit une loi gamma de paramètres (a, b) et on note $X \sim \Gamma(a, b)$ si X est à valeurs dans $[0, +\infty[$ et pour tout intervalle $[c, d] \subset \mathbb{R}$:

$$\mathbb{P}_X([c, d]) = \mathbb{P}(X \in [c, d]) = \frac{b^a}{\Gamma(a)} \int_c^d x^{a-1} e^{-bx} dx,$$

autrement dit, X a pour densité $f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}$ sur $\mathbb{R}^+$.

2.2 Fonction de répartition

Nous introduisons dans ce paragraphe la notion de fonction de répartition d'une variable aléatoire (discrète ou continue). Cette fonction caractérise la loi d'une variable aléatoire, et nous sera utile dans la suite pour dire qu'une suite de variables aléatoires converge vers une variable limite.

Définition 2.2.1. Soit $X : \ \to X()$ une variable aléatoire. On appelle fonction de répartition de répartition de X, et on note F_X, la fonction de $\mathbb{R}$ dans l'intervalle $[0, 1]$ définie par

$$F_X(x) = \mathbb{P}(X \leqslant x).$$

Proposition 2.2.2. *Soit $X : \ \to X()$ une variable aléatoire. Alors sa fonction de répartition F_X vérifie les propriétés suivantes :*

1. *F_X est croissante.*
2. *F_X est continue à droite.*
3. *$F_X(-\infty) = 0$ et $F_X(+\infty) = 1$*

Être continu à droite signifie que si la fonction "saute", sa valeur au point de saut est la valeur à droite de celui-ci, *i.e.* les points gris sur la figure ci-après.

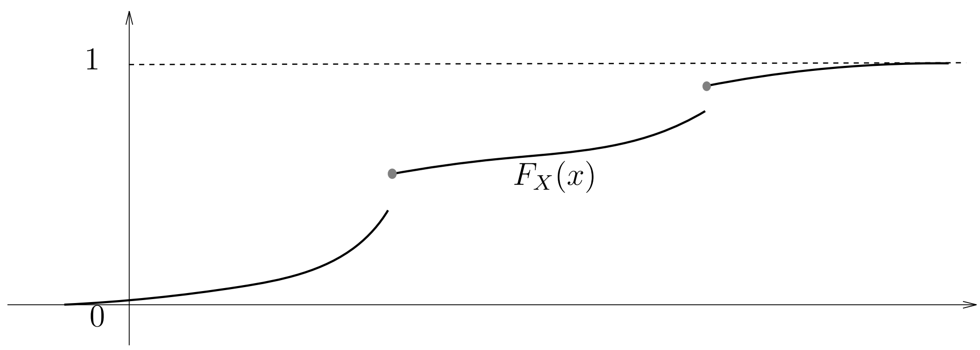


FIGURE 2.2 – Fonction de répartition générique.

Remarque 2.2.3. D'après la définiton de la fonction de répartition, pour tous réels a et b, avec $a < b$ on a : $\mathbb{P}(a < X \leqslant b) = F_X(b) - F_X(a)$. En particulier, pour tout $x \in \mathbb{R}$: $\mathbb{P}(X > x) = 1 - F_X(x)$.

2.2.1 Fonction de répartition d'une variable discrète

Soit X une variable aléatoire discrète pouvant prendre les valeurs $x_1, x_2, \dots, x_n, \dots$ de probabilités respectivement $p_1, p_2, \dots, p_n, \dots$ avec $x_1 < x_2 < \dots < x_n < \dots$. Alors la fonction de répartition de X est donnée par la formule :

$$F_X(x) = \sum_{i=1}^{i=k} p_i,$$

où k est l'indice tel que $x_k \leqslant x < x_{k+1}$. La fonction $x \mapsto F_X(x)$ est alors une fonction constante par morceaux, dont le graphe a l'allure ci-dessous.

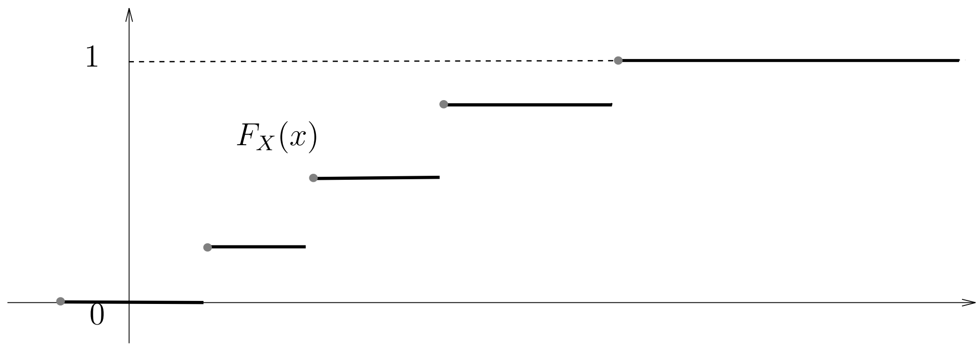


FIGURE 2.3 – Fonction de répartition d'une variable discrète.

Exemple 2.2.4. Ci-dessous, la fonction F_X lorsque $X \sim \mathcal{B}(p)$. La fonction fait un "saut" d'une hauteur p en zéro, et d'une hauteur de $(1-p)$ en un.

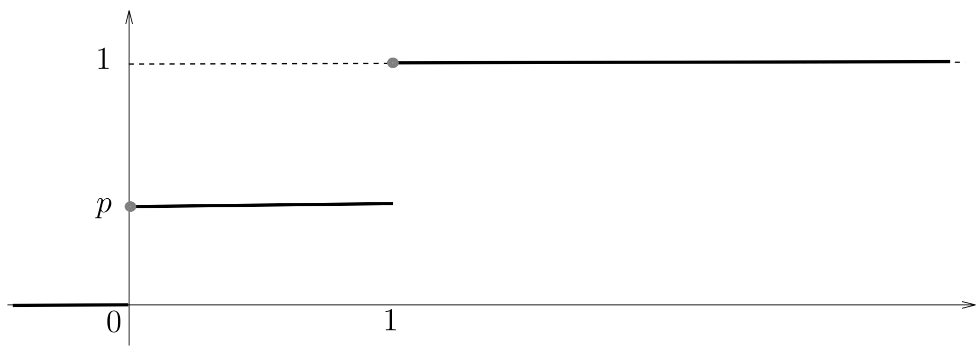


FIGURE 2.4 – Fonction de répartition d'une variable de Bernoulli $\mathcal{B}(p)$.

Exemple 2.2.5. Ci-après, la fonction de répartition d'une variable S de l'exemple 2.1.7, la somme de deux dés. La fonction F_S fait un "saut" d'une hauteur $1/36$ en zéro, d'une hauteur de $2/36$ en un, d'une hauteur $3/36$ en deux etc.

Exemple 2.2.6. Soit X une variable de loi géométrique sur $\{1, 2, \dots\}$, *i.e.* telle que

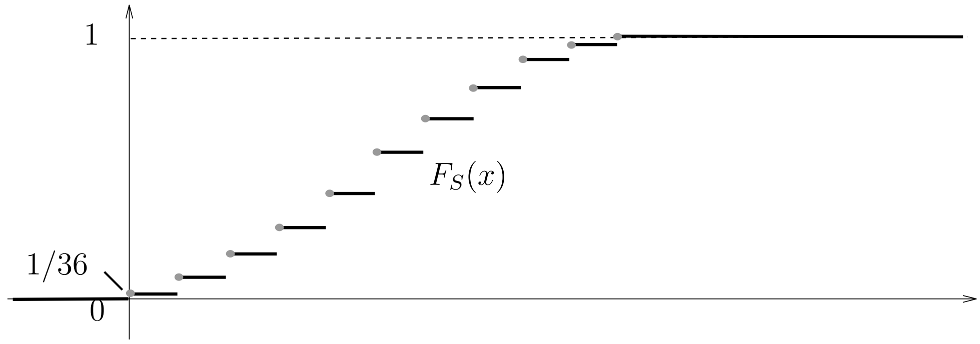


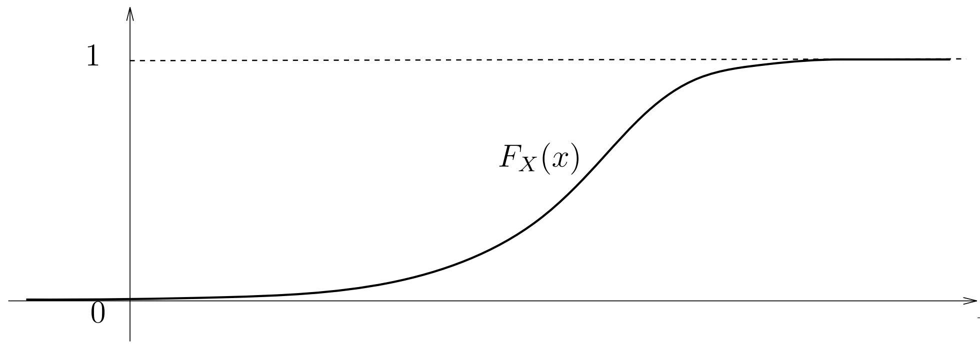
FIGURE 2.5 – Fonction de répartition de la variable S (somme de deux dés).

$\mathbb{P}(X = k) = p(1-p)^{k-1}$. Alors, pour tout entier $m \geqslant 1$, on a

$$\mathbb{P}(X \leqslant m) = 1 - \mathbb{P}(X > m) = 1 - \sum_{k=m+1}^{+\infty} p(1-p)^{k-1} = 1 - (1-p)^m.$$

2.2.2 Fonction de répartition d'une variable continue

Soit X une variable aléatoire continue de densité $f(x)$. Alors, la fonction de répartition de X est la primitive de f : $F_X(x) = \int_{-\infty}^{x} f(u)du$. Dans ce cas, la fonction F_X est une fonction continue à gauche et à droite : on peut la tracer sans lever le stylo.



].

FIGURE 2.6 – Fonction de répartition d'une variable continue.

Exemple 2.2.7. Considérons le cas d'une variable X de loi uniforme sur l'intervalle $[0,1]$. Sa densité f_X est constante sur l'intervalle $[0,1]$ et vaut zéro ailleurs. On en déduit que F_X vaut zéro sur $]-\infty, 0]$, vaut 1 sur $[1, +\infty[$ et :

$$F_X(x) = \int_{-\infty}^{x} f(u)du = \int_0^x 1 \times du = x, \quad \text{pour} \quad x \in [0,1].$$

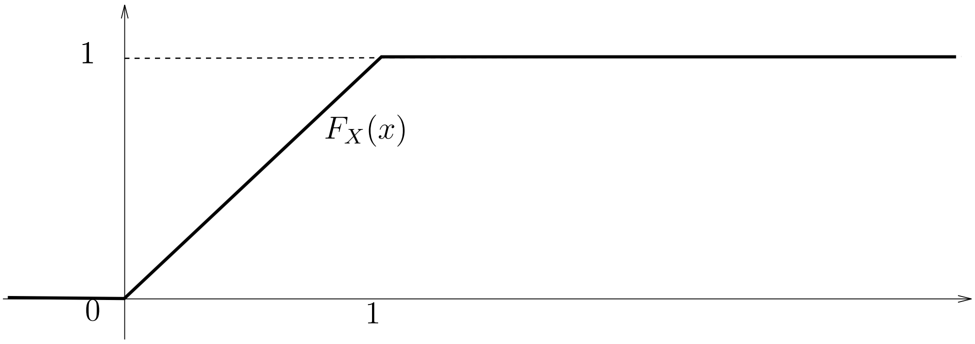


FIGURE 2.7 – Fonction de répartition d'une variable uniforme.

Exemple 2.2.8. Considérons le cas d'une variable X exponentielle de paramètre λ. Sa densité f_X est nulle sur $]-\infty, 0]$ et est donnée par $f_X(x) = \lambda \exp(-\lambda x)$ sur $[0, +\infty[$. On en déduit que F_X vaut zéro sur $]-\infty, 0]$, et vaut, pour $x > 0$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(u)du = \int_0^x \lambda \exp(-\lambda u)\, du \\ &= [-\exp(-\lambda u)]_0^x = 1 - \exp(-\lambda x). \end{aligned}$$

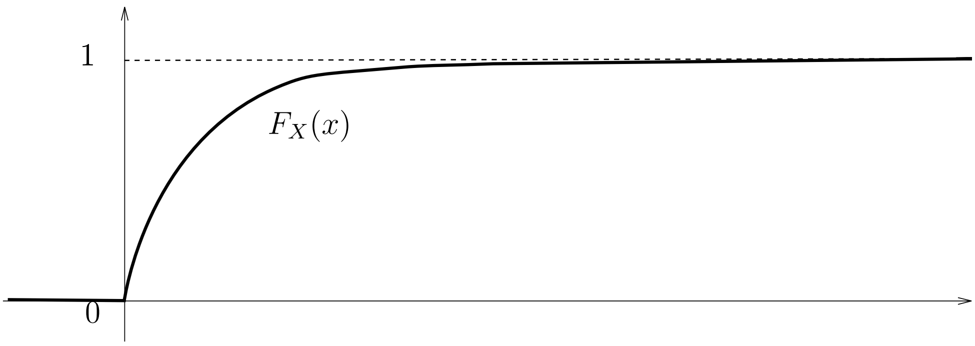


FIGURE 2.8 – Fonction de répartition d'une variable exponentielle.

2.3 Moments d'une variable aléatoire

Dans cette section, nous nous intéressons aux notions de moyenne et de variance d'une variable aléatoire. Ces deux notions seront fondamentales dans la partie "statistique" du cours. Nous définissons tout d'abord la notion d'espérance mathématique.

2.3.1 Espérance d'une variable aléatoire

La notion d'espérance généralise la notion bien connue de moyenne. Il s'agit précisément d'une moyenne pondérée. Dans les deux prochaines sections, nous donnons la définition de l'espérance mathématique d'une variable aléatoire discrète puis d'une variable continue.

Espérance d'une variable discrète

La notion de moyenne pondérée nous est tous familère, il suffit de penser au calcul de la moyenne au baccalauréat ou les différentes matières ont des coefficients distincts : pour un bac S option SVT, un 18 en bio est "plus intéressant" qu'un 18 en sport... L'espérance d'une variable aléatoire discrète est précisément un moyenne pondérée :

Définition 3.3.1. Soit X une variable aléatoire discrète à valeurs dans un ensemble au plus dénombrable $\{x_1, \ldots, x_n, \ldots\}$. On note $p_i := \mathbb{P}(X = x_i)$. Alors l'espérance de X, que l'on note $\mathbb{E}[X]$, est donnée par la formule :

$$\mathbb{E}[X] := \sum_{i=1}^{\infty} x_i p_i = \sum_{i=1}^{\infty} x_i \mathbb{P}_X(\{x_i\}) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i).$$

Plus généralement, si h est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, alors l'espérance de la variable $h(X)$ est donnée par la formule

$$\mathbb{E}[h(X)] := \sum_{i=1}^{\infty} h(x_i) p_i = \sum_{i=1}^{\infty} h(x_i) \mathbb{P}_X(\{x_i\}) = \sum_{i=1}^{\infty} h(x_i) \mathbb{P}(X = x_i).$$

Exemple 2.3.2. Par exemple, si X suit une loi de Bernoulli $\mathcal{B}(p)$ sur $\{0, 1\}$, alors l'espérance de X vaut

$$\mathbb{E}[X] = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = 0 \times (1 - p) + 1 \times p = p.$$

Exemple 2.3.3. Par exemple, si X suit une loi uniforme sur $\{1, 2, \ldots, n\}$, alors l'espérance de X vaut

$$\begin{aligned} \mathbb{E}[X] &= 1 \times \mathbb{P}(X = 1) + 2 \times \mathbb{P}(X = 2) + \ldots + n \times \mathbb{P}(X = n) \\ &= \frac{1 + 2 + \ldots + n}{n} = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}. \end{aligned}$$

Espérance d'une variable continue

On peut généraliser la définition précédente au cadre continu, en remplaçant la somme discrète par une intégrale.

Définition 2.3.4. Soit X une variable aléatoire continue à valeur dans un intervalle $X() \subset \mathbb{R}$ et admettant de densité $f(x)$. Alors l'espérance de X, que l'on note $\mathbb{E}[X]$, est donnée par la formule :

$$\mathbb{E}[X] := \int_{X(\Omega)} x f(x) dx.$$

Plus généralement, si h est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, alors l'espérance de la variable $h(X)$ est donnée par la formule

$$\mathbb{E}[h(X)] := \int_{X(\Omega)} h(x) f(x) dx.$$

Exemple 2.3.5. Par exemple, si X suit une loi uniforme sur l'intervalle $[0,1]$, *i.e.* X admet la densité $f \equiv 1$ sur l'intervalle $[0,1]$, alors l'espérance de X vaut

$$\mathbb{E}[X] = \int_0^1 xf(x)dx = \int_0^1 xdx = \left[\frac{x^2}{2}\right]_0^1 = 1/2.$$

De même, l'espérance de la variable X^2 vaut :

$$\mathbb{E}[X^2] = \int_0^1 x^2 f(x)dx = \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = 1/3.$$

Exemple 2.3.6. Si X suit une loi exponentielle de paramètre λ, *i.e.* si X admet la densité $f(x) = \lambda e^{-\lambda x}$ sur l'intervalle $[0, +\infty[$, alors en intégrant par partie, on obtient que l'espérance de X vaut

$$\begin{aligned}\mathbb{E}[X] &= \int_0^{+\infty} xf(x)dx = \int_0^{+\infty} x\lambda e^{-\lambda x}dx = \left[-xe^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x}dx \\ &= \left[\frac{-e^{-\lambda x}}{\lambda}\right]_0^{+\infty} = \frac{1}{\lambda}.\end{aligned}$$

Remarque 2.3.7. L'espérance mathématique n'est pas toujours définie. C'est en particulier le cas de la loi de Cauchy dont la densité sur $\mathbb{R}$ est donnée par $f(x) = \frac{1}{\pi(1+x^2)}$. Alors on a

$$\mathbb{E}[|X|] = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)}dx = 2\int_0^{+\infty} \frac{x}{\pi(1+x^2)}dx = +\infty.$$

Propriétés de l'espérance

Nous donnons maintenant quelques propriétés de l'espérance, qui sont vérifiées que l'on se place dans le cas discret ou continu.

Proposition 2.3.8 (Linéarité de l'espérance). *Soient X et Y deux variables aléatoires et $c \in \mathbb{R}$ une constante. Alors on a :*

1. $\mathbb{E}[c] = c$;
2. $\mathbb{E}[cX + Y] = c \times \mathbb{E}[X] + \mathbb{E}[Y]$.

Proposition 2.3.9 (Positivité de l'espérance). *Soient X et Y deux variables aléatoires telles que $X \leqslant Y$ avec probabilité un, alors $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$.*

2.3.2 Variance et autres moments

Définition 2.3.10. Soient X une variable aléatoire et m un entier strictement positif. On dit que X admet un moment d'ordre m si $\mathbb{E}[|X|^m] < +\infty$. Si c'est le cas, on appelle moment d'ordre m la quantité $\mathbb{E}[X^m]$, c'est-à-dire selon que l'on est dans le cas discret ou continu :

$$\mathbb{E}[X^m] := \int_{X(\Omega)} x^m f(x)dx,$$

$$\mathbb{E}[X^m] := \sum_i x_i^m p_i = \sum_{i=1}^{\infty} x_i^m \mathbb{P}_X(\{x_i\}) = \sum_{i=1}^{\infty} x_i^m \mathbb{P}(X = x_i).$$

Exemple 2.3.11. Dans les exemples ci-dessus, on a vu que la loi uniforme sur $[0, 1]$ admet un moment d'ordre deux puisque $\mathbb{E}[X^2] = 1/3 < +\infty$. En revanche, la loi de Cauchy de densité $f(x) = \frac{1}{\pi(1+x^2)}$ n'admet pas de moment d'ordre un puisque $\mathbb{E}[|X|] = +\infty$.

Définition 2.3.12. Soient X une variable aléatoire qui admet des moments d'ordre un et deux, *i.e.* $\mathbb{E}[|X|] < +\infty$, $\mathbb{E}[|X|^2] < +\infty$. On appelle variance de X et on note $\mathrm{var}(X)$ la quantité

$$\mathrm{var}(X) := \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

La variance traduit la dispersion de la distribution de la variance autour de sa valeur moyenne. Étant un carré, la dimension de la variance n'est pas celle de la moyenne. C'est pourquoi on utilise plus souvent l'écart type, noté souvent σ, qui est la racine de la variance. On dit aussi que la variance traduit la notion d'incertitude. Plus la variance est faible, moins le résultat de l'expérience aléatoire est incertain. Le cas extrême est celui d'une variable aléatoire de variance nulle, qui est en fait déterministe.

Définition 2.3.13. On dit qu'une variable aléatoire Y est centrée réduite si sa moyenne $\mathbb{E}[Y]$ est nulle et si sa variance $\mathrm{var}(Y)$ est égale à un. Si X est une variable aléatoire qui admet des moments d'ordre un et deux, alors la variable $Y := \frac{X-\mathbb{E}[X]}{\sqrt{\mathrm{var}(X)}}$ est centrée réduite.

Exemple 2.3.14. Par exemple, si X suit une loi de Bernoulli $\mathcal{B}(p)$ sur $\{0, 1\}$, on a vu que l'espérance de X vaut

$$\mathbb{E}[X] = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = 0 \times (1 - p) + 1 \times p = p.$$

Le moment d'ordre deux vaut lui aussi p :

$$\mathbb{E}[X^2] = 0^2 \times \mathbb{P}(X = 0) + 1^2 \times \mathbb{P}(X = 1) = 0 \times (1 - p) + 1 \times p = p,$$

de sorte que la variance de X vaut $\mathrm{var}(X) = p - p^2 = p(1 - p)$.

Exemple 2.3.15. Dans l'exemple de la loi uniforme sur l'intervalle $[0, 1]$, on a vu que $\mathbb{E}[X] = 1/2$ et $\mathbb{E}[X^2] = 1/3$, on a donc $\mathrm{var}(X) = 1/3 - 1/4 = 1/12$.

Proposition 2.3.16. *Soient X et Y deux variables aléatoires, et a et b deux constantes réelles. Alors on a $var(aX+b) = a^2 var(X)$ et*

$$var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$$

où $cov(X,Y)$ est la covariance de X et Y définie par :

$$cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])].$$

2.3.3 Moments des variables usuelles

On explicite ici les premiers moments des variables usuelles. On traite dans un premier temps le cas des variables discrètes, puis celui des variables continues admettant un densité.

Moments des variables discrètes usuelles

Le cas de la variable de Bernoulli a déja été traité dans l'exemple 2.3.14.

Loi de Bernoulli : on a vu que si X suit une loi de Bernoulli de paramètre $p \in [0,1]$, alors $\mathbb{E}[X] = p$ et $\mathrm{var}(X) = p(1-p)$.

Loi uniforme : si X suit une loi uniforme sur un ensemble fini $E = \{1, \ldots, n\}$ alors on a vu dans l'exemple 2.3.3 que $\mathbb{E}[X] = n(n+1)/2$. Le calcul du moment d'ordre deux montre que $\mathrm{var}(X) = (n^2-1)/12$, en effet, on a

$$\begin{aligned} \mathbb{E}[X^2] &= 1^2 \times \mathbb{P}(X=1) + 2^2 \times \mathbb{P}(X=2) + \ldots + n^2 \times \mathbb{P}(X=n) \\ &= \frac{1^2+2^2+\ldots+n^2}{n} = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}. \end{aligned}$$

Loi binomiale : si $X_1, \ldots, X_n$ sont des variables "indépendantes" de loi de Bernoulli de paramètre p, alors la somme $S_n = X_1 + \ldots + X_n$ suit une loi binomiale $\mathcal{B}(n,p)$. Par linéarité de l'espérance, on en déduit que $\mathbb{E}[S_n] = np$ et $\mathrm{var}(S_n) = np(1-p)$.

Loi géométrique : si X suit une loi géométrique de paramètre p sur $\{1, 2, \ldots\}$ alors $\mathbb{E}[X] = 1/p$ et $\mathrm{var}(X) = 1-p$. En effet, on a

$$\begin{aligned} \mathbb{E}[X] &= \textstyle\sum_{k=1}^{+\infty} k \times p(1-p)^{k-1} = p \times \left(\sum_{k=0}^{+\infty} k \times (1-p)^{k-1}\right) \\ &= p \times \left(\sum_{k=0}^{+\infty} \times (1-p)^k\right)' = p \times \left(\frac{-1}{p}\right)' = 1/p. \end{aligned}$$

Par ailleurs, on montre que $\mathbb{E}[X^2] = (1-p) + 1/p^2$, d'où le résultat.

Loi de Poisson : si X suit une loi de Poisson de paramètres λ, alors $\mathbb{E}[X] = \lambda$ et $\mathrm{var}(X) = \lambda$. En effet, on a

$$\begin{aligned}
\mathbb{E}[X] &= \textstyle\sum_{k=0}^{+\infty} k \times e^{-\lambda}\frac{\lambda^k}{k!} = \lambda \times \left(e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!}\right) \\
&= \lambda \times \left(e^{-\lambda} \textstyle\sum_{\ell=0}^{+\infty} \frac{\lambda^\ell}{\ell!}\right) = \lambda. \\
\mathbb{E}[X^2] &= \textstyle\sum_{k=0}^{+\infty} k^2 \times e^{-\lambda}\frac{\lambda^k}{k!} = \sum_{k=0}^{+\infty} (k(k-1)+k) \times e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \lambda^2 \left(\textstyle\sum_{k=2}^{+\infty} e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!}\right) + \lambda \left(\sum_{k=1}^{+\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!}\right) = \lambda^2 + \lambda.
\end{aligned}$$

Moments des variables usuelles continues

On donne maintenant les moyennes et variances des variables continues les plus usuelles. Le cas d'une variable uniforme sur l'intervalle $[0, 1]$ a déjà été traité dans l'exemple 2.3.5.

Loi uniforme : si X suit une loi uniforme sur un intervalle $[a, b]$ alors l'espérance de X est $\mathbb{E}[X] = (b - a)/2$ et sa variance $\mathrm{var}(X) = (b - a)^2/12$.

Loi normale ou gaussienne : si X suit une loi normale de paramètres (μ, σ^2) alors la moyenne de X est $\mathbb{E}[X] = \mu$, et sa variance $\mathrm{var}(X) = \sigma^2$.

Loi exponentielle : si X suit une loi exponentielle de paramètre λ, on a vu dans l'exemple 5.1.10 que $\mathbb{E}[X] = 1/\lambda$. De même, on montre que $\mathbb{E}[X^2] = 1/\lambda^2 + 1/\lambda$. On a donc $\mathrm{var}(X) = 1/\lambda^2$.