MATHF-105 : Probabilités
Résumé

R. Petit

Année académique 2015 - 2016

Contents

1 Rappels

1.1 Rappel sur les séries

Les fonctions logarithmique et exponentielle ont un développement de Taylor exact. Pour la fonction logarithmique, on a, pour $x \in (-1, 1)$:

$$\log(1-x) = -\sum_{k \geqslant 1} \frac{x^k}{k}.$$

Si on pose $S_n \coloneqq \sum_{k=1}^{n} u_k$, on a $(S_n)_{n \in \mathbb{N}}$, la suite des sommes partielles, et $n \mapsto S_n$, une application croissante si (u_n) est une suite positive. Il y a donc deux situations distinctes possibles :

- (S_n) est une suite bornée ($\exists M \in \mathbb{R}$ t. q. $\forall n \in \mathbb{N} : S_n \leqslant M$) et donc converge vers $S \in \mathbb{R}$;
- (S_n) n'est pas bornée ($\forall M \in \mathbb{R} : \exists n \in \mathbb{N}$ t. q. $S_n > M$) et donc diverge vers $+\infty$.

1.1.1 Exemple sur les séries

Prenons $u_n \coloneqq x^n$, avec $x > 0$.

- Si $x = 1$, on a $n \to +\infty \Rightarrow S_n \to +\infty$;
- si $x \neq 1$, on a $(1-x)S_n = x - x^{n+1}$, et donc :

$$S_n \coloneqq x\frac{1-x^n}{1-x}.$$

  - Si $x < 1$, alors $x^n \to 0$ pour $n \to +\infty$, et donc $S_n \to \frac{x}{1-x}$;
  - si $x > 1$, alors $x^n \to +\infty$ pour $n \to +\infty$, et donc $S_n \to +\infty$.

1.1.2 Conclusion de la suite géométrique

On voit alors :

$$\sum_{n \geqslant 1} x^n = \begin{cases} \frac{x}{1-x} & \text{si } x \in [0,1) \\ +\infty & \text{sinon} \end{cases}.$$

Si la suite commence à l'indice 0, on a :

$$\sum_{n \geqslant 0} x^n = 1 + \sum_{n \geqslant 1} x^n = \begin{cases} 1 + \frac{x}{1-x} = \frac{1}{1-x} & \text{si } x \in [0,1) \\ +\infty & \text{sinon} \end{cases}.$$

1.2 Rappels d'analyse

Définition 1.1. Une fonction $f : X \to Y$ est dite mesurable si :

$$\forall A \subset \mathcal{B}(Y) : \{\omega \in \Omega \text{ t. q. } X(\omega) \in A\} \in \mathcal{F},$$

où $\mathcal{B}(Y)$ représente la tribu des boréliens (voir définition 2.9).

Théorème 1.2. *Dans $\mathbb{R}$, toute série absolument convergente est convergente.*

Théorème 1.3. *Dans* $\mathbb{R}$*, toute intégrale impropre absolument convergente est convergente.*

2 Espaces de probabilités

2.1 Définition

Définition 2.1. L'ensemble Ω est l'**espace des chances**, l'ensemble des résultats possibles d'un phénomène aléatoire.

Remarque.

- Ω peut être fini (dénombrable) ou infini ;
- $\Omega = \{0,1\}^{\mathbb{N}}$ est l'ensemble des suites à valeur dans $\{0,1\}$;
- Ω peut être un espace dit *fonctionnel* quand le résultat d'une expérience est une fonction.

Définition 2.2. Un événement E est un ensemble de réalisations possibles à une expérience tel que $E \subseteq \Omega$.

Remarque. L'ensemble $\mathcal{P}(\Omega)$ n'est pas toujours dénombrable. Et donc l'ensemble $\mathcal{P}(\Omega)$ est-il le bon ensemble pour décrire les événements ?

- Si $|\Omega| \in \mathbb{N}$: oui ;
- si $|\Omega| \notin \mathbb{N}$: non.

Définition 2.3. $\mathcal{F}$ est la **classe des événements**. On mesure la *probabilité d'occurrence* d'un événement $A \in \mathcal{F}$. On introduit une fonction d'ensemble $\mathbb{P}$ où :

$$\mathbb{P} : \mathcal{F} \to [0,1] : A \mapsto \mathbb{P}(A).$$

On impose :

(i) $\mathbb{P}(\emptyset) = 0$;

(ii) $\mathbb{P}(\Omega) = 1$;

(iii) $\forall A, B \in \mathcal{F} : A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Proposition 2.4. *Soient $A_1, \ldots, A_n \in \mathcal{F}$. On a :*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} (-1)^{i-1} \sum_{1 \leqslant k_1 < \cdots < k_i \leqslant n} \mathbb{P}\left(\bigcap_{\gamma=1}^{i} A_{k_\gamma}\right).$$

2.1.1 Loi uniforme sur un ensemble fini (ou dénombrable)

Définition 2.5. Soient $m < n \in \mathbb{N}$. On définit l'**intervalle entier** $[\![m,n]\!]$ par :

$$[\![m,n]\!] : \{x \in \mathbb{N} \text{ t. q. } m \leqslant x \leqslant n\}.$$

Définition 2.6. Soit $\Omega = [\![1,n]\!]$. Soit $A \subseteq \Omega$. La loi uniforme est donnée par :

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{n}.$$

Remarque. Il arrive que $|A|$ soit difficile à déterminer et qu'il faille aller chercher du côté de l'analyse combinatoire.

2.1.2 Loi uniforme sur un ensemble infini (intervalle)

Définition 2.7. Soit $\Omega = [0, 1]$ et soit $A = [a, b] \subseteq \Omega$. La loi uniforme est donnée par :

$$\mathbb{P}(A) = (b - a).$$

Remarque. La définition de loi uniforme sur un intervalle fait intervenir la notion de mesure et donc de mesurabilité. Or il existe des parties de Ω sur lesquelles la mesure n'a pas de sens. En général, $\mathcal{P}(\Omega)$ est *trop grand*, et il faut donc remplacer l'utilisation de l'ensemble des parties par la notion de tribu.

Définition 2.8. Soit Ω un ensemble de chances et $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ une famille de parties de Ω. On dit que $\mathcal{F}$ est une tribu s'il respecte les trois propriétés suivantes :

- $\emptyset \in \mathcal{F}$;
- $\forall A : A \in \mathcal{F} \Rightarrow A^{\complement} \in \mathcal{F}$;
- $\forall A_1, \ldots, A_n, \ldots : A_1, \ldots, A_n, \ldots \in \mathcal{F} \Rightarrow \bigcup_{k \geqslant 1} A_k \in \mathcal{F}$.

Une autre appellation pour une tribu est une σ-algèbre.

Remarque.

- On remarque que $\mathcal{P}(\Omega)$ est une tribu, mais une tribu trop grande pour être intéressante ;
- Soit $A \in \mathcal{P}(\Omega)$. Alors $T \coloneqq \{\emptyset, A, A^{\complement}, \mathcal{P}(\Omega)\}$ est une tribu. T est la plus petite tribu contenant A, et on l'appelle la **tribu engendrée par** A, que l'on note $\sigma(A)$.

Définition 2.9. Soit I une partie de $\mathcal{P}(\Omega)$. On appelle la *tribu engendrée par* I la plus petite tribu contenant I et on la note $\sigma(I)$.

En prenant $I \coloneqq \{$ intervalles ouverts de $[0, 1]\}$, on obtient $\sigma(I)$ que l'on appelle **tribu des boréliens**. [1]

Définition 2.10. Soit Ω un ensemble de chances et $\mathcal{F} \subset \mathcal{P}(\Omega)$ une tribu sur Ω. Une probabilité sur $(\Omega, \mathcal{F}, \mathbb{P})$ est une fonction $\mathbb{P}$ définie par :

$$\mathbb{P} : \mathcal{F} \to [0, 1] : A \mapsto \mathbb{P}(A),$$

où $\mathbb{P}$ satisfait :

(i) $\mathbb{P}(\emptyset) = 0$;

(ii) $\forall A \in \mathcal{F} : \mathbb{P}(aA) + \mathbb{P}(A^{\complement}) = 1$;

(iii) $\forall A_1, \ldots, A_n, \ldots$ disjoints deux à deux, on a :

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

Définition 2.11. On appelle $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités.

Remarque. Probabiliser un expérience revient à déterminer :

- Ω, l'espace des chances ;
- $\mathcal{F}$, la classe des événements ;
- $\mathbb{P}$, la fonction d'ensembles sur $\mathcal{F}$.

[1] Le nom de *borélien* vient du mathématicien français Émile Borel suite à ses travaux sur la théorie de la mesure.

2.2 Modèles

2.2.1 Modèles discrets

Remarque. On prend Ω un ensemble fini ou dénombrable. On prend également $\mathcal{F} = \mathcal{P}(\Omega)$.

Si Ω est fini, on parle de tirages, et si Ω est infini dénombrable, on parle de populations.

On pose :

$$\mathbb{P} : \{k\} \mapsto p_k \in [0, 1],$$

où :

$$\sum_{k \in \Omega} p_k = 1$$

et pour $A = \{k_1, \dots, k_n\} \in \mathcal{F}$:

$$\mathbb{P}(A) = \sum_{\gamma=1}^{n} p_{k_\gamma}.$$

Définition 2.12 (Modèle de Bernoulli)**.** On prend $\Omega = \{0, 1\}$ où :

$$\begin{cases} p_0 & = 1 - p \\ p_1 & = p \end{cases}.$$

Remarque. Il est évident que $p + (1 - p) = 1 = P(\Omega)$.

Définition 2.13 (Modèle binomial)**.** On prend $\Omega = [\![0, N]\!]$ (et donc $\mathcal{F} = \mathcal{P}(\Omega)$) et $p \in [0, 1]$. Le modèle binomial est défini par $p_k = \binom{n}{k} p^k (1 - p)^{N-k}$ pour tout $k \in [\![0, N]\!]$.

Remarque. On remarque que $\sum_{k \geqslant 1} p_k = 1$ car les p_k représentent les termes du binôme de Newton $(p + (1 - p))^N = 1^N = 1$.

Définition 2.14 (Modèle géométrique)**.** On prend $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{F}(\Omega) \simeq \mathbb{R}$, et $p \in (0, 1)$. Le modèle géométrique est défini par $p_k = (1 - p)^{k-1} p$ pour tout $k \in \mathbb{N}$.

Remarque. On remarque que :

$$\sum_{k \geqslant 1} p_k = \sum_{k \geqslant 1} p(1 - p)^{k-1} = p \sum_{k \geqslant 0} (1 - p)^k = p \frac{1}{1 - (1 - p)} = \frac{p}{p} = 1,$$

où on utilise la formule de la somme des termes d'une suite géométrique u définie par $u_n = u_{n-1} q$ pour $n \geqslant 1$ (avec $0 < q < 1$) qui donne :

$$\sum_{k=0}^{N} u_k = u_0 \frac{1 - q^{N+1}}{1 - q},$$

et pour la série, il suffit de passer à la limite :

$$\lim_{N \to +\infty} \sum_{k=0}^{N} u_k = \lim_{N \to +\infty} u_0 \frac{1 - q^{N+1}}{1 - q} = u_0 \frac{1}{1 - q}.$$

Définition 2.15 (Modèle de Poisson)**.** On prend $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\Omega)$, et un paramètre $\lambda \in \mathbb{R}_0^+$. Le modèle poissonien est défini par $p_k = \exp(-\lambda) \frac{\lambda^k}{k!}$ pour tout $k \in \mathbb{N}$.

Remarque. On remarque que $\mathbb{P}(\Omega) = 1$ en utilisant la formule de Taylor de l'exponentielle :

$$\exp(x) = \sum_{k \geqslant 0} \frac{x^k}{k!}.$$

On a effectivement :

$$\mathbb{P}(\Omega) = \sum_{k \geqslant 0} \mathbb{P}(\{k\}) = \sum_{k \geqslant 0} p_k = \sum_{k \geqslant 0} \exp(-\lambda) \frac{\lambda^k}{k!} = \exp(-\lambda)\exp(\lambda) = 1.$$

2.2.2 Modèles continus (à densité)

Remarque. On prend Ω un intervalle (fini ou infini[2]) sur $\mathbb{R}$, et $\mathcal{F} = \mathcal{B}(I)$, la tribu des boréliens sur I[3].

Définition 2.16. Soit $f : I \to \mathbb{R}^+$ une fonction intégrable telle que $\int_{\mathbb{R}} f(x)\,dx = 1$. Soit $A \in \mathcal{F}$, on pose $\mathbb{P}(A) = \int_A f(x)\,dx$. f est appelée fonction de densité de modèle stochastique.

Définition 2.17 (Loi uniforme continue)**.** On prend $I = [a, b]$ avec $a < b ß\mathbb{R}$. Le modèle uniforme est défini par f constante :

$$f(x) = \begin{cases} 0 & \text{si } x \notin [a, b] \\ \frac{1}{b-a} & \text{si } x \in [a, b] \end{cases}.$$

Remarque. On remarque effectivement $\int_{\mathbb{R}} f(x)\,dx = 1$:

$$\int_{\mathbb{R}} f(x)\,dx = \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{b} f(x)\,dx + \int_{b}^{+\infty} f(x)\,dx = 0 + \frac{1}{b-a}\int_a^b dx + 0 = 1.$$

Définition 2.18 (Modèle exponentiel)**.** [4] On prend $I = \mathbb{R}^+$ et $\lambda > 0$. Le modèle exponentiel est défini par :

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}.$$

Remarque. On peut calculer l'intégrale impropre comme suit :

$$\begin{aligned} \int_{\mathbb{R}} f(x)\,dx &= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{+\infty} f(x)\,dx = 0 + \lim_{M \to +\infty} \int_0^M f(x)\,dx \\ &= \lim_{M \to +\infty} \big[-\exp(-\lambda x)\big]_0^M = \lim_{M \to +\infty} \big(1 - \exp(-\lambda M)\big) = 1. \end{aligned}$$

Définition 2.19 (Modèle gaussien)**.** [5] On prend $I = \mathbb{R}$, et $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_0^+$. Le modèle gaussien est défini par :

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

[2] On parle d'intervalle fini pour $[a, b]$, avec $a < b \in \mathbb{R}$ et d'intervalle semi-infini pour $(-\infty, b]$ ou $[a, +\infty)$ et d'intervalle infini pour $(-\infty, +\infty) = \mathbb{R}$.

[3] Ou encore la tribu engendrée par les intervalles de I.

[4] Également appelé *modèle des files d'attente*.

[5] Également appelé *modèle des erreurs* ou encore *modèle normal*.

Remarque. Pour que $\mathbb{P}$ soit une probabilité, il faut que f soit définie positive. Or f est une exponentielle multipliée par un coefficient positif. Il faut également $\int_{\mathbb{R}} f(x)\,dx = 1$, ce qui peut se vérifier par :

$$\int_{\mathbb{R}} f(x)\,dx,$$

en posant $y \coloneqq x - \mu$, et donc $dy = dx$:

$$\int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2\sigma^2}\right).$$

En posant $z \coloneqq \frac{y}{\sigma}$ (et donc $dz = \frac{dx}{\sigma}$), on obtient :

$$\int_{\mathbb{R}} f(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{z^2}{2}\right) dz.$$

Une primitive de $\exp\left(-\frac{z^2}{2}\right)$ est :

$$\int_{-\infty}^{z} \exp\left(-\frac{x^2}{2}\right) \frac{dx}{\sqrt{2\pi}} = \mathrm{Erf}(z).$$

On écrit alors :

$$\begin{aligned}
\mathbb{P}(\Omega)^2 &= \left(\int_{\mathbb{R}} \exp\left(-\frac{x^2}{2}\right) \frac{dx}{\sqrt{2\pi}}\right) \left(\int_{\mathbb{R}} \exp\left(-\frac{y^2}{2}\right) \frac{dy}{\sqrt{2\pi}}\right) \\
&= \iint_{\mathbb{R}^2} \exp\left(-\frac{x^2+y^2}{2}\right) \frac{dx\,dy}{2\pi}.
\end{aligned}$$

En passant en coordonnées polaires, on obtient :

$$\mathbb{P}(\Omega)^2 = \int_{-\pi}^{+\pi} \int_{\mathbb{R}} \exp\left(-\frac{r^2}{2}\right) \frac{r\,dr\,d\theta}{2\pi} = \int_{-\pi}^{+\pi} \frac{d\theta}{2\pi} \int_{\mathbb{R}} r \exp\left(-\frac{r^2}{2}\right) dr = \left[-\exp\left(-\frac{r^2}{2}\right)\right]_0^{+\infty} = 1.$$

On en déduit alors $\mathbb{P}(\Omega) = 1$ également. $\mathbb{P}$ est donc bien une probabilité.

Définition 2.20. On a défini une probabilité sur $\left(R^+, (R^+)\right)$ via la fonction $f(r) = r \exp\left(-\frac{r^2}{2}\right)$. On l'appelle la *probabilité de Rayleigh*.

2.2.3 Divergence sur la fonction Gamma d'Euler

Définition 2.21 (Fonction Gamma d'Euler)**.** La fonction Gamma d'Euler est définie comme suit :

$$\Gamma : \mathbb{R}_0^+ \to \mathbb{R} : x \mapsto \int_0^{+\infty} \exp(-x)x^{t-1}\,dx.$$

Remarque. On note $\gamma \coloneqq -\Gamma'(1) > 0$ la constante d'Euler-Mascheroni. La question $\gamma \overset{?}{\in} \mathbb{Q}$ est toujours ouverte.

Proposition 2.22. $\forall t > 0 : \Gamma(t+1) = t\Gamma(t)$.

Démonstration. Soit $t > 0$. Par l'intégration par parties, on a :

$$\Gamma(t+1) = \int_0^{+\infty} \exp(-x)x^t \, dx = \left[-x^t \exp(-x)\right]_0^{+\infty} + t\int_0^{+\infty} \exp(-x)x^{t-1} \, dx = t\Gamma(t).$$

□

Remarque. Par la proposition 2.22, on peut définir la factorielle de tout nombre naturel par :

$$\forall n \in \mathbb{N}^* : n! = \Gamma(n+1)$$

Proposition 2.23 (Formule des compléments)**.** *Soit* $t \in (0,1)$. *Alors :*

$$\Gamma(t)\Gamma(1-t) = \frac{\pi}{\sin(\pi t)}.$$

2.2.4 Retour aux modèles stochastiques

Définition 2.24 (Modèle Gamma)**.** [6] On prend $\Omega = \mathbb{R}^+$. Le modèle Gamma est défini par :

$$f_t(x)\frac{x^t - \exp(-x)}{\Gamma(t)}.$$

2.3 Notion de variables aléatoires

2.3.1 Cas discret

Définition 2.25. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une variable aléatoire discrète[7] est une application $X : \Omega \to E$ où E est un ensemble fini ou infini dénombrable. On demande à cette application d'être mesurable.

Remarque.

- Bien souvent, on a $E = \Omega$, et $X(\omega) = \omega$. Dans ce cas, on *identifie* l'espace des chances avec l'espace d'arrivée. La probabilité $\mathbb{P}$ s'appelle alors la **loi** de la variable aléatoire X.
- Il arrive parfois que l'espace de probabilités soit plus gros que l'espace d'état.

Définition 2.26. Plus formellement, la **loi** d'une v.a.d. X est l'ensemble :

$$\{\mathbb{P}(X = x) \text{ t.q. } x \in E\}.$$

Définition 2.27. Pour toute valeur $k \in E$ que peut prendre la variable aléatoire X, on note $\mathbb{P}(X = k)$ la probabilité que la variable X prenne la valeur k. C'est équivalent à $\mathbb{P}\left(X(\omega) = k\right)$ pour $\omega \in \Omega$.

Définition 2.28. Lorsqu'une v.a.d. X suit une certaine loi $\mathcal{L}$, on note $X \sim \mathcal{L}$.

Par exemple, une variable Y suivant une poisson de paramètre λ se note $Y \sim \mathcal{P}(\lambda)$.

[6] Le modèle Γ est une généralisation du modèle exponentiel (définition 2.18).
[7] Souvent écrite v.a.d. ou V.A.-D.

2.3.2 Cas absolument continu

Définition 2.29. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une variable aléatoire absolument continue[8] est une application $X : \Omega \to \mathbb{R}$ mesurable au sens où :

$$\forall A \in \mathcal{B}(\mathbb{R}) : \{\omega \in \Omega \,\text{t. q.}\, X(\omega) \in A\} \in \mathcal{F},$$

et *absolument continue* au sens où :

$$\exists f_X : \mathbb{R} \to \mathbb{R}^+$$

mesurable et telle que :

$$\int_{\mathbb{R}} f_X(x)\,dx = 1,$$

avec :

$$\mathbb{P}(X \in A) = \int_A f_X(x)\,dx. \tag{1}$$

Définition 2.30. On appelle f_X la **densité** de X.

Remarque. La *loi* de X est donnée par (1).

Définition 2.31. On note $F_X(t) = \mathbb{P}(X \leqslant t)$, ou encore $F_X(t) = \int_{-\infty}^{t} f(x)\,dx$ (en prenant $A = (-\infty, t]$).

Remarque. La fonction $t \mapsto F_X(t)$ est continue et est (presque) partout dérivable avec :

$$\frac{\partial F_X}{\partial t}(t) = f_X(t) \geqslant 0.$$

Donc F_X est croissante avec :

$$\lim_{t \to -\infty} F_X(t) = 0,$$

et :

$$\lim_{t \to +\infty} F_X(t) = 1.$$

Remarque. On peut associer une fonction de répartition F_X à toute variable aléatoire X, même si X est une v.a.d. Dans ce cas, on construit F_X constante par morceaux (et présente donc des points de discontinuité).

Définition 2.32. Si F_X est continue, on dit que X est continue.

Remarque. Donc si X est continue, alors $\mathbb{P}(X = x) = F_X(x) - \lim_{y \to x} F_X(y) = 0$. Ce résultat peut également être observé en utilisant le fait que $\mathbb{P}(X = x) = \int_x^x f(x)\,dx$, et une intégration sur un point est nulle.
Remarque. Il existe des fonction continues nulle part dérivables. On peut donc avoir $F_X(t)$ continue mais pas sous la forme suivante :

$$F_X(t) = \int_{-\infty}^{t} f(x)\,dx, \tag{2}$$

pour une fonction f_X donnée.

Définition 2.33. On dit qu'une variable fonction $f : \mathbb{R} \to \mathbb{R}$ est **absolument continue** si elle admet une représentation intégrale de type (2).

Définition 2.34. Soit E un ensemble. La fonction 1_E est appelée **fonction indicatrice** est est définie telle que :

$$\forall x : 1_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{sinon} \end{cases}.$$

[8] Souvent écrite v.a.c. ou V.A.-C.

Exemples

1. Si $X_1 \sim U_{[a,b]}$ est une v.a.c. uniforme sur $[a, b]$, alors :

$$F_{X_1}(t) = \begin{cases} 0 & \text{si } t \leqslant a \\ t - a & \text{si } a < t < b \\ 1 & \text{si } t > b \end{cases}.$$

2. Si $X_2 \sim \text{Exp}(\lambda)$ est une v.a.c. exponentielle de paramètre λ, alors :

$$F_{X_2}(t) = \int_{-\infty}^{t} \lambda \exp(-\lambda t) 1_{(0,+\infty)}(t) = -\exp(-\lambda t) 1_{(0,+\infty)}(t).$$

3. Si $x_3 \sim \mathcal{N}(\mu, \sigma^2)$ est une v.a.c. normale de moyenne μ est de variance σ^2, alors :

$$F_{X_3}(t) = \int_{-\infty}^{t} f(x)\, dx = \text{Erf}\left(\frac{t-\mu}{\sigma}\right).$$

4. Si $X_4 \sim \mathcal{C}$ est une v.a.c. de Cauchy de densité donnée par :

$$f_{X_4}(x) = \frac{1}{\pi(1+x^2)},$$

alors :

$$F_{X_4}(t) = \frac{1}{2} + \frac{1}{\pi} \arctan(t).$$

2.4 Théorème de de Moivre-Laplace

Soient $p \in (0, 1)$ et $n \geqslant 1$. On pose $X_{n,p} \sim \mathcal{B}(n, p)$.
Soit $Y_{n,p}$ défini par :

$$Y_{n,p} := \frac{X_{n,p} - np}{\sqrt{np(1-p)}}.$$

On remarque que $Y_{n,p}$ est une binomiale renormalisée.

Théorème 2.35 (Théorème de de Moivre-Laplace)**.** *Si* $t \in \mathbb{R}$*, alors :*

$$\mathbb{P}(Y_{n,p} \leqslant t) \overset{n\to+\infty}{\to} F_{\mathcal{N}(0,1)}(t).$$

Remarque. La signification de ce théorème est qu'une binomiale renormalisée se comporte comme une gaussienne $\mathcal{N}(0, 1)$ lorsque $n \to +\infty$.

Proposition 2.36 (Formule de Stirling)**.**

$$n! \overset{n\to+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

2.5 Convergence en loi

Définition 2.37.

- Soit Z une v.a.c. Soit $\{Z_n, n \geqslant 1\}$ une suite de v.a. quelconques. On dit que Z_n *converge en loi vers* Z si :

$$\forall x : F_{Z_n}(x) = \mathbb{P}(Z_n \leqslant x) \overset{n \to +\infty}{\to} F_Z(x).$$

  On note cela :

$$Z_n \overset{\mathcal{D}}{\to} Z.$$

- Soient Z une v.a.d. et une $\{Z_n, n \geqslant 1\}$ une suite de variables aléatoires discrètes. On dit que Z_n *covnerge en loi vers* Z si :

$$\forall x \in E : \mathbb{P}(Z_n = x) \overset{n \to +\infty}{\to} \mathbb{P}(Z = x).$$

  On note cela :

$$Z_n \overset{\mathcal{D}}{\to} Z.$$

Remarque. Un exemple typique de convergence en loi de variables discrètes est :

$$\mathcal{B}\left(n, \frac{\lambda}{n}\right) \overset{\mathcal{D}}{\to} \mathcal{P}(\lambda).$$

3 Espérance

3.1 Pari de pascal

Le terme *espérance* vient de Blaise Pascal et de son traitement de la question « Faut-il croire en Dieu ? ». On pose la variable X qui décrit le résultat de l'existence de Dieu définie comme suit :

$$X = \begin{cases} 0 & \text{si Dieu n'existe pas} \\ +\infty & \text{sinon} \end{cases}.$$

On a alors $\mathbb{P}(X = 0) = p$ et $\mathbb{P}(X = +\infty) = 1 - p$. Prenons $p < 1$ (car si $p = 1$, on suppose que Dieu n'existe pas). Alors $\bar{X}$, la valeur moyenne de X est donnée par :

$$\bar{X} = p \cdot 0 + (1 - p) \cdot +\infty.$$

Blaise Pascal a appelé cette valeur **espérance** et l'a noté $\mathbb{E}(X)$.

3.2 Espérance et variables aléatoires

Remarque. Il existe plusieurs méthodes pour décrire le comportement d'une variable aléatoire. On s'intéresse ici aux **indicateurs de position**. Il existe d'autres types d'indicateurs dont les **indicateurs de répartition** qui seront vus plus loin.

Définition 3.1. On considère X une variable aléatoire sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

1. <u>Méthode de la médiane :</u>

   On évalue le nombre $\widetilde{x}$ tel que $\mathbb{P}(X \leqslant \widetilde{x}) = \mathbb{P}(X \geqslant \widetilde{x}) = \frac{1}{2}$.

   Lorsque X est continue, la médiane existe toujours. Si X est discrète, la médiane n'existe pas obligatoirement et n'est pas forcément unique.

2. <u>Méthode de l'espérance :</u>

   On évalue une *moyenne pondérée* des valeurs que peut prendre X par leur probabilité.

3.3 Définition de l'espérance

3.3.1 Cas positif

Définition 3.2 (Cas discret)**.** Soit X une v.a.d. à valeurs positives. On note $p_k := \mathbb{P}(X = x_k)$. On pose :

$$\mathbb{E}(X) = \sum_{k \geqslant 0} p_k x_k.$$

Remarque. Dans ce cas, l'espérance fait toujours sens et existe toujours mais peut valoir $+\infty$.

Définition 3.3 (Cas absolument continu)**.** Soit X une v.a.c. définie positive de densité f_X et de répartition F_X. On note :

$$\begin{cases} F_X(t) = \mathbb{P}(X \leqslant t) \\ f_X(t) = \frac{\partial}{\partial t} F_X(t) \end{cases}.$$

On définit alors :

$$\mathbb{E}(X) = \int_0^{+\infty} xf(x)\,dx.$$

Remarque.

- L'intégrale démarre en 0 car la variable aléatoire X est définie positive ;
- à nouveau, l'espérance existe toujours mais peut valoir $+\infty$.

3.3.2 Cas général

Définition 3.4 (Cas discret)**.** Soit X une v.a.d. à valeurs dans $E = \{x_0, \dots, x_n\} \subset \mathbb{R}$ fini ou infini dénombrable (typiquement $E = \mathbb{Z}$). On pose $p_n \coloneqq \mathbb{P}(X = x_n)$. On considère la série à termes positifs :

$$\sum_{x_n \in E} |x_n|\, p_n.$$

Si la série vaut $+\infty$, on dit que X *n'est pas intégrable* et on ne peut pas définir son espérance.

Si la série est finie, alors le théorème 1.2 entraine que la série :

$$\sum_{x_n \in E} x_n p_n$$

converge également.

On définit alors :

$$\mathbb{E}(X) = \sum_{x_n \in E} x_n p_n. \tag{3}$$

Définition 3.5 (Cas absolument continu (à densité))**.** Soit X une v.a.c. de densité f_X sur $\mathbb{R}$ telle que $\forall A \in \mathcal{B}(\mathbb{R})$:

$$\mathbb{P}(X \in A) = \int_A f(x)\,dx.$$

On considère :

$$I = \int_{-\infty}^{+\infty} |x|\, f(x)\,dx. \tag{4}$$

Si $I = +\infty$, on dit que N *n'est pas intégrable* et on ne peut pas définir son espérance.

Si $E < +\infty$, le théorème 1.3 entraine que l'intégrale :

$$\int_{-\infty}^{+\infty} xf_X(x)\,dx$$

converge également.

On définit alors :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)\,dx. \tag{5}$$

3.4 Exemples d'espérance

Exemple 1. (exemple de 3.2.) Soit $X \sim \mathcal{B}(n, p)$ une binomiale. Par définition, on évalue :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} k = \sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} k = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{\gamma=0}^{n-1} \binom{n-1}{\gamma} p^{\gamma} (1-p)^{n-1-\gamma} = np(p + (1-p))^{n-1} = np.\end{aligned}$$

Exemple 2. (exemple de 3.2.) Soit $X \sim \mathcal{P}(\lambda)$ une poisson de paramètre λ. On évalue :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k \geqslant 0} \mathbb{P}(X = k) k = \sum_{k \geqslant 0} \exp(-\lambda) \frac{\lambda^k}{k!} k = \exp(-\lambda) \sum_{k \geqslant 1} \frac{\lambda^k}{k!} k \\ &= \exp(-\lambda) \lambda \sum_{k \geqslant 1} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \exp(-\lambda) \sum_{\gamma \geqslant 0} \frac{\lambda^{\gamma}}{\gamma!} = \lambda \exp(-\lambda) \exp(\lambda) = \lambda.\end{aligned}$$

Exemple 3. (exemple de 3.2.) Soit $X \sim$ Bâle. La loi de X est donnée par $\mathbb{P}(X = k) = \frac{6}{(\pi k)^2}$ pour tout $k \geqslant 1$. On évalue :

$$\mathbb{E}(X) = \sum_{k \geqslant 1} \mathbb{P}(X = k) k \sum_{k \geqslant 1} \frac{6}{\pi^2} \sum_{k \geqslant 1} \frac{1}{k^2} k = \frac{6}{\pi^2} \sum_{k \geqslant 1} \frac{1}{k} = +\infty.$$ [9]

Exemple 4. (exemple de 3.3.) Soit $X \sim \mathrm{Exp}(\lambda)$ une exponentielle négative[10]. On sait :

$$f_X(t) = \lambda \exp(-\lambda t),$$

et donc on calcule :

$$\mathbb{E}(X) = \int_0^{+\infty} x f_X(x)\, dx = \int_0^{+\infty} x \lambda \exp(-\lambda x)\, dx.$$

On pose $y := \lambda x$ (et donc $dy = \lambda\, dx$), et on obtient :

$$E(X) = \frac{1}{\lambda} \int_0^{+\infty} y \exp(-y)\, dy = \frac{1}{\lambda}.$$

Exemple 5. (exemple de 3.3.) Soit $X \sim U_{(a,b)}$ une uniforme sur (a, b) où $0 \leqslant a < b \in \mathbb{R}$. On sait :

$$f_X(t) = \frac{1}{b-a} 1_{(a,b)}(t),$$

et donc, on calcule :

$$\mathbb{E}(X) = \int_0^{+\infty} x f_X(x)\, dx = \int_a^b \frac{x}{b-a} = \left[\frac{x^2}{2(b-a)} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

Exemple 6. (exemple de 3.3.) Soit $X \sim \frac{1}{2}\mathcal{C}$ une demi-Cauchy. On sait (pour $t \in [0, +\infty)$) :

$$f_X(t) = \frac{2}{\pi(1 + t^2)},$$

[9] La série $\sum_{k \geqslant 1} \frac{1}{k} = +\infty$ se démontre en utilisant le fait que $\sum_{i=2^{\alpha}}^{2^{\alpha+1}} \frac{1}{i} \gneqq \frac{1}{2} \forall \alpha \in \mathbb{N}$ et donc en faisant tendre $\alpha \to +\infty$, on obtient $+\infty$.

[10] la notion d'exponentielle *négative* vient du fait que le paramètre de la fonction exponentielle est négatif, mais la fonction exponentielle est définie positive.

et donc, on calcule :

$$\mathbb{E}(X) = \int_0^{+\infty} x f_X(x)\,dx = \int_0^{+\infty} \frac{2x}{(1+x^2)}\frac{dx}{\pi} = \frac{1}{\pi}\left[\log(1+x^2)\right]_0^{+\infty} = +\infty$$

Exemple 7. (exemple de 3.5.) Soit $X \sim \mathcal{C}$ une Cauchy. On sait :

$$f_X(t) = \frac{1}{\pi(1+t^2)},$$

et donc, on calcule :

$$\int_{-\infty}^{+\infty} x f_X(x)\,dx = 2\int_0^{+\infty} |x|\, f_X(x)\,dx = 2\infty = +\infty.$$

Proposition 3.6 (Critère de d'Alembert)**.** *Soit* f *une fonction définie positive sur* $[1, +\infty)$*. On suppose* $f(x) \sim \frac{c}{x^a}$ *quand* $x \to +\infty$ *avec* $a, c \in \mathbb{R}$*.*

Si $a < 1$*, alors :*

$$\int_1^{+\infty} f(x)\,dx < +\infty,$$

et si $a \geqslant 1$*, alors :*

$$\int_1^{+\infty} f(x)\,dx = +\infty.$$

Remarque. Dans les cas des Cauchy, on a $a = 1$ et $c = \frac{1}{\pi}$. Donc, par d'Alembert, on sait que l'intégrale est infinie. On n'a pas besoin de primitive explicite.

Remarque. Pour les variables aléatoires continues n'ayant pas d'espérance, on peut s'intéresser à la médiane $m \in \mathbb{R}$ telle que :

$$\int_{-\infty}^{m} f(x)\,dx = \int_m^{+\infty} f(x)\,dx = \frac{1}{2}.$$

Si f s'annule en certaines x, il se peut que m ne soit pas unique.

Une variable $X \sim \mathcal{C}$ Cauchy est paire, et donc $m = 0$.

3.5 Espérance de fonctions de variables aléatoires

Définition 3.7. Soit X une v.a.c. réelle et $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable. La quantité $Y = g(X)$ est une variable aléatoire car c'est une application : $Y : \Omega \to \mathbb{R} : \omega \mapsto (g \circ X)(\omega)$.

Théorème 3.8 (Principe de transfert)**.** *Soient* X *une v.a. et* $Y := g(X)$*. On suppose* $\mathbb{E}(|Y|) < +\infty$[11]*, alors :*

$$\mathbb{E}(Y) = \begin{cases} \displaystyle\sum_{x_n \in E} g(x_n)\mathbb{P}(X = x_n) & \text{si } X \text{ est discrète} \\ \displaystyle\int_{\mathbb{R}} g(x) f_X(x)\,dx & \text{si } X \text{ est absolument continue} \end{cases}. \tag{6}$$

Remarque. Ce théorème signifie que pour déterminer l'espérance de $g(X)$, on intègre $g(x)$ le long de la loi de X.

[11] Ainsi, $\mathbb{E}(Y)$ a un sens.

Exemple 8 (Calcul du moment d'ordre 2). On prend $g : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$.

Cas discret : soit X à valeurs dans $E := \{x_0, x_1, \ldots\}$. On prend $Y := g(X)$. Alors, l'espérance est donnée par :

$$\mathbb{E}(Y) = \mathbb{E}(g(Y)) = \sum_{x_n \in E} (x_n)^2 \mathbb{P}(X = x_n).$$

Si $X \sim \mathcal{P}(\lambda)$ est une poisson de paramètre λ, on a $E = \mathbb{N}$. Dès lors, on considère $x_n = n$. L'espérance est alors :

$$\begin{aligned}
\mathbb{E}(Y) = \mathbb{E}(X^2) &= \sum_{n \geqslant 0} n^2 \mathbb{P}(X = n) = \sum_{n \geqslant 1} n^2 \exp(-\lambda) \frac{\lambda^n}{n!} = \exp(-\lambda) \sum_{n \geqslant 1} n \frac{\lambda^n}{(n + -1)!} \\
&= \exp(-\lambda) \sum_{n \geqslant 1} \big((n-1) + 1\big) \frac{\lambda^n}{(n-1)!} = \exp(-\lambda) \left[\lambda \sum_{n \geqslant 1} \frac{\lambda^{n-1}}{(n-1)!} + \sum_{n \geqslant 1} (n-1) \frac{\lambda^n}{(n-1)!} \right] \\
&= \lambda + \exp(-\lambda) \sum_{n \geqslant 2} \frac{\lambda^n}{(n-2)!} = \lambda + \lambda^2 \exp(-\lambda) \sum_{n \geqslant 2} \frac{\lambda^{n-2}}{(n-2)!} \\
&= \lambda + \lambda^2.
\end{aligned}$$

Cas absolument continu : soit X à valeurs dans $\mathbb{R}$ de densité f_X. On prend $Y := g(X)$. Alors, l'espérance est donnée par :

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{\mathbb{R}} x^2 f_X(x) \, dx.$$

Si $X \sim \mathrm{Exp}(\lambda)$ est une exponentielle négative de paramètre λ, on a (en posant $y := \lambda x$ et donc $dy = \lambda \, dx$) :

$$\mathbb{E}(Y) = \mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \lambda \exp(-\lambda x) \, dx = \frac{1}{\lambda^2} \int_{\mathbb{R}} y^2 \exp(-y) \, dy = \frac{\Gamma(3)}{\lambda^2} = \frac{2!}{\lambda^2} = \frac{2}{\lambda^2}.$$

Définition 3.9. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités. On note :

$$\mathcal{L}_1 (\Omega, \mathcal{F}, \mathbb{P}) := \{X \text{ v.a. t. q. } \mathbb{E}(|X|) < +\infty\}. \tag{7}$$

Théorème 3.10 (Propriété fondamentale de l'espérance)**.** *L'espace $\mathcal{L}_1 (\Omega, \mathcal{F}, \mathbb{P})$ est un espace vectoriel réel de dimension infinie. Donc :*

$$\forall X, Y \in \mathcal{L}_1 (\Omega, \mathcal{F}, \mathbb{P}), \lambda, \mu \in \mathbb{R} : \lambda X + \mu Y \in \mathcal{L}_1 (\Omega, \mathcal{F}, \mathbb{P}).$$

De plus, $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.

Remarque. On dit alors que l'espérance est un opérateur linéaire.

Théorème 3.11. *Si $X \geqslant 0$ est une v.a. définie positive, alors son espérance est positive.*

4 Variance

4.1 Définitions

Définition 4.1. Soit $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. On pose $Y \coloneqq (X - \mathbb{E}(X))^2$. Par définition, Y est positive, et donc on peut définir son espérance. On pose :

$$\mathrm{Var}(X) \coloneqq \mathbb{E}(Y). \tag{8}$$

Remarque. Il est possible que $\mathrm{Var}(X) = \mathbb{E}(Y) = +\infty$. Dans ce cas, on dit que X est de variance infinie.
Remarque. La variance est un indicateur de répartition par rapport à la moyenne. Dans le cas où la variance est infinie, on regarde les quantités $\mathbb{P}(|X - \mathbb{E}(X)| \geqslant x)$ pour $x > 0$ par analogie à la médiane.

Proposition 4.2. *Soit* X *une v.a. de variance* $\mathrm{Var}(X) < +\infty$. *Alors :*

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2. \tag{9}$$

Démonstration. On observe que :

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

□

Remarque. La variance est une valeur positive. Donc on a $\mathbb{E}(X^2) - \mathbb{E}(X)^2 \geqslant 0$, et donc $\mathbb{E}(X^2) \geqslant \mathbb{E}(X)^2$.

Proposition 4.3. *La proposition 4.2 peut se retrouver à l'aide de l'inégalité de Cauchy-Schwartz.*

Démonstration. Dans le cas discret, on pose $y_n \coloneqq x_n\sqrt{\mathbb{P}(X = x_n)}$, et $z_n \coloneqq \sqrt{\mathbb{P}(X = x_n)}$. L'inégalité de Cauchy-Schwartz implique :

$$\left|\sum_{n=1}^{N} y_n z_n\right| \leqslant \sqrt{\left(\sum_{n=1}^{N} y_n^2\right)\left(\sum_{n=1}^{N} z_n^2\right)}.$$

Les valeurs étant positives, on peut passer au carré. En faisant tendre $N \to +\infty$, on obtient :

$$\left(\sum_{n\geqslant 1} y_n z_n\right)^2 \leqslant \left(\sum_{n\geqslant 1} y_n^2\right)\left(\sum_{n\geqslant 1} z_n^2\right).$$

On sait que :

$$\begin{cases} \sum_{n\geqslant 1} z_n^2 = \sum_{n\geqslant 1} \mathbb{P}(X = x_n) = 1, \\ \sum_{n\geqslant 1} y_n^2 = \sum_{n\geqslant 1} (x_n)^2\mathbb{P}(X = x_n) = \mathbb{E}(X^2), \\ \sum_{n\geqslant 1} y_n z_n = \sum_{n\geqslant 1} y_n z_n = \sum_{n\geqslant 1} x_n\mathbb{P}(X = x_n) = \mathbb{E}(X). \end{cases}$$

On a bien :

$$\mathbb{E}(X)^2 \leqslant \mathbb{E}(X^2),$$

qui est l'inégalité (9).

Dans le cas absolument continu, on a X une v.a. de densité f_X sur $\mathbb{R}$. On pose $g(x) = \sqrt{f_X(x)}$ et $h(x) = x\sqrt{f_X(x)}$. L'inégalité de Cauchy-Scwhartz implique :

$$\left|\int_{\mathbb{R}} h(x)g(x)\,dx\right| \leqslant \sqrt{\left(\int_{\mathbb{R}} h(x)^2\,dx\right)\left(\int_{\mathbb{R}} g(x)^2\,dx\right)}.$$

À nouveau, en mettant au carré, on obtient :

$$\begin{aligned}\left(\int_{\mathbb{R}} x f_X(x)\,dx\right)^2 &\leqslant \left(\int_{\mathbb{R}} \left(x\sqrt{f_X(x)}\right)^2 dx\right)\left(\int_{\mathbb{R}} \left(\sqrt{f_X(x)}\right) dx\right) \\ &= \left(\int_{\mathbb{R}} x^2 f_X(x)\,dx\right)\left(\int_{\mathbb{R}} f_X(x)\,dx\right).\end{aligned}$$

On sait que :

$$\begin{cases} \left(\int_{\mathbb{R}} x f_X(x)\,dx\right)^2 = \mathbb{E}(X)^2 \\ \int_{\mathbb{R}} x^2 f_X(x)\,dx = \mathbb{E}(X^2) \\ \int_{\mathbb{R}} f_X(x)\,dx = 1. \end{cases}$$

On a bien :

$$\mathbb{E}(X)^2 \leqslant \mathbb{E}(X)^2,$$

qui est l'inégalité (9). □

Définition 4.4. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités. On pose :

$$\mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P}) \coloneqq \{V \text{ v.a. t. q. } \mathbb{E}(X^2) < +\infty\}.$$

Théorème 4.5. *L'espace $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$ est un espace vectoriel réel de dimension infinie. Donc :*

$$\forall X, Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P}), \lambda, \mu \in \mathbb{R} : \lambda X + \mu Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P}).$$

Remarque. L'application $X \mapsto \mathbb{E}(X^2)$ est une forme quadratique et donc n'est pas linéaire.

Théorème 4.6. *L'espace $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$ est inclus dans l'espace $\mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. Donc si $\mathbb{E}(X^2) < +\infty$, alors $\mathbb{E}(|X|) < +\infty$.*

Théorème 4.7 (Inégalité de Cauchy-Schwartz sur les espaces $\mathcal{L}_i$)**.** *Soient $X, Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$. Alors :*

$$\left|\mathbb{E}(XY)\right| \leqslant \mathbb{E}\left(|XY|\right) \leqslant \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}. \tag{10}$$