(b) *Loi exponentielle de paramètre* $\lambda > 0$.

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

Les lois exponentielles possèdent la propriété caractéristique suivante : si $a, b > 0$,

$$P(X > a + b) = P(X > a)\, P(X > b),$$

ce qu'on interprète en disant que la probabilité que $X - a > b$ sachant que $X > a$ coïncide avec la probabilité que $X > b$. C'est la propriété d'absence de mémoire de la loi exponentielle, qui explique qu'elle soit utilisée par exemple pour modéliser les temps de vie de machine sans usure.

(c) *Loi gaussienne, ou normale,* $\mathcal{N}(m, \sigma^2)$ ($m \in \mathbb{R}$, $\sigma > 0$).

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(x-m)^2}{2\sigma^2}\Big).$$

Avec la loi de Poisson, c'est la loi la plus importante en théorie des probabilités. Sa densité est la fameuse courbe en cloche. Les paramètres m et σ s'interprètent comme

$$m = E[X]\ , \quad \sigma^2 = E[(X-m)^2].$$

On remarque aussi que $X - m$ suit la loi $\mathcal{N}(0, \sigma^2)$. La loi gaussienne jouera un rôle important dans le Chapitre 10.

Par convention on dira qu'une v.a. constante égale à m suit la loi gaussienne $\mathcal{N}(m, 0)$. Si X suit la loi $\mathcal{N}(m, \sigma^2)$, pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda X + \mu$ suit la loi $\mathcal{N}(\lambda m + \mu, \lambda^2\sigma^2)$.

8.1.6 Fonction de répartition d'une variable aléatoire réelle

Si X est une v.a. réelle, la *fonction de répartition* de X est la fonction $F_X : \mathbb{R} \longrightarrow [0,1]$ définie par

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

La fonction F_X est croissante, continue à droite et a pour limite 0 en $-\infty$ et 1 en $+\infty$.

Inversement, si on se donne une fonction F ayant ces propriétés, on a vu dans le cours d'intégration qu'il existe une (unique) mesure de probabilité μ telle que $\mu(]-\infty, t]) = F(t)$ pour tout $t \in \mathbb{R}$. Cela montre qu'on peut interpréter F comme la fonction de répartition d'une v.a. réelle.

Il découle des résultats du cours d'intégration que F_X caractérise la loi P_X de X. On a en particulier

$$\begin{aligned} P(a \le X \le b) &= F_X(b) - F_X(a-) \qquad \text{si } a \le b,\\ P(a < X < b) &= F_X(b-) - F_X(a) \qquad \text{si } a < b, \end{aligned}$$

et les sauts de F_X correspondent aux atomes de P_X.

$$\lim_{\substack{t \to t_0 \\ t > t_0}} F_X(t) = F_X(t_0).$$

Soit $(X_n)_n$ suite de v.a

Soit X v.a

Définitions. a) $X_n \xrightarrow{p.s.} X$ si

$$P\left(\{\omega\in\Omega : \lim_{n\infty} X_n(\omega) = X(\omega)\}\right) = 1$$

(presque sûre)

b) $X_n \xrightarrow{P} X$ si

$$\forall \varepsilon>0, \quad P(|X_n - X| > \varepsilon) \xrightarrow[n\infty]{} 0$$

(en probabilités)

c) $X_n \xrightarrow{\mathcal{L}} X$ si

$\forall f \in C_b(\mathbb{R},\mathbb{R})$ (continue bornée)

$$E(f(X_n)) \xrightarrow[n\infty]{} E(f(X))$$

(en loi)

Théorème. $X_n \xrightarrow{p.s.} X \Longrightarrow X_n \xrightarrow{P} X \Longrightarrow X_n \xrightarrow{\mathcal{L}} X$

contreex 1 $P(X_n = 1) = P(X_n = 0) = \frac{1}{2}$ X_n v.a. indép.

$$\forall f \in C_b(\mathbb{R},\mathbb{R}), \quad E(f(X_n)) = \frac{f(0)+f(1)}{2} \quad (\forall n)$$

$X_n \xrightarrow{\mathcal{L}} X$ avec $P(X=0) = P(X=1) = \frac{1}{2}$.

Mais $\forall \varepsilon > 0$, $P(|X_n - X| > \varepsilon) = P(X_n = 1 \text{ et } X = 0)$
$+ P(X_n = 0 \text{ et } X = 1)$

$X_n \overset{P}{\not\to} X$ $\quad = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \underset{n\infty}{\not\to} 0$

Contre-ex 2.

X_n v.a. indépendantes

$P(X_n=1)=p_n$, $P(X_n=0)=1-p_n$

$X_n \xrightarrow{P} 0 \iff \lim_n p_n = 0$

$$\boxed{X_n \xrightarrow{p.s.} 0 \iff \sum_{n=0}^{\infty} p_n \text{ converge}}$$ (*)

En particulier si $p_n = \frac{1}{n+1}$, $X_n \cancel{\xrightarrow{p.s.}} 0$

$X_n \xrightarrow{P} 0$

(*) :

$$\left\{ \omega : \lim X_n(\omega) = 0 \right\}$$

$$= \bigcap_{k>0} \bigcup_{N \in \mathbb{N}} \bigcap_{n \geq N} \left\{ \omega : |X_n(\omega)| < \frac{1}{k} \right\}$$

$$\Rightarrow P(\{\omega : \lim X_n(\omega) = 0\}) = P\left(\bigcup_N \bigcap_{n \geq N} \{\omega : X_n = 0\}\right)$$

si $\sum_{n=0}^{\infty} p_n$ converge

$$P(\{\omega : \lim X_n(\omega) \neq 0\}) = P\left(\bigcap_N \bigcup_{n \geq N} \{\omega : X_n(\omega) = 1\}\right)$$

$$= \lim_{N\infty} P\left(\bigcup_{n \geq N} \{\omega : X_n(\omega) = 1\}\right)$$

$$\leq \lim_{N\infty} \sum_{n=N}^{+\infty} p_n = 0$$

si $\sum_{n=0}^{\infty} p_n = +\infty$

$$P\Big(\bigcup_{N} \bigcap_{n\geq N} \{X_n = 0\}\Big)$$

$$\leq P\Big(\bigcup_{N=0}^{p} \bigcap_{m\geq N} \{X_n = 0\}\Big)$$

$$\leq P\Big(\bigcap_{m\geq p} \{X_n = 0\}\Big) \leq \underbrace{P\Big(\bigcap_{m=p}^{q} \{X_n = 0\}\Big)}_{\prod_{m=p}^{q} (1-p_m)}$$

$$(1 - x \leq e^{-x}) \qquad \leq \prod_{m=p}^{q} e^{-p_m}$$

$$= e^{-\sum_{n=p}^{q} p_m} \qquad (\forall q > p)$$

$$\Rightarrow P\Big(\bigcap_{m\geq p} \{X_n = 0\}\Big) = 0$$

$$\Rightarrow P\Big(\bigcup_{N} \bigcap_{n\geq N} \{X_n = 0\}\Big) = 0 \qquad \square$$

c-à-d $X_n \overset{p.s.}{\not\to} 0$

$p.s. \Rightarrow P.$

Soit $\varepsilon > 0$, $P(|X_m - X| > \varepsilon)$

$$X_n \xrightarrow{p.s.} X \iff P\left(\bigcap_{k>0} \bigcup_N \bigcap_{n \geq N} \{|X_n - X| < \tfrac{1}{k}\}\right) = 1$$

$$\Rightarrow \forall k, P\left(\bigcup_N \bigcap_{n \geq N} \{|X_n - X| < \tfrac{1}{k}\}\right) = 1$$

$$\Rightarrow \forall k, P\left(\bigcap_N \bigcup_{n \geq N} \{|X_n - X| \geq \tfrac{1}{k}\}\right) = 0$$

$$\Rightarrow \forall k, \lim_{N\infty} \underbrace{P\left(\bigcup_{n \geq N} \{|X_n - X| \geq \tfrac{1}{k}\}\right)}_{\geq P(\{|X_N - X| \geq \frac{1}{k}\})} = 0$$

$$\Rightarrow \forall k, \lim_{N\infty} P(\{|X_N - X| \geq \tfrac{1}{k}\}) = 0.$$

$P \Rightarrow \mathcal{L}.$

Soit $f \in C_b(\mathbb{R}, \mathbb{R})$.

$$\lim_{R \to \infty} P(|X| > R) = 0$$

Soit $\varepsilon > 0$, Soit R tq $P(|X| > R) < \varepsilon$

f uniformément continue sur $[-2R;2R]$.

$\exists\, \eta > 0$, $\forall -2R \leq x, y \leq 2R$, $|x-y| \leq \eta \Rightarrow |f(x) - f(y)| < \varepsilon$

$$\left| E(f(X_n)) - Ef(X) \right| \leq E(|f(X_n) - f(X)|)$$

Or $|f(X_n) - f(X)| \leq \varepsilon$ si $|X_n|, |X| \leq 2R$
$|X_n - X| \leq \eta$

$$|f(X_n) - f(X)| \leq \chi_{\{|X| \leq R\}} |f(X_n) - f(X)|$$
$$+ \chi_{\{|X| > R\}} 2\|f\|_\infty$$

$$\Rightarrow |E(f(X_n) - f(X))| \leq P(|X| > R) \cdot 2\|f\|_\infty$$
$$+ P(|X_n - X| \geq \eta)\, 2\|f\|_\infty$$
$$+ P(|X_n - X| < \eta)\, \varepsilon$$

$$\leq 2\|f\|_\infty \varepsilon + \varepsilon + 2\|f\|_\infty \underbrace{P(|X_n - X| \geq \eta)}_{\xrightarrow[n\infty]{} 0}$$

donc $\limsup_n |E(f(X_n)) - E(f(X))| = 0$

donc $X_n \xrightarrow{\mathcal{L}} X$.

Chapitre 10

Convergence de variables aléatoires

La première partie de ce chapitre présente les différentes notions de convergence de variables aléatoires, et les liens existant entre ces notions. On établit ensuite la loi forte des grands nombres, qui est l'un des deux théorèmes limites fondamentaux de la théorie des probabilités. Le troisième paragraphe présente la convergence en loi des variables aléatoires : ce type de convergence est sans doute le plus délicat à comprendre, en partie parce qu'il s'agit d'une convergence de mesures (ce sont les lois des variables aléatoires qui convergent et non les variables elle-mêmes). La notion de convergence en loi, et le théorème important reliant cette convergence à celle des fonctions caractéristiques, permettent d'arriver au deuxième théorème limite fondamental qui est le théorème central limite.

10.1 Les différentes notions de convergence

Soient $(X_n)_{n\geq 1}$, X des variables aléatoires à valeurs dans $\mathbb{R}^d$, définies sur un espace de probabilité $(\Omega, \mathcal{A}, P)$. On a déjà rencontré plusieurs notions de convergence de la suite (X_n) vers X. En particulier

$$X_n \xrightarrow[n\to\infty]{\text{p.s.}} X \qquad \text{si } P(\{\omega \in \Omega : X(\omega) = \lim_{n\to\infty} X_n(\omega)\}) = 1,$$

et, pour $p \in [1, \infty[$,

$$X_n \xrightarrow[n\to\infty]{L^p} X \qquad \text{si } \lim_{n\to\infty} E[|X_n - X|^p] = 0.$$

Définition 10.1.1 *On dit que la suite (X_n) converge en probabilité vers X, et on note*

$$X_n \xrightarrow[n\to\infty]{\text{(P)}} X$$

si pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0.$$

Proposition 10.1.1 *Soit $\mathcal{L}^0_{\mathbb{R}^d}(\Omega, \mathcal{A}, P)$ l'espace de toutes les variables aléatoires à valeurs dans $\mathbb{R}^d$, et soit $L^0_{\mathbb{R}^d}(\Omega, \mathcal{A}, P)$ son quotient par la relation d'équivalence $X \sim Y$ ssi $X = Y$ p.s. Alors, la formule*

$$d(X, Y) = E[|X - Y| \wedge 1]$$

définit une distance sur $L^0_{\mathbb{R}^d}(\Omega, \mathcal{A}, P)$ qui est compatible avec la convergence en probabilité, au sens où une suite (X_n) converge en probabilité vers X ssi $d(X_n, X)$ tend vers 0. De plus, l'espace $L^0_{\mathbb{R}^d}(\Omega, \mathcal{A}, P)$ est complet pour la distance d.

Preuve. Il est facile de vérifier que d est une distance. De plus, si la suite (X_n) converge en probabilité vers X, on a pour tout $\varepsilon > 0$,

$$E[|X_n - X| \wedge 1] \leq E[|X_n - X| 1_{\{|X_n - X| \leq \varepsilon\}}] + E[(|X_n - X| \wedge 1) 1_{\{|X_n - X| > \varepsilon\}}] \leq \varepsilon + P(|X_n - X| > \varepsilon).$$

D'après la définition de la convergence en probabilité, cela entraîne $\limsup d(X_n, X) \leq \varepsilon$, et puisque ε était arbitraire on a $d(X_n, X) \longrightarrow 0$. Inversement, si $d(X_n, X) \longrightarrow 0$, alors, pour tout $\varepsilon \in]0, 1]$,

$$P(|X_n - X| > \varepsilon) \leq \varepsilon^{-1} E[|X_n - X| \wedge 1] = \varepsilon^{-1} d(X_n, X) \underset{n\to\infty}{\longrightarrow} 0.$$

Il reste à voir que L^0 est complet pour la distance d. Soit donc (X_n) une suite de Cauchy pour la distance d. On peut trouver une sous-suite $Y_k = X_{n_k}$ telle que, pour tout $k \geq 1$,

$$d(Y_k, Y_{k+1}) \leq 2^{-k}.$$

Alors

$$E[\sum_{k=1}^{\infty} (|Y_{k+1} - Y_k| \wedge 1)] = \sum_{k=1}^{\infty} d(Y_k, Y_{k+1}) < \infty,$$

ce qui entraîne $\sum_{k=1}^{\infty} (|Y_{k+1} - Y_k| \wedge 1) < \infty$ p.s., et donc aussi $\sum_{k=1}^{\infty} |Y_{k+1} - Y_k| < \infty$ p.s. (p.s. il ne peut y avoir qu'un nombre fini de valeurs de k pour lesquelles $|Y_{k+1} - Y_k| \geq 1$). On définit ensuite une v.a. X dans L^0 en posant

$$X = Y_1 + \sum_{k=1}^{\infty} (Y_{k+1} - Y_k).$$

Par construction, la suite (Y_k) converge p.s. vers X, et cela entraîne

$$d(Y_k, X) = E[|Y_k - X| \wedge 1] \underset{k\to\infty}{\longrightarrow} 0,$$

par convergence dominée. Donc la suite (Y_k) converge en probabilité vers X, et cela est aussi vrai pour la suite de départ (X_n). □

La preuve précédente montre en particulier que de toute suite qui converge en probabilité on peut extraire une sous-suite qui converge p.s. (vers la même limite). Nous reprenons cette propriété dans l'énoncé suivant.

Proposition 10.1.2 *Si la suite (X_n) converge p.s., ou dans L^p, vers X, elle converge aussi en probabilité vers X. Inversement, si la suite (X_n) converge en probabilité vers X, il existe une sous-suite (X_{n_k}) qui converge p.s. vers X.*

Preuve. La deuxième assertion a déjà été vue. Pour la première, si X_n converge p.s. vers X,

$$d(X_n, X) = E[|X_n - X| \wedge 1] \underset{n\to\infty}{\longrightarrow} 0,$$

par convergence dominée. Si X_n converge dans L^p vers X,

$$d(X_n, X) \leq \|X_n - X\|_1 \leq \|X_n - X\|_p \underset{n\to\infty}{\longrightarrow} 0.$$

□

En résumé la convergence en probabilité est plus faible à la fois que la convergence p.s. et que la convergence dans L^p pour n'importe quel $p \in [1, \infty[$ (et a fortiori pour $p = \infty$). Dans l'autre sens, la convergence en probabilité entraîne la convergence p.s. pour une sous-suite, et la proposition ci-dessous donne des conditions qui permettent de déduire la convergence L^p de la convergence en probabilité.

Proposition 10.1.3 *Soit (X_n) une suite de v.a. convergeant en probabilité vers X. Supposons qu'il existe $r \in]1, \infty[$ tel que la suite (X_n) soit bornée dans L^r. Alors, pour tout $p \in [1, r[$, la suite (X_n) converge vers X dans L^p.*

Preuve. Par hypothèse, il existe une constante C telle que $E[|X_n|^r] \leq C$ pour tout n. Le lemme de Fatou entraîne alors $E[|X|^r] \leq C$ et donc $X \in L^r$. Ensuite, en utilisant l'inégalité de Hölder, on a pour tout $p \in [1, r[$ et tout $\varepsilon > 0$,

$$\begin{aligned} E[|X_n - X|^p] &= E[|X_n - X|^p 1_{\{|X_n - X| \leq \varepsilon\}}] + E[|X_n - X|^p 1_{\{|X_n - X| > \varepsilon\}}] \\ &\leq \varepsilon^p + E[|X_n - X|^r]^{p/r}\, P(|X_n - X| > \varepsilon)^{1-p/r} \\ &\leq \varepsilon^p + 2^p C^{p/r}\, P(|X_n - X| > \varepsilon)^{1-p/r}. \end{aligned}$$

En utilisant l'hypothèse de convergence en probabilité, il vient

$$\limsup_{n\to\infty} E[|X_n - X|^p] \leq \varepsilon^p$$

d'où le résultat annoncé puisque ε est arbitraire. □

10.2 La loi forte des grands nombres

Notre objectif est de montrer que si (X_n) est une suite de v.a. indépendantes et de même loi, dans L^1, alors les moyennes $\frac{1}{n}(X_1 + \cdots + X_n)$ convergent p.s. vers $E[X_1]$. Nous avons déjà obtenu ce résultat sous l'hypothèse supplémentaire que $E[|X_1|^4] < \infty$, mais nous cherchons maintenant à l'établir sous des hypothèses optimales. Nous commençons par un résultat préliminaire important.

Théorème 10.2.1 (Loi du tout ou rien) *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, à valeurs dans des espaces mesurables quelconques. Pour tout $n \geq 1$ soit $\mathcal{B}_n$ la tribu*

$$\mathcal{B}_n = \sigma(X_k\,;\, k \geq n).$$

Commençons par quelques notations. Pour tout entier $k \in \mathbb{N}$, posons

$$M_k = \sup_{0 \le n \le k} (S_n - na),$$
$$M'_k = \sup_{0 \le n \le k} (S_{n+1} - S_1 - na).$$

Alors M_k et M'_k ont même loi : en effet d'une part les vecteurs $(X_1, \ldots, X_k)$ et $(X_2, \ldots, X_{k+1})$ ont même loi et d'autre part on peut écrire $M_k = F_k(X_1, \ldots, X_k)$ et $M'_k = F_k(X_2, \ldots, X_{k+1})$ avec la même fonction (déterministe) $F_k : \mathbb{R}^k \longrightarrow \mathbb{R}$. Il en découle que

$$M = \lim_{k\to\infty} \uparrow M_k$$

et

$$M' = \lim_{k\to\infty} \uparrow M'_k$$

ont aussi même loi (écrire $P(M' \le x) = \lim \downarrow P(M'_k \le x) = \lim \downarrow P(M_k \le x) = P(M \le x)$).

Par ailleurs, il découle des définitions que pour tout $k \ge 1$,

$$M_{k+1} = \sup\Big(0, \sup_{1\le n\le k+1} (S_n - na)\Big) = \sup(0, M'_k + X_1 - a),$$

ce qu'on peut encore réécrire sous la forme

$$M_{k+1} = M'_k - \inf(a - X_1, M'_k).$$

Puisque M'_k a même loi que M_k (et que ces deux v.a. sont clairement dans L^1), on trouve

$$E[\inf(a - X_1, M'_k)] = E[M'_k] - E[M_{k+1}] = E[M_k] - E[M_{k+1}] \le 0$$

grâce à l'inégalité triviale $M_k \le M_{k+1}$. On peut maintenant appliquer le théorème de convergence dominée à la suite des v.a. $\inf(a - X_1, M'_k)$, qui sont dominées en valeur absolue par $|a - X_1|$ (rappelons que $M'_k \ge 0$). Il vient alors

$$E[\inf(a - X_1, M')] = \lim_{k\to\infty} E[\inf(a - X_1, M'_k)] \le 0.$$

Si on avait $P(M = \infty) = 1$, on aurait aussi $P(M' = \infty) = 1$, puisque les v.a. M et M' ont même loi, et donc $\inf(a - X_1, M') = a - X_1$ p.s. Mais alors l'inégalité précédente donnerait $E[a - X_1] \le 0$, ce qui est absurde puisqu'on a choisi $a > E[X_1]$. Cette contradiction termine la preuve. $\square$

10.3 La convergence en loi

Rappelons que $C_b(\mathbb{R}^d)$ désigne l'espace des fonctions continues bornées de $\mathbb{R}^d$ dans $\mathbb{R}$, qu'on munit de la norme sup

$$\|\varphi\| = \sup_{x\in\mathbb{R}^d} |\varphi(x)|.$$

Définition 10.3.1 *Une suite (μ_n) de mesures de probabilité sur $\mathbb{R}^d$ converge étroitement vers une mesure de probabilité μ sur $\mathbb{R}^d$ (on note $\mu_n \xrightarrow{(e)} \mu$) si*

$$\forall \varphi \in C_b(\mathbb{R}^d) \ , \qquad \int \varphi \, d\mu_n \underset{n\to\infty}{\longrightarrow} \int \varphi \, d\mu.$$

Une suite (X_n) de v.a. à valeurs dans $\mathbb{R}^d$ converge en loi vers une v.a. X à valeurs dans $\mathbb{R}^d$ (on note $X_n \xrightarrow{(\mathrm{loi})} X$) si la suite (P_{X_n}) converge étroitement vers P_X. Cela équivaut encore à

$$\forall \varphi \in C_b(\mathbb{R}^d) \ , \qquad E[\varphi(X_n)] \underset{n\to\infty}{\longrightarrow} E[\varphi(X)].$$

Remarques. (i) Il y a un abus de langage à dire que la suite de v.a. (X_n) converge en loi vers X, car la v.a. limite X n'est pas définie de manière unique : seule sa loi P_X l'est (pour cette raison on écrira parfois qu'une suite de v.a. (X_n) converge en loi vers μ mesure de probabilité sur $\mathbb{R}^d$, et il faudra évidemment comprendre que la suite (P_{X_n}) converge étroitement vers μ). Notons aussi qu'on peut considérer la convergence en loi de v.a. définies sur des espaces de probabilité différents (ici nous supposerons toujours implicitement qu'elles sont définies sur le même espace de probabilité), ce qui rend la convergence en loi très différente des autres convergences discutées ci-dessus.

(ii) L'espace des mesures de probabilité sur $\mathbb{R}^d$ peut être vu comme un sous-ensemble du dual $C_b(\mathbb{R}^d)^*$. La convergence étroite correspond alors à la topologie faible * sur le dual (topologie de la convergence simple, les éléments du dual étant vus comme des fonctions sur $C_b(\mathbb{R}^d)$).

Exemples. (a) Si les v.a. X_n et X sont à valeurs dans $\mathbb{Z}^d$, alors X_n converge en loi vers X si et seulement si

$$\forall x \in \mathbb{Z}^d \ , \quad P(X_n = x) \underset{n\to\infty}{\longrightarrow} P(X = x)$$

(l'implication $\Leftarrow$ demande un petit raisonnement : l'argument est facile si on sait, ce qui sera établi plus tard, qu'on peut remplacer $C_b(\mathbb{R}^d)$ par $C_c(\mathbb{R}^d)$ dans la définition de la convergence étroite).

(b) Si les X_n sont des v.a. à densité, $P_{X_n}(dx) = p_n(x)dx$, si on suppose

$$p_n(x) \longrightarrow p(x) \ , \quad dx \text{ p.p.}$$

et s'il existe une fonction $q \geq 0$ telle que $\int_{\mathbb{R}^d} q(x)dx < \infty$ et

$$\forall n \ , \quad p_n(x) \leq q(x) \ , \quad dx \text{ p.p.}$$

alors p est une densité de probabilité sur $\mathbb{R}^d$, et X_n converge en loi vers la loi $p(x)dx$. Cela découle du théorème de convergence dominée.

(c) Si X_n est de loi uniforme sur $\{\frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n}{2^n}\}$, alors X_n converge en loi vers la loi uniforme sur $[0, 1]$. Ce résultat découle de l'approximation de l'intégrale d'une fonction continue par ses sommes de Riemann.

(d) Si X_n est de loi gaussienne $\mathcal{N}(0, \sigma_n^2)$ et si $\sigma_n \longrightarrow 0$, alors X_n converge en loi vers la v.a. constante égale à 0.

Proposition 10.3.1 *Si la suite (X_n) converge en probabilité vers X alors la suite (X_n) converge en loi vers X.*

Preuve. Supposons d'abord que X_n converge p.s. vers X. Alors, pour toute fonction $\varphi \in C_b(\mathbb{R}^d)$, $\varphi(X_n)$ converge p.s. vers $\varphi(X)$ et donc le théorème de convergence dominée entraîne $E[\varphi(X_n)] \longrightarrow E[\varphi(X)]$, d'où la convergence en loi recherchée.

Dans le cas général, raisonnons par l'absurde en supposant que X_n ne converge pas en loi vers X, donc qu'il existe une fonction $\varphi \in C_b(\mathbb{R}^d)$ telle que $E[\varphi(X_n)]$ ne converge pas vers $E[\varphi(X)]$. On peut trouver une sous-suite (n_k) et $\varepsilon > 0$ tels que $|E[\varphi(X_{n_k})] - E[\varphi(X)]| \geq \varepsilon$ pour tout k. Mais, d'après un résultat de la partie 1, il existe une sous-sous-suite (n_{k_ℓ}) telle que $(X_{n_{k_\ell}})$ converge p.s. vers X. La première partie de la preuve donne alors une contradiction. $\square$

Remarque. Il existe un cas où la réciproque de la proposition est vraie. C'est le cas où la v.a. limite X est constante (p.s.). En effet, si X_n converge en loi vers $a \in \mathbb{R}^d$, il découle de la propriété (ii) de la proposition qui suit que pour tout $\varepsilon > 0$,

$$\liminf_{n\to\infty} P_{X_n}(B(a,\varepsilon)) \geq 1$$

où $B(a,\varepsilon)$ est la boule ouverte de centre a et de rayon ε. C'est exactement dire que X_n converge en probabilité vers a.

Si (X_n) est une suite de v.a. convergeant en loi vers X, il n'est pas toujours vrai qu'on ait

$$P(X_n \in B) \longrightarrow P(X \in B)$$

pour tout borélien B de $\mathbb{R}^d$ (prendre $B = \{0\}$ dans l'exemple (d) ci-dessus). On a cependant le résultat suivant.

Proposition 10.3.2 *Soient $(\mu_n), \mu$ des mesures de probabilité sur $\mathbb{R}^d$. Les quatre assertions suivantes sont équivalentes.*

(i) *La suite (μ_n) converge étroitement vers μ.*

(ii) *Pour tout ouvert G de $\mathbb{R}^d$,*
$$\liminf \mu_n(G) \geq \mu(G).$$

(iii) *Pour tout fermé F de $\mathbb{R}^d$,*
$$\limsup \mu_n(F) \leq \mu(F).$$

(iv) *Pour tout borélien B de $\mathbb{R}^d$ tel que $\mu(\partial B) = 0$,*
$$\lim \mu_n(B) = \mu(B).$$

Preuve. Commençons par montrer (i)⇒(ii). Si G est un ouvert de $\mathbb{R}^d$, on peut trouver une suite (φ_p) de fonctions continues bornées telles que $0 \leq \varphi_p \leq 1_G$ et $\varphi_p \uparrow 1_G$ (par exemple $\varphi_p(x) = p\,\mathrm{dist}(x, G^c) \wedge 1$). Alors,

$$\liminf_{n\to\infty} \mu_n(G) \geq \sup_p \Big(\liminf_{n\to\infty} \int \varphi_p d\mu_n\Big) = \sup_p \Big(\int \varphi_p d\mu\Big) = \mu(G).$$

L'équivalence (ii)⇔(iii) est immédiate par passage au complémentaire.

Montrons que (ii) et (iii) entraînent (iv). Si $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\limsup \mu_n(B) \leq \limsup \mu_n(\overline{B}) \leq \mu(\overline{B})$$
$$\liminf \mu_n(B) \geq \liminf \mu_n(\overset{\circ}{B}) \geq \mu(\overset{\circ}{B}).$$

Si $\mu(\partial B) = 0$ on a $\mu(\overline{B}) = \mu(\overset{\circ}{B}) = \mu(B)$ et on obtient (iv).

Il reste à montrer l'implication (iv)⇒(i). Soit $\varphi \in C_b(\mathbb{R}^d)$. Quitte à décomposer $\varphi = \varphi^+ - \varphi^-$ on peut supposer $\varphi \geq 0$. Soit $K > 0$ tel que $0 \leq \varphi \leq K$. Alors le théorème de Fubini montre que

$$\int \varphi(x)\mu(dx) = \int \Big(\int_0^K 1_{\{t \leq \varphi(x)\}} dt \Big) \mu(dx) = \int_0^K \mu(E_t^\varphi) dt,$$

où $E_t^\varphi = \{x \in \mathbb{R}^d : \varphi(x) \geq t\}$. De même, pour tout n,

$$\int \varphi(x)\mu_n(dx) = \int_0^K \mu_n(E_t^\varphi) dt.$$

Remarquons que $\partial E_t^\varphi \subset \{x \in \mathbb{R}^d : \varphi(x) = t\}$, et qu'il existe au plus une infinité dénombrable de valeurs de t telles que

$$\mu(\{x \in \mathbb{R}^d : \varphi(x) = t\}) > 0$$

(en effet il y a au plus k valeurs distinctes de t telles que $\mu(\{x \in \mathbb{R}^d : \varphi(x) = t\}) \geq \frac{1}{k}$). Donc (iv) entraîne

$$\mu_n(E_t^\varphi) \underset{n\to\infty}{\longrightarrow} \mu(E_t^\varphi) \ , \quad dt \text{ p.p.}$$

et par convergence dominée on obtient

$$\int \varphi(x)\mu_n(dx) = \int_0^K \mu_n(E_t^\varphi) dt \underset{n\to\infty}{\longrightarrow} \int_0^K \mu_n(E_t^\varphi) dt = \int \varphi(x)\mu(dx).$$

□

Conséquence. Une suite (X_n) de v.a. réelles converge en loi vers une v.a. X si et seulement si les fonctions de répartition $F_{X_n}(x)$ convergent vers $F_X(x)$ en tout point x où F_X est continue. L'implication ⇒ découle immédiatement de la propriété (iv) ci-dessus. Dans l'autre sens, on observe que sous la condition de convergence des fonctions de répartition (en tout point où F_X est continue), on a pour tout $x \in \mathbb{R}$,

$$\liminf F_{X_n}(x-) \geq F_X(x-),$$
$$\limsup F_{X_n}(x) \leq F_X(x).$$

Il découle de cette observation que la condition (ii) de la proposition est satisfaite pour $\mu_n = P_{X_n}$ et $\mu = P_X$ lorsque G est un intervalle ouvert. Il suffit ensuite d'écrire un ouvert quelconque comme réunion dénombrable disjointe d'intervalles ouverts pour aboutir au résultat désiré.

Rappelons la notation $C_c(\mathbb{R}^d)$ pour l'espace des fonctions continues à support compact sur $\mathbb{R}^d$.

Proposition 10.3.3 *Soient (μ_n) et μ des mesures de probabilité sur $\mathbb{R}^d$. Soit H un sous-ensemble de $C_b(\mathbb{R}^d)$ dont l'adhérence (pour la norme sup) contient $C_c(\mathbb{R}^d)$. Les propriétés suivantes sont équivalentes :*

(i) *La suite (μ_n) converge étroitement vers μ.*

(ii) *On a*

$$\forall \varphi \in C_c(\mathbb{R}^d) \ , \quad \int \varphi \, d\mu_n \underset{n\to\infty}{\longrightarrow} \int \varphi \, d\mu.$$

(iii) *On a*

$$\forall \varphi \in H \ , \quad \int \varphi \, d\mu_n \underset{n\to\infty}{\longrightarrow} \int \varphi \, d\mu.$$

Preuve. Il est évident que (i)$\Rightarrow$(ii) et (i)$\Rightarrow$(iii). Supposons ensuite que (ii) est satisfaite. Soit $\varphi \in C_b(\mathbb{R}^d)$ et soit (f_k) une suite de fonctions dans $C_c(\mathbb{R}^d)$ telles que $0 \leq f_k \leq 1$ et $f_k \uparrow 1$ quand $k \to \infty$. Alors pour tout k, $\varphi f_k \in C_c(\mathbb{R}^d)$ et donc

$$\int \varphi f_k \, d\mu_n \underset{n\to\infty}{\longrightarrow} \int \varphi f_k \, d\mu.$$

Par ailleurs,

$$\begin{aligned} \Big| \int \varphi d\mu_n - \int \varphi f_k \, d\mu_n \Big| &\leq \Big(\sup_{x\in\mathbb{R}} |\varphi(x)| \Big) \Big(1 - \int f_k d\mu_n \Big), \\ \Big| \int \varphi d\mu - \int \varphi f_k \, d\mu \Big| &\leq \Big(\sup_{x\in\mathbb{R}} |\varphi(x)| \Big) \Big(1 - \int f_k d\mu \Big). \end{aligned}$$

Donc, pour tout k,

$$\begin{aligned} \limsup_{n\to\infty} \Big| \int \varphi \, d\mu_n - \int \varphi \, d\mu \Big| &\leq \Big(\sup_{x\in\mathbb{R}} |\varphi(x)| \Big) \Big(\limsup_{n\to\infty} (1 - \int f_k d\mu_n) + (1 - \int f_k d\mu) \Big) \\ &= 2 \Big(\sup_{x\in\mathbb{R}} |\varphi(x)| \Big) (1 - \int f_k d\mu) \Big). \end{aligned}$$

Il suffit maintenant de faire tendre k vers ∞ pour trouver que $\int \varphi d\mu_n$ converge vers $\int \varphi d\mu$, et on a établi (i).

Il reste à montrer (iii)$\Rightarrow$(ii). On suppose donc que la propriété (iii) est satisfaite. Ensuite, si $\varphi \in C_c(\mathbb{R}^d)$, on peut pour chaque entier $k \geq 1$ trouver une fonction $\varphi_k \in H$ telle que $\|\varphi - \varphi_k\| \leq 1/k$. Mais alors, pour tout $k \geq 1$,

$$\begin{aligned} &\limsup_{n\to\infty} | \int \varphi d\mu_n - \int \varphi d\mu | \\ &\leq \limsup_{n\to\infty} \Big(| \int \varphi d\mu_n - \int \varphi_k d\mu_n | + | \int \varphi_k d\mu_n - \int \varphi_k d\mu | + | \int \varphi_k d\mu - \int \varphi d\mu | \Big) \leq \frac{2}{k}. \end{aligned}$$

Comme k est arbitraire cela donne $\int \varphi d\mu_n \longrightarrow \int \varphi d\mu$, d'où la propriété (ii). $\square$