Université d'Artois
Faculté des Sciences Jean Perrin
Probabilités (Master 1 Mathématiques-Informatique)
Daniel Li

Chapitre 3

Théorèmes de convergence pour les sommes de variables aléatoires indépendantes

1 Existence de suites de variables aléatoires indépendantes suivant une loi donnée

Définition 1.1 *Soit $\mathscr{F}$ un ensemble de v.a.r. $X\colon \Omega \to \mathbb{R}$. On dit que les v.a.r. de $\mathscr{F}$ sont indépendante si $X_1, \ldots, X_n$ sont indépendantes pour tout choix d'un nombre fini de v.a.r. $X_1, \ldots, X_n$ dans $\mathscr{F}$.*

Pour une suite $(X_n)_{n\geqslant 1}$ de *v.a.r.*, dire qu'elle est indépendante revient à dire que $X_1, \ldots, X_n$ sont indépendantes pour tout $n \geqslant 1$.

1.1 Cas de deux variables aléatoires réelles

Soit Q une mesure de probabilité sur $\mathbb{R}$.
Considérons sa fonction de répartition, définie, pour $x \in \mathbb{R}$, par :

$$F(x) = Q(]-\infty, x])\,.$$

On sait que F est croissante et continue à droite.

Lorsque F est continue et strictement croissante, F a un inverse. Dans le cas général, on peut définir un **pseudo-inverse** (ou **inverse généralisé de Paul Lévy**), à partir de la remarque suivante : pour tout $t \in]\,0,1\,[$, l'ensemble :

$$\{s \in \mathbb{R}\,;\ F(s) \geqslant t\}$$

est un *intervalle*. On définit le pseudo-inverse de F en posant, pour $t \in]\,0,1\,[$:

$$\boxed{G(t) = \inf\{s \in \mathbb{R}\,;\ F(s) \geqslant t\}}\,.$$

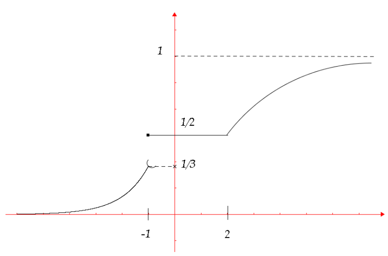
Comme F est continue à droite, la borne inférieure est atteinte et l'on a :

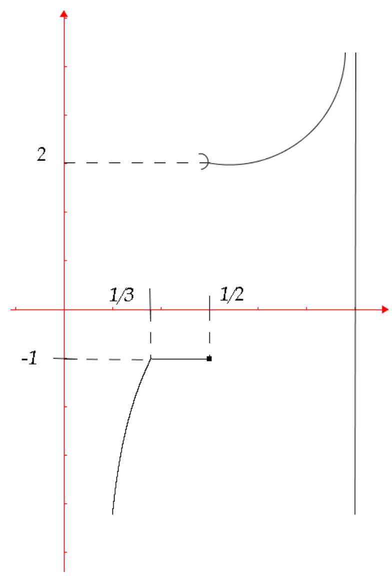
$$\boxed{\{s \in \mathbb{R}\,;\ F(s) \geqslant t\} = [G(t), +\infty\,[}\,. \tag{1.1}$$

G est clairement *croissante*, et elle est **continue à gauche** ; en effet, si $t_n \nearrow t$, alors :

$$\bigcap_{n \geqslant 1} [G(t_n), +\infty\,[= [G(t), +\infty\,[\,,$$

car, si $s \geqslant G(t_n)$ pour tout $n \geqslant 1$, on a $F(s) \geqslant t_n$ pour tout $n \geqslant 1$; donc $F(s) \geqslant t$.





On pourra vérifier à titre d'exercice que $G[F(s)] \leqslant s$ pour tout $s \in \mathbb{R}$ et que $F[G(t)] \geqslant t$ pour tout $t \in]\,0,1\,[$.

Proposition 1.2 (Paul Lévy) *Le pseudo-inverse* $G\colon]\,0,1\,[\to \mathbb{R}$ *de la fonction de répartition de* Q *est une v.a.r. et sa loi est* Q*.*

Plus généralement, si $U\colon \Omega \to]\,0,1\,[$ *est une v.a.r. de loi uniforme sur* $]\,0,1\,[$*, alors* $H = G \circ U\colon \Omega \to \mathbb{R}$ *est une v.a.r. dont la loi est* Q*.*

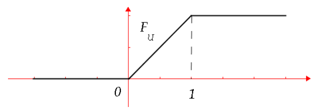
Preuve. G est mesurable car elle est croissante ; c'est donc une *v.a.r.*.

Notons maintenant que (1.1) s'écrit :

$$\boxed{G(t) \leqslant s \quad \Longleftrightarrow \quad F(s) \geqslant t}\,;$$

donc :

$$\begin{aligned} F_H(x) &= \mathbb{P}_H\big(]-\infty, x]\big) = \mathbb{P}\,(H \leqslant x) = \mathbb{P}\,\big(\{\omega\,;\ G\big(U(\omega)\big) \leqslant x\}\big) \\ &= \mathbb{P}\,\big(\{\omega\,;\ F(x) \geqslant U(\omega)\}\big) = F_U\big(F(x)\big) = F(x)\,, \end{aligned}$$



ce qui prouve que $\mathbb{P}_H = Q$. $\square$

Notons qu'avec $\Omega =]\,0,1\,[$, on peut prendre $U(\omega) = \omega$.

On en déduit le résultat suivant.

Proposition 1.3 *Etant données deux lois de probabilité Q_1 et Q_2 sur $\mathbb{R}$, on peut trouver un espace de probabilité $(\Omega, \mathscr{A}, \mathbb{P})$ et deux v.a.r.* **indépendantes** *$X_1, X_2 \colon \Omega \to \mathbb{R}$ dont les lois sont Q_1 et Q_2 respectivement.*

Preuve. Par la proposition précédente, il existe deux *v.a.r.* $G_1, G_2 \colon]\,0,1\,[\to \mathbb{R}$ de lois respectives Q_1 et Q_2. Il s'agit de les rendre indépendantes.

Considérons $\Omega =]\,0,1\,[\times]\,0,1\,[$, muni de sa tribu borélienne et de la mesure de Lebesgue $\mathbb{P} = \lambda_2$. Posons, pour $\omega = (t_1, t_2) \in]\,0,1\,[^2$:

$$\begin{cases} X_1(t_1,t_2) = G_1(t_1) \\ X_2(t_1,t_2) = G_2(t_2)\,. \end{cases}$$

Les *v.a.r.* X_1 et X_2 ont pour lois Q_1 et Q_2 :

$$F_{X_j}(x) = \mathbb{P}\left(X_j(t_1,t_2) \leqslant x\right) = \mathbb{P}\left(G_j(t_j) \leqslant x\right) = Q_j(]-\infty, x])\,,$$

et comme :

$$\begin{aligned} \mathbb{P}_{(X_1,X_2)}([a,b]\times[c,d]) &= \mathbb{P}\left(G_1(x_1) \in [a,b] \text{ et } G_2(x_2) \in [c,d]\right) \\ &= \lambda_2\big(G_1^{-1}([a,b]) \times G_2^{-1}([c,d])\big) \\ &= \lambda\big(G_1^{-1}([a,b])\big).\lambda\big(G_2^{-1}([c,d])\big) \\ &= \lambda_2\big(G_1^{-1}([a,b])\times]\,0,1\,[\big).\lambda_2\big(]\,0,1\,[\times G_2^{-1}([c,d])\big) \\ &= \lambda_2\big(X_1^{-1}([a,b])\big).\lambda_2\big(X_2^{-1}([c,d])\big) \\ &= \mathbb{P}_{X_1}([a,b]).\mathbb{P}_{X_2}([c,d])\,, \end{aligned}$$

on a $\mathbb{P}_{(X_1,X_2)} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2}$, ce qui prouve que X_1 et X_2 sont indépendantes. $\square$

Cette méthode s'étend pour obtenir une *suite infinie* de *v.a.r.* indépendantes de lois données. Mais pour cela, il faudrait d'abord définir ce qu'est un produit infini d'espaces de probabilité. On ne le fera pas et nous allons voir une autre méthode. Il est néanmoins essentiel d'en retenir l'idée : les *v.a.r. indépendantes* s'obtiennent comme *fonctions définies sur des espaces-produits et qui ne sont fonctions que de coordonnées différentes.*

1.2 Cas général

Nous allons définir une suite *infinie* particulière de *v.a.r.* indépendantes.

Commençons par rappeler que tout nombre réel $t \in [0,1[$ a un *unique développement dyadique* :

$$\boxed{t = \sum_{n=1}^{\infty} \frac{\varepsilon_n(t)}{2^n}},$$

avec $\varepsilon_n(t) = 0$ ou 1 et les $\varepsilon_n(t)$ n'étant pas tous égaux à 1 à partir d'un certain rang (développement dyadique *propre* – de toute façon, l'ensemble des t qui ont un développement dyadique impropre est dénombrable, et donc de mesure nulle).

Prouvons d'abord que les fonctions $\varepsilon_n \colon [0,1[\to \{0,1\}$ sont mesurables ; ce seront donc des *v.a.r.*, définies sur $\Omega = [0,1[$, avec pour probabilité la mesure de Lebesgue.

Notons que si

$$t = \sum_{k=1}^{+\infty} \frac{\varepsilon_k(t)}{2^k},$$

alors, si $E(x)$ est la partie entière de x, et $A_n \in \mathbb{N}$:

$$2^{n-1}t = A_n + \frac{\varepsilon_n(t)}{2} + \frac{\varepsilon_{n+1}(t)}{2^2} \cdots = E\left(2^{n-1}t\right) + \frac{\varepsilon_n(t)}{2} + \sum_{k=1}^{+\infty} \frac{\varepsilon_{n+k}(t)}{2^{k+1}}.$$

Comme le développement dyadique de t est propre, une infinité des $\varepsilon_k(t)$, pour $k \geqslant n+1$, est nulle ; on a donc :

$$\sum_{k=1}^{+\infty} \frac{\varepsilon_{n+k}(t)}{2^{k+1}} < \sum_{k=1}^{+\infty} \frac{1}{2^{k+1}} = \frac{1}{2}.$$

Ainsi :

$$\begin{cases} \varepsilon_n(t) = 0 \iff 0 \leqslant 2^{n-1}t - E\left(2^{n-1}t\right) < \dfrac{1}{2} \\ \varepsilon_n(t) = 1 \iff \dfrac{1}{2} \leqslant 2^{n-1}t - E\left(2^{n-1}t\right) < 1. \end{cases}$$

Introduisons alors la fonction indicatrice $R \colon \mathbb{R} \to \{0,1\}$ définie par :

$$R(t) = \begin{cases} 0 & \text{si} \quad 0 \leqslant t - E(t) < 1/2 \\ 1 & \text{si} \quad 1/2 \leqslant t - E(t) < 1; \end{cases}$$

on a, d'après ce qui précède :

$$\varepsilon_n(t) = R(2^{n-1}t).$$

Il est alors clair que chaque ε_n est mesurable et que ($\mathbb{P}$ étant la mesure de Lebesgue sur $[0,1[$) :

$$\mathbb{P}(\varepsilon_n = 0) = \mathbb{P}(\varepsilon_n = 1) = \frac{1}{2};$$

Ce sont donc des *v.a.r.* de Bernoulli.

Pour voir que la suite $(\varepsilon_n)_{n\geqslant 1}$ est indépendante, raisonnons par récurrence ; supposons $\varepsilon_1, \ldots, \varepsilon_n$ indépendantes, et soit, pour $a_1, \ldots, a_{n+1} \in \{0, 1\}$:

$$A_n = \{t \in [0,1[\,;\ \varepsilon_1(t) = a_1, \ldots, \varepsilon_n(t) = a_n\}\,.$$

A_n est l'*intervalle* $[t_n, 1[$ des nombres t dont le développement dyadique commence par :

$$0, a_1 a_2 \ldots a_n\,.$$

La *v.a.r.* ε_{n+1} prend la valeur 0 sur la première moitié de cet intervalle et la valeur 1 sur la seconde moitié. Donc :

$$\mathbb{P}\left(A_n \cap \{\varepsilon_{n+1} = a_{n+1}\right) = \frac{1}{2}\,\mathbb{P}\left(A_n\right),$$

de sorte que, grâce à l'hypothèse de récurrence :

$$\mathbb{P}\left(\varepsilon_1 = a_1, \ldots, \varepsilon_n = a_n, \varepsilon_{n+1} = a_{n+1}\right) = \frac{1}{2^{n+1}} = \prod_{k=1}^{n+1} \mathbb{P}\left(\varepsilon_k = a_k\right),$$

ce qui achève la preuve. □

Proposition 1.4 *Soit* $(Z_n)_{n\geqslant 1}$ *une suite indépendante de v.a.r. de Bernoulli, prenant les valeurs* 0 *et* 1 *avec probabilité* $1/2$*. Alors la somme :*

$$U = \sum_{n=1}^{\infty} \frac{Z_n}{2^n}$$

suit la loi uniforme sur $[0, 1[$.

Preuve. Comme on ne s'intéresse qu'aux lois, on peut prendre des réalisations particulires des *v.a.r.* ; et, avec les réalisations ci-dessus en termes de développement dyadique ; dans ce cas, la proposition exprime simplement le fait que, pour tout $t \in [0, 1[$, on a :

$$t = \sum_{n=1}^{\infty} \frac{\varepsilon_n(t)}{2^n}\,.$$ □

Théorème 1.5 *Pour toute suite* $\mathbb{P}_k$ *de lois de probabilité sur* $\mathbb{R}$*, il existe une suite* indépendante *de v.a.r.* $X_k \colon [0, 1[\to \mathbb{R}$ *telle que, pour tout* $k \geqslant 1$*, la loi* $\mathbb{P}_{X_k}$ *de* X_k *soit* $\mathbb{P}_k$*.*

Preuve. Donnons-nous une *bijection* :

$$\begin{array}{ccc} \mathbb{N}^* \times \mathbb{N}^* & \longrightarrow & \mathbb{N}^* \\ (k, l) & \longmapsto & n(k, l)\,, \end{array}$$

et posons, pour tout $k \geqslant 1$:

$$U_k = \sum_{l=1}^{\infty} \frac{\varepsilon_{n(k,l)}}{2^l} \,.$$

La Proposition 1.4 nous dit que les U_k, $k \geqslant 1$, suivent toutes la loi uniforme sur $[0,1[$; elles sont de plus indépendantes, car la définition de chacune d'elles ne fait intervenir que des éléments différents de la suite indépendante $(\varepsilon_n)_{n\geqslant 1}$. Notons maintenant G_k le pseudo-inverse de la fonction de répartition de $\mathbb{P}_k$; la Proposition 1.2 nous dit que $X_k = G_k \circ U_k$ suit la loi $\mathbb{P}_k$. De plus, la suite $(X_k)_{k\geqslant 1}$ est indépendante puisque $(U_k)_{k\geqslant 1}$ l'est. □

2 Modes de convergence

2.1 Convergence presque sûre et convergence en moyenne

Ces convergences sont bien connues ; nous allons simplement faire quelques rappels.

Soit $(X_n)_{n\geqslant 1}$ une suite de *v.a.r.*, définies sur le *même* espace de probabilité $(\Omega, \mathscr{A}, \mathbb{P})$.

• On dit que $(X_n)_{n\geqslant 1}$ *converge presque sûrement* (en abrégé : *p.s.*) vers la *v.a.r.* X, et l'on notera $X_n \xrightarrow[n\to\infty]{p.s.} X$, si :

$$\mathbb{P}\big(\{\omega \in \Omega \,;\; X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)\}\big) = 1\,.$$

• Si $r \geqslant 1$, et que $X_n, X \in L^r(\mathbb{P})$ (c'est-à-dire qu'elles ont des moments d'ordre r), on dit que $(X_n)_{n\geqslant 1}$ *converge en moyenne d'ordre* r, et l'on écrit $X_n \xrightarrow[n\to\infty]{L^r} X$ ou $X_n \xrightarrow[n\to\infty]{\|\;\|_r} X$, si :

$$\|X_n - X\|_r \xrightarrow[n\to\infty]{} 0.$$

Les cas les plus usuels sont $r = 1$; on dit alors qu'il y a *convergence en moyenne*, et $r = 2$; on dit alors qu'il y a *convergence en moyenne quadratique*.

On a les liens suivant entre ces convergences.

1) Comme $\mathbb{P}(\Omega) = 1$, si $r_1 \leqslant r_2$, on a :

$$\|X_n - X\|_{r_1} \leqslant \|X_n - X\|_{r_2} \,;$$

en particulier *la convergence en moyenne quadratique entraîne la convergence en moyenne.*

2) Le *Théorème de convergence dominée* dit que si $X_n \xrightarrow[n\to\infty]{p.s.} X$ et s'il existe $Y \in L^r(\Omega)$ telle que $|X_n| \leqslant Y$ *p.s.*, pour tout $n \geqslant 1$, alors $X_n \xrightarrow[n\to\infty]{L^r} X$.

3) D'autre part, si $X_n \xrightarrow[n\to\infty]{L^r} X$, on peut extraire une *sous-suite* $(X_{n_k})_{k\geqslant 1}$ qui converge *p.s.* vers X.

2.2 Convergence en probabilité

Définition 2.1 *On dit que la suite de v.a.r. $X_n)_{n\geqslant 1}$* **converge en probabilité** *vers la v.a.r. X si, pour tout $\varepsilon > 0$, on a :*

$$\boxed{\mathbb{P}\left(\{|X_n - X| \geqslant \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0}.$$

On notera $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$.

On notera que, dans la dans la définition de la convergence en probabilité, on peut toujours se resteindre à prendre ε plus petit qu'un nombre $\varepsilon_0 > 0$ donné à l'avance, puisque, plus ε est petit, plus la probabilité $\mathbb{P}(|X_n - X| \geqslant \varepsilon)$ est grande. D'autre part, on peut remplacer $\mathbb{P}(|X_n - X| \geqslant \varepsilon)$ par $\mathbb{P}(|X_n - X| > \varepsilon)$.

Il est utile de remarquer que l'on a le critère suivant :

Proposition 2.2 *La suite $(X_n)_{n\geqslant 1}$ converge en probabilité vers X si et seulement si :*

$$(\forall \varepsilon > 0) \quad (\exists N \geqslant 1) \quad n \geqslant N \Longrightarrow \mathbb{P}(|X_n - X| \geqslant \varepsilon) \leqslant \varepsilon.$$

Preuve. Que la condition soit nécessaire est clair. Inversement, si l'on a cette condition, pour tout $a > 0$ donné, on a, si $0 < \varepsilon \leqslant a$, pour $n \geqslant N$:

$$\mathbb{P}(|X_n - X| \geqslant a) \leqslant \mathbb{P}(|X_n - X| \geqslant \varepsilon) \leqslant \varepsilon,$$

ce qui montre que $\mathbb{P}(|X_n - X| \geqslant a) \underset{n\to\infty}{\longrightarrow} 0$. □

Exemple. Sur $\Omega = [0, 1[$, la suite définie, pour $n = 2^k + l$, $0 \leqslant l \leqslant 2^k - 1$, $k \geqslant 0$, par :

$$X_n = \mathbb{1}_{\left[\frac{l}{2^k}, \frac{l+1}{2^k}\right]}$$

converge en probabilité vers 0, car, pour $\varepsilon \leqslant 1$, on a :

$$\mathbb{P}(|X_n| \geqslant \varepsilon) = \frac{1}{2^k} \quad \text{si } 2^k \leqslant n < 2^{k+1}.$$

Toutefois, cette suite ne converge en aucun point de Ω.

Il est immédiat de voir que si $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$ et $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} Y$, alors $X_n + Y_n \xrightarrow[n\to\infty]{\mathbb{P}} X + Y$ et $aX_n \xrightarrow[n\to\infty]{\mathbb{P}} aX$, pour tout $a \in \mathbb{R}$. On a aussi :

Proposition 2.3 *Si $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$ et $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} Y$, alors $X_n Y_n \xrightarrow[n\to\infty]{\mathbb{P}} XY$.*

Preuve. Soit $\varepsilon > 0$ et $\delta > 0$ donnés, arbitraires.

On a $|X_nY_n - XY| \leqslant |X|\,|Y_n - Y| + |Y_n|\,|X_n - X|$, donc :

$$\mathbb{P}\,(|X_nY_n - XY| \geqslant \varepsilon) \leqslant \mathbb{P}\,(|X|\,|Y_n - Y| \geqslant \varepsilon/2) + \mathbb{P}\,(|Y_n|\,|X_n - X| \geqslant \varepsilon/2).$$

Soit $a > 0$ tel que $\mathbb{P}\,(|X| \geqslant a) \leqslant \delta/4$. Comme $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} Y$, il existe $n_0 \geqslant 1$ tel que, pour $n \geqslant n_0$, on ait $\mathbb{P}\,(|Y_n - Y| \geqslant \varepsilon/2a) \leqslant \delta/4$. Alors, pour $n \geqslant n_0$:

$$\begin{aligned}\mathbb{P}\,(|X|\,|Y_n - Y| \geqslant \varepsilon/2) &\leqslant \mathbb{P}\,(|X| \geqslant a) + \mathbb{P}\,(|X| \leqslant a \text{ et } |X|\,|Y_n - Y| \geqslant \varepsilon/2)\\ &\leqslant \mathbb{P}\,(|X| \geqslant a) + \mathbb{P}\,(|Y_n - Y| \geqslant \varepsilon/2a) \leqslant \delta/2.\end{aligned}$$

Choisissons maintenant $b > 0$ tel que $\mathbb{P}\,(|Y| \geqslant b) \leqslant \delta/8$. On a :

$$\begin{aligned}\mathbb{P}\,(|Y_n|\,|X_n - X| \geqslant \varepsilon/2) \leqslant{}& \mathbb{P}\,(|Y_n - Y| \geqslant \varepsilon/2a) + \mathbb{P}\,(|Y| \geqslant b)\\ &+ \mathbb{P}\,(|Y_n - Y| \leqslant \varepsilon/2a\,; |Y| \leqslant b \text{ et } |Y_n|\,|X_n - X| \geqslant \varepsilon/2);\end{aligned}$$

et comme $|Y_n - Y| \leqslant \varepsilon/2a$ et $|Y| \leqslant b$ entraînent $|Y_n| \leqslant b + \varepsilon/2a$, on obtient :

$$\begin{aligned}\mathbb{P}\,(|Y_n|\,|X_n - X| \geqslant \varepsilon/2) \leqslant{}& \mathbb{P}\,(|Y_n - Y| \geqslant \varepsilon/2a) + \mathbb{P}\,(|Y| \geqslant b)\\ &+ \mathbb{P}\,(|X_n - X| \geqslant \varepsilon/[2b + \varepsilon/a]).\end{aligned}$$

Alors, comme $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$, il existe $n_1 \geqslant 1$, que l'on peut choisir $\geqslant n_0$, tel que, pour $n \geqslant n_1$, on ait $\mathbb{P}\,(|X_n - X| \geqslant \varepsilon/[2b + \varepsilon/a]) \leqslant \delta/8$. Comme $n_1 \geqslant n_0$, on a donc, pour $n \geqslant n_1$:

$$\mathbb{P}\,(|Y_n|\,|X_n - X| \geqslant \varepsilon/2) \leqslant \delta/4 + \delta/8 + \delta/8 = \delta/2\,,$$

et donc $\mathbb{P}\,(|X_nY_n - XY| \geqslant \varepsilon) \leqslant \delta$. □

Remarque. Soit $L^0(\Omega, \mathbb{P})$ l'espace de toutes les (classes de) *v.a.r.* $X : \Omega \to \mathbb{R}$. La convergence en probabilité peut être définie, comme on pourra le vérifier en exercice, par l'une des deux distances suivantes :

$$\begin{aligned}&d_1(X, Y) = \mathbb{E}\left(\frac{|X - Y|}{1 + |X - Y|}\right)\\ &d_2(X, Y) = \inf\{\varepsilon > 0\,;\ \mathbb{P}\,(|X - Y| \geqslant \varepsilon) \leqslant \varepsilon\}\,.\end{aligned}$$

Pour ces distances, $L^0(\mathbb{P})$ est un espace vectoriel topologique, non localement convexe, et il est complet pour ces distances.

Voyons quels liens il y a avec les convergences précédentes.

Proposition 2.4 *La convergence en moyenne d'ordre r entraîne la convergence en probabilité : si* $X_n \xrightarrow[n\to\infty]{L^r} X$, *alors* $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$.

Preuve. Il suffirait de le faire pour $r = 1$, puisque la convergence dans L^r entraîne la convergence en moyenne ; mais cela résulte immédiatement de l'inégalité **très utile** suivante :

Lemme 2.5 (inégalité de Markov) *Pour toute* $Y \in L^r(\mathbb{P})$*, on a, pour tout* $\varepsilon > 0$ *:*

$$\boxed{\mathbb{P}\left(|Y| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^r}\, \mathbb{E}\left(|Y|^r\right)}\,.$$

mais dont la preuve est immédiate :

$$\mathbb{E}\left(|Y|^r\right) \geqslant \int_{\{|Y| \geqslant \varepsilon\}} |Y|^r \, d\mathbb{P} \geqslant \varepsilon^r \, \mathbb{P}\left(|Y| \geqslant \varepsilon\right). \qquad \square$$

Cette inégalité est souvent appelée **inégalité de Bienaymé-Tchebychev**, qui en est un cas particulier, pour $r = 2$, lorsque l'on centre la *v.a.r.* :

$$\mathbb{P}\left(|Y - \mathbb{E}\,(Y)| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} \operatorname{Var}(Y)\,.$$

Proposition 2.6

1) *La convergence presque sûre entraîne la convergence en probabilité : si* $X_n \xrightarrow[n\to\infty]{p.s.} X$*, alors* $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$*.*

2) *Si* $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$*, on peut extraire une* sous-suite *qui converge presque sûrement vers* X*.*

On retrouve le fait que s'il y a convergence en moyenne d'ordre r, alors il y a une sous-suite qui converge presque sûrement.

Cette proposition peut se prouver directement, mais il est plus intéressant de passer par le lemme suivant, qui est **fondamental** pour la convergence presque sûre : il exprime la convergence presque sûre comme une convergence *uniforme* en probabilité. Il explique aussi pourquoi, lorsque l'on veut montrer des convergences presque sûres, on a besoin de montrer ce que l'on appelle des *inégalités maximales*, c'est-à-dire des inégalités portant sur la borne supérieure de familles de fonctions, ou plutôt de leur valeurs absolues.

Lemme 2.7 (Lemme fondamental pour la convergence presque sûre)

Une suite de v.a.r. $(X_n)_{n \geqslant 1}$ *converge presque sûrement vers* 0 *si et seulement si :*

$$\boxed{Y_n = \sup_{k \geqslant n} |X_k| \xrightarrow[n\to\infty]{\mathbb{P}} 0}\,.$$

Preuve. Remarquons tout d'abord que $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$ si et seulement si, pour tout $\varepsilon > 0$, on a $\mathbb{P}\left(|Y_n| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$ (on a mis $> \varepsilon$ au lieu de $\geqslant \varepsilon$).

Pour tout $\varepsilon > 0$ et tout $n \geqslant 1$, posons :

$$E_n(\varepsilon) = \{|X_n| > \varepsilon\}$$

et :

$$E(\varepsilon) = \varlimsup_{n\to\infty} E_n(\varepsilon) = \bigcap_{n \geqslant 1} \bigcup_{k \geqslant n} E_k(\varepsilon)\,.$$

Remarquons maintenant que $X_n(\omega) \xrightarrow[n\to\infty]{} 0$ si et seulement si :

$$(\forall \varepsilon > 0) \quad (\exists n \geqslant 1) \quad (\forall k \geqslant n) \qquad\qquad |X_k(\omega)| \leqslant \varepsilon\,;$$

par conséquent :

$$\{\omega \in \Omega\,;\ X_n(\omega) \ \ \not\xrightarrow[n\to\infty]{} 0\} = \bigcup_{\varepsilon>0} E(\varepsilon) = \bigcup_{j\geqslant 1} E(1/j)\,,$$

de sorte que :

$$X_n \xrightarrow[n\to\infty]{p.s.} 0 \qquad \Longleftrightarrow \qquad \mathbb{P}\big(E(\varepsilon)\big) = 0\,, \quad \forall \varepsilon > 0\,.$$

Mais $\mathbb{P}\big(E(\varepsilon)\big) = \lim_{n\to\infty} \mathbb{P}\big(\bigcup_{k\geqslant n} E_k(\varepsilon)\big)$, et cela donne le résultat puisque :

$$\bigcup_{k\geqslant n} E_k(\varepsilon) = \Big\{\sup_{k\geqslant n} |X_k| > \varepsilon\Big\}\,. \qquad \square$$

Preuve de la Proposition 2.6. Le 1) résulte immédiatement du lemme.

Pour le 2), on utilisera un autre lemme, qui donne un **critère pratique de convergence presque sûre**.

Lemme 2.8 *Soit $(X_k)_{k\geqslant 1}$ une suite de v.a.r. . Si l'on a :*

$$\sum_{k=1}^{\infty} \mathbb{P}\left(|X_k| > \varepsilon\right) < +\infty$$

pour tout $\varepsilon > 0$, alors $X_k \xrightarrow[k\to\infty]{p.s.} 0$.

Preuve. Il suffit de remarquer que la condition entraîne :

$$\mathbb{P}\big(E(\varepsilon)\big) = \mathbb{P}\Big(\overline{\lim_{k\to\infty}}\ E_k(\varepsilon)\Big) = 0,$$

par le Lemme de Borel-Cantelli. $\square$

On peut alors terminer la preuve de la Proposition 2.6 : si $X_n \xrightarrow[k\to\infty]{\mathbb{P}} X$, on peut trouver une suite strictement croissante d'entiers n_k, $k \geqslant 1$, tels que :

$$\mathbb{P}\left(|X_{n_k} - X| > 1/2^k\right) \leqslant \frac{1}{2^k}\,.$$

Alors, pour tout $\varepsilon > 0$, on a, lorsque $1/2^k \leqslant \varepsilon$:

$$\mathbb{P}\left(|X_{n_k} - X| > \varepsilon\right) \leqslant \mathbb{P}\left(|X_{n_k} - X| > 1/2^k\right) \leqslant \frac{1}{2^k}\,,$$

de sorte que $\sum_{k=1}^{\infty} \mathbb{P}\left(|X_{n_k} - X| > \varepsilon\right) < +\infty$. $\square$

Bien sûr, la condition : $\sum_{k=1}^{\infty} \mathbb{P}\left(|X_{n_k} - X| > \varepsilon\right) < +\infty$, $\forall \varepsilon > 0$, est équivalente à $\sum_{k=1}^{\infty} \mathbb{P}\left(|X_{n_k} - X| \geqslant \varepsilon\right) < +\infty$, $\forall \varepsilon > 0$. Elle est vérifiée dès qu'il existe un $r \geqslant 1$ tel que :

$$\sum_{k=1}^{\infty} \|X_k\|_{L^r(\mathbb{P})}^r = \sum_{k=1}^{\infty} \mathbb{E}\left(|X_k|^r\right) < +\infty\,,$$

d'après l'inégalité de Markov. Mais, dans ce cas, la conclusion $X_k \xrightarrow[k\to\infty]{p.s.} 0$ résulte du Théorème de convergence monotone :

$$\int_\Omega \Big(\sum_{k=1}^{\infty} |X_k|^r\Big)\,d\mathbb{P} = \sum_{k=1}^{\infty}\bigg(\int_\Omega |X_k|^r\,d\mathbb{P}\bigg) = \sum_{k=1}^{\infty} \mathbb{E}\left(|X_k|^r\right) < +\infty\,,$$

qui entraîne $\sum_{k=1}^{\infty} |X_k|^r < +\infty$ presque sûrement, et donc, en particulier, $X_k \xrightarrow[k\to\infty]{p.s.} 0$.

3 Convergence des séries de variables aléatoires réelles indépendantes

3.1 Loi du 0–1 de Kolmogorov

Soit $\left(\mathscr{B}_n\right)_{n\geqslant 1}$ une suite de sous-tribus de $\mathscr{A}$.

On note :

$$\sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \ldots)$$

la tribu engendrée par $\mathscr{B}_n, \mathscr{B}_{n+1}, \ldots$.

Définition 3.1 *La tribu :*

$$\mathscr{B}_\infty = \bigcap_{n\geqslant 1} \sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \ldots)$$

est appelée **tribu asymptotique** *de la suite* $\left(\mathscr{B}_n\right)_{n\geqslant 1}$.

Intuitivement, les événements de $\mathscr{B}_\infty$ décrivent se qui se passe à l'infini, relativement à la suite $\left(\mathscr{B}_n\right)_{n\geqslant 1}$.

L'exemple de base est donné par la :

Proposition 3.2 *Si $A_n \in \mathscr{B}_n$ pour tout $n \geqslant 1$, alors :*

$$\varliminf_{n\to\infty} A_n \in \mathscr{B}_\infty \quad \textit{et} \quad \varlimsup_{n\to\infty} A_n \in \mathscr{B}_\infty\,.$$

Remarque. Etant donnés $A_n \in \mathscr{A}$, $n \geqslant 1$, on peut toujours choisir $\mathscr{B}_n = \{\emptyset, A_n, A_n^c, \Omega\}$; les événements $\underline{\lim}_{n\to\infty} A_n$ et $\overline{\lim}_{n\to\infty} A_n$ sont donc toujours asymptotiques (pour ce choix de tribus).

Preuve. Comme $\overline{\lim}_{n\to\infty} A_n = \Big(\underline{\lim}_{n\to\infty} A_n^c\Big)^c$, il suffit de montrer que $\underline{\lim}_{n\to\infty} A_n \in \mathscr{B}_\infty$.

Or, pour tout $n \geqslant m$, on a :

$$\bigcap_{k\geqslant n} A_k \in \sigma(\mathscr{B}_m, \mathscr{B}_{m+1}, \dots)\,.$$

Comme la suite $\big(\bigcap_{k\geqslant n} A_k\big)_{n\geqslant 1}$ est *croissante*, on a, pour tout $m \geqslant 1$:

$$\underline{\lim}_{n\to\infty} A_n = \bigcup_{n\geqslant 1}\Big(\bigcap_{k\geqslant n} A_k\Big) = \bigcup_{n\geqslant m}\Big(\bigcap_{k\geqslant n} A_k\Big) \in \sigma(\mathscr{B}_m, \mathscr{B}_{m+1}, \dots)\,,$$

ce qui donne bien le résultat annoncé. □

Proposition 3.3 *Si, pour tout $n \geqslant 1$, $X_n \colon \Omega \to \overline{\mathbb{R}}$ est une v.a.r. $\mathscr{B}_n$-mesurable, alors $\underline{\lim}_{n\to\infty} X_n$ et $\overline{\lim}_{n\to\infty} X_n$ sont des v.a.r. $\mathscr{B}_\infty$-mesurables.*

Preuve. Il suffit de faire la preuve pour $X = \underline{\lim}_{n\to\infty} X_n$. La *v.a.r.* $Y_n = \inf_{k\geqslant n} X_k$ est $\sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \dots)$-mesurable ; donc $\sup_{n\geqslant m} Y_n$ est $\sigma(\mathscr{B}_m, \mathscr{B}_{m+1}, \dots)$-mesurable. Mais, pour tout $m \geqslant 1$, $X = \sup_{n\geqslant m} Y_n$; donc X est $\sigma(\mathscr{B}_m, \mathscr{B}_{m+1}, \dots)$-mesurable pour tout $m \geqslant 1$, et donc $\mathscr{B}_\infty$-mesurable. □

Corollaire 3.4 *L'ensemble de convergence :*

$$\{\omega \in \Omega\,;\ \lim_{n\to+\infty} X_n(\omega) \quad \textit{existe}\}$$

est dans la tribu asymptotique $\mathscr{B}_\infty$.

Remarque. Les *v.a.r.* X_n étant données, on peut toujours prendre $\mathscr{B}_n = \sigma(X_n)$, ou $\mathscr{B}_n = \sigma(X_n, X_{n+1}, \dots)$; dans ces deux cas, on a $\sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \dots) = \sigma(X_n, X_{n+1}, \dots)$, et l' on a donc la même tribu asymptotique.

L'intérêt de la tribu asymptotique est :

Théorème 3.5 (loi du 0–1 de Kolmogorov)

Si les tribus $\mathscr{B}_n$, pour $n \geqslant 1$, sont **<u>indépendantes</u>**, *alors tout* $\boxed{B \in \mathscr{B}_\infty}$ *est indépendant de lui-même, et donc* $\boxed{\mathbb{P}(B) = 0 \text{ ou } \mathbb{P}(B) = 1}$.

Preuve. Pour tout $m < n$, les deux tribus :

$$\sigma(\mathscr{B}_1, \ldots, \mathscr{B}_m) \quad \text{et} \quad \sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \ldots)$$

sont indépendantes ; donc $\sigma(\mathscr{B}_1, \ldots, \mathscr{B}_m)$ et $\mathscr{B}_\infty$ sont indépendantes pour tout $m \geqslant 1$.

Notons :

$$\mathscr{C} = \bigcup_{m \geqslant 1} \sigma(\mathscr{B}_1, \ldots, \mathscr{B}_m).$$

$\mathscr{C}$ est stable par intersection finie (car la suite $\big(\sigma(\mathscr{B}_1, \ldots, \mathscr{B}_m)\big)_{m \geqslant 1}$ est croissante), contient Ω, et tout élément de $\mathscr{C}$ est indépendant de tout élément de $\mathscr{B}_\infty$. Par le critère d'indépendance vu au Chapitre 1 (Proposition 4.9), les tribus :

$$\sigma(\mathscr{C}) = \sigma(\mathscr{B}_1, \mathscr{B}_2, \ldots) \quad \text{et} \quad \mathscr{B}_\infty$$

sont indépendantes.

Comme $\mathscr{B}_\infty \subseteq \sigma(\mathscr{B}_1, \mathscr{B}_2, \ldots)$, $\mathscr{B}_\infty$ est indépendante d'elle-même, et donc, pour tout $B \in \mathscr{B}_\infty$, on a :

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

donc $\mathbb{P}(B) = 0$ ou $\mathbb{P}(B) = 1$. □

Corollaire 3.6 *Si les v.a.r. X_n, $n \geqslant 1$, sont* **indépendantes***, alors les v.a.r. $\varliminf_{n\to\infty} X_n$ et $\varlimsup_{n\to\infty} X_n$ sont p.s. constantes.*

Preuve. Prenons $\mathscr{B}_n = \sigma(X_n)$. On sait que $X = \varliminf_{n\to\infty} X_n$ est $\mathscr{B}_\infty$-mesurable.

Si $\mathbb{P}(X = -\infty) = 1$ (respectivement $\mathbb{P}(X = +\infty) = 1$), alors $X \overset{p.s.}{=} -\infty$ (respectivement $+\infty$).

Sinon, comme les tribus $\mathscr{B}_1, \mathscr{B}_2, \ldots$ sont indépendantes, on a, par la loi du 0–1 : $\mathbb{P}(X = -\infty) = \mathbb{P}(X = +\infty) = 0$. X est donc *p.s.* à valeurs dans $\mathbb{R}$. De plus, pour tout $x \in \mathbb{R}$:

$$F_X(x) = \mathbb{P}(X \leqslant x) = 0 \text{ ou } 1 ;$$

donc si l'on pose $a = \sup\{x \in \mathbb{R} ; F_X(x) = 0\}$, alors $a \in \mathbb{R}$, et $F_X(x) = 0$ pour $x < a$, $F_X(x) = 1$ pour $x > a$. Comme F_X est continue à droite, on a $F_X(a) = 1$. Donc $F_X = \mathbb{1}_{[a, +\infty[}$, de sorte que $X \overset{p.s.}{=} a$. □

Remarque. Nous avons vu que si les événements A_n, $n \geqslant 1$, sont indépendants, alors $\varlimsup_{n\to\infty} A_n$ est asymptotique, et donc de probabilité 0 ou 1 par la loi du 0–1. On peut en fait voir cela directement, et même le préciser.

Proposition 3.7 (Lemme de Borel-Cantelli) *Si les événements A_n, $n \geqslant 1$, sont* **indépendants**, *on a :*

a) $\quad \mathbb{P}(\overline{\lim_{n\to\infty}} A_n) = 0 \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty$ *;*

b) $\quad \mathbb{P}(\overline{\lim_{n\to\infty}} A_n) = 1 \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty.$

Preuve. Nous avons déjà vu que, même sans indépendance :

$$\sum_{n=1}^{+\infty} \mathbb{P}(A_n) < +\infty \quad \Longrightarrow \quad \mathbb{P}(\overline{\lim}_{n\to\infty} A_n) = 0\,.$$

Mais, d'autre part, on a aussi :

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty \quad \Longrightarrow \quad \mathbb{P}(\overline{\lim}_{n\to\infty} A_n) = 1\,;$$

en effet, comme les A_n^c, $n \geqslant 1$, sont indépendants, on a :

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{k=n}^{N} A_k^c\Big) &= \prod_{k=n}^{N} \mathbb{P}(A_k^c) = \prod_{k=n}^{N} \big(1 - \mathbb{P}(A_k)\big) \\ &\leqslant \prod_{k=n}^{N} \exp\big(-\mathbb{P}(A_k)\big) = \exp\Big(-\sum_{k=n}^{N} \mathbb{P}(A_k)\Big) \xrightarrow[N\to+\infty]{} 0\,;\end{aligned}$$

donc $\mathbb{P}\big(\bigcap_{k\geqslant n} A_k^c\big) = 0$, et donc :

$$\mathbb{P}\Big(\underline{\lim}_{n\to\infty} A_n^c\Big) = \mathbb{P}\Big(\bigcap_{k\geqslant n} A_k^c\Big) = 0\,. \qquad \square$$

Exemple. Si l'on joue à "pile ou face", et que A_n est l'événement "obtenir "face" au $n^{ème}$ lancer", comme $\mathbb{P}(A_n) = 1/2$ pour tout $n \geqslant 1$, on a $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty$, donc $\mathbb{P}(\overline{\lim_{n\to\infty}} A_n) = 1$: presque sûrement, on aura une infinité de fois "face". De même, on aura une infinité de fois "pile".

Remarque. L'indépendance est essentielle. Par exemple, si $\Omega = [0,1]$ et $A_n = [0, 1/n]$, on a $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty$, mais $\overline{\lim_{n\to\infty}} A_n = \{0\}$.

3.2 Convergence des séries de variables aléatoires réelles indépendantes

Nous supposerons ici que les *v.a.r.* X_n prennent leurs valeurs dans $\mathbb{R}$, et pas dans $\overline{\mathbb{R}}$.

Commençons par un corollaire de la loi du 0–1.

Proposition 3.8 *Soit* $(X_n)_{n\geqslant 1}$ *une suite de v.a.r.* indépendantes. *Alors :*

$$C = \left\{\omega \in \Omega\,;\ \sum\nolimits_{n\geqslant 1} X_n(\omega) \quad \textit{converge}\right\}$$

est de probabilité 0 ou 1.

Preuve. Posons $S_n = X_1 + \cdots + X_n$.

Il faut faire attention que les *v.a.r.* $S_1, S_2, \ldots$ ne sont <u>pas</u> indépendantes.

Néanmoins, si l'on prend $\mathscr{B}_n = \sigma(X_n)$, les tribus $\mathscr{B}_1, \mathscr{B}_2, \ldots$ sont indépendantes, et, pour tout $n \geqslant 1$:

$$C = \Big\{\omega \in \Omega\,;\ \sum_{k\geqslant n} X_k(\omega) \quad \text{converge}\Big\} \in \sigma(\mathscr{B}_n, \mathscr{B}_{n+1}, \ldots)\,;$$

donc $C \in \mathscr{B}_\infty$. □

Un premier critère de convergence pour la série est donné par le :

Théorème 3.9 (Théorème de Kolmogorov) *Si les v.a.r.* X_n, $n \geqslant 1$, *sont dans* $L^2(\mathbb{P})$, *sont* <u>**indépendantes**</u> *et* <u>centrées</u> *:* $\mathbb{E}\,(X_n) = 0$, *alors :*

$$\sum_{n=1}^{\infty} \|X_n\|_2^2 = \sum_{n=1}^{\infty} \mathbb{E}\,(X_n^2) < +\infty \qquad \Longrightarrow \qquad \sum_{n=1}^{\infty} X_n \quad \textit{converge p.s.}.$$

Exemple. Soit $(\varepsilon_n)_{n\geqslant 1}$ une suite indépendante de *v.a.r.* telles que $\mathbb{P}\,(\varepsilon_n = -1) = \mathbb{P}\,(\varepsilon_n = 1) = 1/2$ (*choix de signes aléatoire*), alors la série $\sum_{n\geqslant 1} \frac{\varepsilon_n}{n^\alpha}$ converge *p.s.* dès que $\alpha > 1/2$ (alors qu'elle ne converge pour *tout* choix de signes que si $\alpha > 1$). Si $\varepsilon_n(\omega_0) = (-1)^n$, pour tout $n \geqslant 1$, alors $\sum_{n\geqslant 1} \frac{\varepsilon_n(\omega_0)}{n^\alpha}$ est alternée et converge pour tout $\alpha > 0$.

Si l'on ne suppose pas les X_n centrées, les $X_n - \mathbb{E}\,(X_n)$ le sont de toute façon, et le théorème devient :

Corollaire 3.10 *Si les* X_n, $n \geqslant 1$, *sont dans* $L^2(\mathbb{P})$ *et sont* **indépendantes**, *alors :*

$$\sum_{n=1}^{\infty} \mathrm{Var}\,(X_n) < +\infty \qquad \Longrightarrow \qquad \sum_{n=1}^{\infty} \big(X_n - \mathbb{E}\,(X_n)\big) \quad \textit{converge p.s.}.$$

Remarque. Les X_n étant indépendantes et centrées, elles sont orthogonales dans $L^2(\mathbb{P})$. Par conséquent, la condition $\sum_{n=1}^{\infty} \|X_n\|_2^2 = \sum_{n=1}^{\infty} \mathbb{E}\,(X_n^2) < +\infty$ équivaut à la convergence de la série $\sum_{n=1}^{\infty} X_n$ *pour la norme* dans l'espace $L^2(\mathbb{P})$ (c'est-à-dire en moyenne quadratique). Il en résulte qu'il y a convergence de la série *en probabilité*. Cela entraîne qu'une *sous-suite* de la suite des sommes partielles converge presque sûrement. Le Théorème de Kolmogorov donne la convergence presque sûre de la suite des sommes partielles elle-même.

En fait, nous verrons un théorème, dû à Paul Lévy, disant que pour des *v.a.r. indépendantes* X_n, la convergence en probabilité de la **série** $\sum_{n\geqslant 1} X_n$ équivaut à sa convergence presque sûre. Mais il est préférable de connaître la remarquable inégalité de Kolmogorov, ci-dessous, qui donne directement le théorème. On peut donner une preuve directe du théorème de Paul Lévy utilisant les mêmes idées (voir D. Revuz, p. 138, ou L. Breiman, p. 45–46), mais on suivra un chemin légèrement différent.

Théorème 3.11 (inégalité de Kolmogorov) *Si les v.a.r.* X_n, $n \geqslant 1$, *sont dans* $L^2(\mathbb{P})$ *et sont* **indépendantes** *et* **centrées**, *alors, pour tout* $a > 0$, *on a :*

$$\boxed{\mathbb{P}\Big(\sup_{n\geqslant 1}\Big|\sum_{k=1}^{n} X_k\Big| > a\Big) \leqslant \frac{1}{a^2}\sum_{k=1}^{\infty}\mathbb{E}\,(X_k^2)}\,.$$

C'est ce que l'on appelle une "*inégalité maximale*" ; il est naturel, comme on l'a dit, de devoir montrer une telle inégalité, puisque l'on cherche à prouver une convergence presque sûre. Cette inégalité de Kolmogorov généralise l'inégalité de Bienaymé-Tchebychev (Markov) qui correspond au cas d'une seule *v.a.r.* (c'est-à-dire si $X_2 = X_3 = \cdots = 0$). Contrairement à celle de Bienaymé-Tchebychev, sa preuve n'est pas évidente.

On peut aussi remarquer que l'orthogonalité des X_n permet d'écrire que $\mathbb{E}\Big(\sum_{k=1}^{n} X_k^2\Big) = \sum_{k=1}^{n}\mathbb{E}\,(X_k^2)$, et donc l'inégalité de Bienaymé-Tchebychev s'écrit :

$$\mathbb{P}\Big(\Big|\sum_{k=1}^{n} X_k\Big| \geqslant a\Big) \leqslant \frac{1}{a^2}\sum_{k=1}^{n}\mathbb{E}\,(X_k^2) \leqslant \frac{1}{a^2}\sum_{k=1}^{\infty}\mathbb{E}\,(X_k^2)\,;$$

on voit donc le renforcement apporté par l'inégalité de Kolmogorov ; c'est ce qui permet de passer de la convergence en probabilité à la convergence presque sûre.

Preuve du Théorème à partir de l'inégalité. On applique l'inégalité pour la suite $X_{m+1}, X_{m+2}, \ldots$:

$$\mathbb{P}\Big(\sup_{n>m}\Big|\sum_{k=m+1}^{n} X_k\Big| > a\Big) \leqslant \frac{1}{a^2}\sum_{k=m+1}^{\infty}\mathbb{E}\,(X_k^2)\,.$$

Posons $S_n = \sum_{k=1}^{n} X_k$.

Comme :

$$S_{n_2} - S_{n_1} \leqslant |S_{n_2} - S_m| + |S_{n_1} - S_m|\,,$$

on a :

$$\sup_{n>m} S_n - \inf_{n>m} S_n \leqslant 2 \sup_{n>m} |S_n - S_m| \,,$$

donc :

$$\mathbb{P}\big(\sup_{n>m} S_n - \inf_{n>m} S_n > 2a\big) \leqslant \mathbb{P}\big(\sup_{n>m} |S_n - S_m| > a\big) \leqslant \frac{1}{a^2} \sum_{k=m+1}^{\infty} \mathbb{E}\,(X_k^2) \xrightarrow[m\to\infty]{} 0\,.$$

Comme

$$\varlimsup_{n\to\infty} S_n - \varliminf_{n\to\infty} S_n = \lim_{m\to\infty} \big[\sup_{n>m} S_n - \inf_{n>m} S_n\big] = \inf_{m\geqslant 1} \big[\sup_{n>m} S_n - \inf_{n>m} S_n\big],$$

on a, pour tout $m \geqslant 1$:

$$\big\{\sup_{n>m} S_n - \inf_{n>m} S_n > 2a\big\} \supseteq \big\{\varlimsup_{n\to\infty} S_n - \varliminf_{n\to\infty} S_n > 2a\big\};$$

on obtient donc :

$$\mathbb{P}\big(\varlimsup_{n\to\infty} S_n - \varliminf_{n\to\infty} S_n > 2a\big) = 0\,.$$

Comme c'est vrai pour tout $a > 0$, cela signifie que :

$$\varlimsup_{n\to\infty} S_n = \varliminf_{n\to\infty} S_n \quad p.s.\,.$$

La suite $(S_n)_{n\geqslant 1}$ a donc presque sûrement une limite. Cette limite est presque sûrement finie car on a vu, juste avant la preuve, que, grâce à l'orthogonalité des X_n dans $L^2(\mathbb{P})$, la série $\sum_{n\geqslant 1} X_n$ converge en probabilité vers une *v.a.r.* $X \in L^2(\mathbb{P})$. Cette *v.a.r.* X est finie *p.s.* et est limite en probabilité de la suite $(S_n)_{n\geqslant 1}$. Elle coïncide donc avec sa limite presque sûre, qui est donc bien finie presque sûrement. □

Preuve de l'inégalité de Kolmogorov. Introduisons la *v.a.r.* :

$$\nu_a \colon \Omega \to \{1, 2, \ldots, +\infty\}$$

définie par :

$$\boxed{\nu_a(\omega) = \begin{cases} +\infty \quad \text{si} \quad \Big|\displaystyle\sum_{k=1}^{n} X_k(\omega)\Big| \leqslant a\,,\ \forall n \geqslant 1\,, \\ \min\Big\{n \geqslant 1\,;\ \Big|\displaystyle\sum_{k=1}^{n} X_k(\omega)\Big| > a\Big\} \quad \text{sinon.} \end{cases}}\,.$$

On a :

$$\boxed{\nu(\omega) < +\infty \qquad \Longrightarrow \qquad \Big|\sum_{k=1}^{\nu_a(\omega)} X_k(\omega)\Big| > a}\,.$$

$\nu_a(\omega)$ est le *premier instant* (c'est-à-dire le premier indice) où la somme va dépasser a.

On a :

$$\{\nu_a \leqslant n\} = \bigcup_{m \leqslant n} \left\{ \left| \sum_{k=1}^{m} X_k(\omega) \right| > a \right\}; \tag{3.1}$$

donc :

$$\begin{aligned}\{\nu_a < +\infty\} = \bigcup_{n \geqslant 1} \{\nu_a \leqslant n\} &= \bigcup_{m \geqslant 1} \left\{ \left| \sum_{k=1}^{m} X_k(\omega) \right| > a \right\} \\ &= \left\{ \sup_{m \geqslant 1} \left| \sum_{k=1}^{m} X_k(\omega) \right| > a \right\},\end{aligned}$$

et il s'agit par conséquent de montrer que :

$$\mathbb{P}\left(\nu_a < +\infty\right) \leqslant \frac{1}{a^2} \sum_{n=1}^{\infty} \mathbb{E}\left(X_n^2\right).$$

Introduisons alors les *v.a.r.* Y_n définies par :

$$Y_n(\omega) = X_n(\omega) \mathbb{1}_{\{\nu_a \geqslant n\}}(\omega) = \begin{cases} X_n(\omega) & \text{si} \quad n \leqslant \nu_a(\omega) \\ 0 & \text{si} \quad n > \nu_a(\omega). \end{cases}$$

On a :

$$\boxed{\sum_{n=1}^{\infty} Y_n(\omega) = \sum_{n=1}^{\nu_a(\omega)} X_n(\omega)},$$

de sorte que, pour calculer la somme des $Y_n(\omega)$, on calcule celle des $X_n(\omega)$, en s'*arrêtant* à l'indice $\nu_a(\omega)$. On dit que la *v.a.r.* ν_a est un **temps d'arrêt**.

Maintenant :

- comme $|Y_n| \leqslant |X_n|$, on a $Y_n \in L^2(\mathbb{P})$;
- la suite $(Y_n)_{n \geqslant 1}$ est orthogonale dans $L^2(\mathbb{P})$. En effet, commençons par remarquer que (3.1) entraîne :

$$\{\nu_a \leqslant n\} \in \sigma(X_1, \ldots, X_n), \quad \forall n \geqslant 1;$$

donc :

$$\{\nu_a \geqslant m\} = \{\nu_a \leqslant m-1\}^c \in \sigma(X_1, \ldots, X_{m-1}),$$

et donc Y_m est $\sigma(X_1, \ldots, X_m)$-mesurable. Par conséquent, pour $m < n$, la *v.a.r.* $\mathbb{1}_{\{\nu_a \geqslant n\}} Y_m$ est $\sigma(X_1, \ldots, X_{n-1})$-mesurable, donc indépendante de X_n. Il en résulte que :

$$\mathbb{E}\left(Y_n Y_m\right) = \mathbb{E}\left(X_n . \mathbb{1}_{\{\nu_a \geqslant n\}} Y_m\right) = \mathbb{E}\left(X_n\right) \mathbb{E}\left(\mathbb{1}_{\{\nu_a \geqslant n\}} Y_m\right) = 0;$$

Comme :

$$\nu_a(\omega) < +\infty \quad \Longrightarrow \quad \left|\sum_{n=1}^{\nu_a(\omega)} X_n(\omega)\right| > a \quad \Longrightarrow \quad \left|\sum_{n=1}^{\infty} Y_n(\omega)\right| > a\,;$$

on obtient :

$$\begin{aligned}
\mathbb{P}\,(\nu_a < +\infty) &\leqslant \mathbb{P}\left(\left|\sum_{n=1}^{\infty} Y_n(\omega)\right| > a\right) \\
&\leqslant \frac{1}{a^2}\mathbb{E}\left(\left[\sum_{n=1}^{\infty} Y_n\right]^2\right) \qquad \text{(Bienaymé-Tchebychev)} \\
&= \frac{1}{a^2}\sum_{n=1}^{\infty} \|Y_n\|_{L^2(\mathbb{P})}^2 \qquad \text{(par orthogonalité)} \\
&\leqslant \frac{1}{a^2}\sum_{n=1}^{\infty} \|X_n\|_{L^2(\mathbb{P})}^2 = \frac{1}{a^2}\sum_{n=1}^{\infty} \mathbb{E}\,(X_n^2)\,. \qquad \square
\end{aligned}$$

Le Théorème de Kolmogorov a une réciproque partielle.

Théorème 3.12 *Si les v.a.r. X_n, $n \geqslant 1$, sont* indépendantes, *et* centrées, **et si** $\boxed{|X_n| \leqslant M,\, \forall n \geqslant 1}$, *alors la convergence p.s. de la série* $\sum_{n\geqslant 1} X_n$ *entraîne* $\sum_{n=1}^{+\infty} \mathbb{E}\,(X_n^2) < +\infty\,.$

Remarquons que, puisque $|X_n| \leqslant M$, on a $X_n \in L^\infty(\mathbb{P})$, donc $X_n \in L^2(\mathbb{P})$.

Nous allons donner deux preuves de ce résultat. La première reprend les idées de la preuve du Théorème de Kolmogorov ; la seconde, plus élémentaire, utilise une inégalité très utile, bien que très simple, due à Paley et Zygmund.

Preuve 1. On utilisera de nouveau le temps d'arrêt ν_a.

1) Fixons $n \geqslant 1$.

a) Par définition, sur l'ensemble $\{\nu_a = n\}$, on a $\left|\sum_{k=1}^{n-1} X_k\right| \leqslant a\,.$

Comme $|X_n| \leqslant M$ par hypothèse, on a :

$$\left|\sum_{k=1}^{n} X_k\right| \leqslant a + M \quad \text{sur} \quad \{\nu_a = n\}\,.$$

b) Comme $\{\nu_a = n\} \in \sigma(X_1, \ldots, X_n)$, on a, par indépendance, pour $k > n$:

$$\int_{\{\nu_a = n\}} X_k^2\, d\mathbb{P} = \mathbb{E}\left(\mathbb{1}_{\{\nu_a = n\}} X_k^2\right) = \mathbb{E}\left(\mathbb{1}_{\{\nu_a = n\}}\right)\mathbb{E}\,(X_k^2) = \mathbb{P}\,(\nu_a = n)\,\mathbb{E}\,(X_k^2)\,.$$

c) De même, puisque $\mathbb{E}\,(X_k) = 0$, on a, pour $k > n$ et $k \neq k'$:

$$\int_{\{\nu_a = n\}} X_{k'} X_k \, d\mathbb{P} \; = \mathbb{E}\,(\mathbb{1}_{\{\nu_a = n\}} X_{k'})\, \mathbb{E}\,(X_k) = 0\,.$$

d) Donc, pour $N \geqslant n$:

$$\int_{\{\nu_a = n\}} \Big(\sum_{k=1}^{N} X_k\Big)^2 d\mathbb{P} \; = \int_{\{\nu_a = n\}} \Big(\sum_{k=1}^{n} X_k\Big)^2 d\mathbb{P} \; + \sum_{k=n+1}^{N} \int_{\{\nu_a = n\}} X_k^2 \, d\mathbb{P}\;.$$

On obtient ainsi :

$$\begin{aligned}\int_{\{\nu_a = n\}} \Big(\sum_{k=1}^{N} X_k\Big)^2 d\mathbb{P} \; &\leqslant (a+M)^2 \mathbb{P}\,(\nu_a = n) + \mathbb{P}\,(\nu_a = n) \sum_{k=n+1}^{N} \mathbb{E}\,(X_k^2) \\ &\leqslant \mathbb{P}\,(\nu_a = n) \left[(a+M)^2 + \sum_{k=n+1}^{N} \mathbb{E}\,(X_k^2) \right].\end{aligned}$$

2) En sommant pour $1 \leqslant n \leqslant N$, cela donne :

$$\int_{\{\nu_a \leqslant N\}} \left(\sum_{k=1}^{N} X_k\right)^2 d\mathbb{P} \; \leqslant \mathbb{P}\,(\nu_a \leqslant N) \left[(a+M)^2 + \sum_{k=1}^{N} \mathbb{E}\,(X_k^2) \right]. \qquad (3.2)$$

3) Notons maintenant que :

$$\nu_a > N \qquad \Longrightarrow \qquad \Big| \sum_{k=1}^{N} X_k \Big| \leqslant a\,;$$

donc :

$$\int_{\{\nu_a > N\}} \Big(\sum_{k=1}^{N} X_k\Big)^2 d\mathbb{P} \; \leqslant a^2\, \mathbb{P}\,(\nu_a > N) \leqslant (a+M)^2\, \mathbb{P}\,(\nu_a > N)\,. \qquad (3.3)$$

4) En additionnant (3.2) et (3.3), on obtient :

$$\mathbb{E}\left[\Big(\sum_{k=1}^{N} X_k\Big)^2 \right] \leqslant (a+M)^2 + \mathbb{P}\,(\nu_a \leqslant N) \sum_{k=1}^{N} \mathbb{E}\,(X_k^2)\,.$$

Mais, les *v.a.r.* X_n, $n \geqslant 1$, étant indépendantes et centrées, sont orthogonales ; donc $\mathbb{E}\left[\Big(\sum_{k=1}^{N} X_k\Big)^2\right] = \sum_{k=1}^{N} \mathbb{E}\,(X_k^2)$, et l'on obtient :

$$\mathbb{P}\,(\nu_a > N) \sum_{k=1}^{N} \mathbb{E}\,(X_k^2) \leqslant (a+M)^2\,.$$

Faisons tendre N vers $+\infty$; cela donne :

$$\mathbb{P}\,(\nu_a = +\infty) \sum_{k=1}^{\infty} \mathbb{E}\,(X_k^2) \leqslant (a+M)^2 \,.$$

5) Utilisons maintenant la convergence *p.s.* de la série $\sum_{n\geqslant 1} X_n$. Elle implique qu'il existe $a > 0$ tel que :

$$\mathbb{P}\left(\sup_{n\geqslant 1} \Big| \sum_{k=1}^{n} X_k \Big| \leqslant a \right) > 0 \,,$$

car sinon on aurait $\mathbb{P}\left(\sup_{n\geqslant 1} \Big| \sum_{k=1}^{n} X_k \Big| > m \right) = 1$, pour tout $m \geqslant 1$, et l'on aurait donc $\mathbb{P}\left(\sup_{n\geqslant 1} \Big| \sum_{k=1}^{n} X_k \Big| = +\infty \right) = 1$.

Mais cela signifie que pour cet $a > 0$:

$$\mathbb{P}\,(\nu_a = +\infty) > 0 \,,$$

ce qui termine la démonstration. □

Preuve 2. On commencera par montrer une inégalité simple, mais utile.

Proposition 3.13 (inégalité de Paley-Zygmund) *Soit X une v.a.r. positive, non nulle, et dans L^2. Alors, pour tout $\lambda \in]0,1[$, on a :*

$$\boxed{\mathbb{P}\big(X \geqslant \lambda\,\mathbb{E}\,(X)\big) \geqslant (1-\lambda)^2 \, \frac{[\mathbb{E}\,(X)]^2}{\mathbb{E}\,(X^2)}} \,.$$

Preuve. Soit A l'événement $\big(X \geqslant \lambda\,\mathbb{E}\,(X)\big)$. Utilisons l'inégalité de Cauchy-Schwarz pour écrire :

$$\begin{aligned} \mathbb{E}\,(X) = \mathbb{E}\,(X\mathbb{1}_{A^c}) + \mathbb{E}\,(X\mathbb{1}_A) &\leqslant \lambda\,\mathbb{E}\,(X) + [\mathbb{E}\,(X^2)]^{1/2}[\mathbb{E}\,(\mathbb{1}_A)]^{1/2} \\ &= \lambda\,\mathbb{E}\,(X) + \mathbb{P}\,(A)^{1/2}[\mathbb{E}\,(X^2)]^{1/2} \,; \end{aligned}$$

d'où $(1-\lambda)\,\mathbb{E}\,(X) \leqslant \mathbb{P}\,(A)^{1/2}[\mathbb{E}\,(X^2)]^{1/2}$, ce qui donne le résultat. □

Preuve 2 du Théorème 3.12. On utilisera seulement le fait que la *v.a.r.* :

$$S = \sup_{n\geqslant 1} |S_n| = \sup_{n\geqslant 1} |X_1 + \cdots + X_n|$$

est finie presque sûrement.

Sans perdre de généralité, on peut supposer que $M = 1$.

Comme $X_1, \ldots, X_N$ sont indépendantes et centrées, elles sont orthogonales dans $L^2(\mathbb{P})$, et donc $\sum\limits_{n=1}^{N} \mathbb{E}\,(X_n^2) = \mathbb{E}\,(S_N^2)$. Notons cette valeur par σ_N^2.

L'inégalité de Paley-Zygmund donne :

$$\mathbb{P}\,(S^2 \geqslant \sigma_N^2/2) \geqslant \mathbb{P}\,(S_N^2 \geqslant \sigma_N^2/2) \geqslant \frac{1}{4}\,\frac{\sigma_N^4}{\mathbb{E}\,(S_N^4)}\,.$$

Majorons $\mathbb{E}\,(S_N^4)$. On a :

$$S_N^4 = \sum \frac{4!}{\alpha_1! \ldots \alpha_N!} X_1^{\alpha_1} \cdots X_N^{\alpha_N}\,,$$

où la somme porte sur tous les N-uples d'entiers tels que $\alpha_1 + \cdots + \alpha_N = 4$. D'où, par indépendance :

$$\mathbb{E}\,(S_N^4) = \sum \frac{4!}{\alpha_1! \cdots \alpha_N!}\,\mathbb{E}\,(X_1^{\alpha_1}) \cdots \mathbb{E}\,(X_N^{\alpha_N})\,.$$

Mais un α_i égal à 1 donne une contribution nulle car les X_i sont centrées ; donc :

$$\mathbb{E}\,(S_N^4) = \sum \frac{4!}{(2\beta_1)! \cdots (2\beta_N)!}\,\mathbb{E}\,(X_1^{2\beta_1}) \cdots \mathbb{E}\,(X_N^{2\beta_N})\,,$$

où la somme porte cette fois sur tous les N-uples d'entiers tels que $\beta_1+\cdots+\beta_N = 2$. Comme $|X_n| \leqslant 1$, on obtient :

$$\begin{aligned}\mathbb{E}\,(S_N^4) &\leqslant 6\Big(\sum_{i=1}^N \mathbb{E}\,(X_i^4) + \sum_{i\neq j} \mathbb{E}\,(X_i^2)\mathbb{E}\,(X_j^2)\Big)\\ &\leqslant 6\Big(\sum_{i=1}^N \mathbb{E}\,(X_i^2) + \sum_{i\neq j} \mathbb{E}\,(X_i^2)\mathbb{E}\,(X_j^2)\Big) \leqslant 6(\sigma_N^2 + \sigma_N^4)\,.\end{aligned}$$

Revenant à l'inégalité de Paley-Zygmund, on obtient donc :

$$\mathbb{P}\,(S^2 \geqslant \sigma_N^2/2) \geqslant \frac{1}{24}\,\frac{\sigma_N^4}{\sigma_N^2 + \sigma_N^4}\,.$$

Si l'on avait $\sigma^2 = \lim_{N\to\infty} \sigma_N^2 = +\infty$, cela donnerait, par passage à la limite monotone : $\mathbb{P}\,(S^2 = +\infty) \geqslant 1/24$, ce qui contredirait $\mathbb{P}\,(S < +\infty) = 1$.

On a donc bien $\sum_{n=1}^\infty \mathbb{E}\,(X_n^2) = \sigma^2 < +\infty$. □

On peut supprimer l'hypothèse que les variables soient centrées ($\mathbb{E}\,(X_n) = 0$) de la façon suivante.

Corollaire 3.14 *Si les X_n sont* indépendantes *et si $|X_n| \leqslant M$, $\forall n \geqslant 1$, alors la convergence p.s. de $\sum_{n\geqslant 1} X_n$ entraîne $\sum_{n=1}^\infty \mathrm{Var}\,(X_n) < +\infty$.*

Preuve. Soit $(X_n')_{n\geqslant 1}$ une suite indépendante de *v.a.r.*, et elle-même indépendante de la suite $(X_n)_{n\geqslant 1}$, telle que X_n' ait la **même loi** que X_n pour tout $n \geqslant 1$.

La *v.a.r.* $X_n^s = X_n - X_n'$ est appelée la **symétrisée** de X_n.

Elle est centrée : $\mathbb{E}(X_n^s) = \mathbb{E}(X_n) - \mathbb{E}(X_n') = 0$, et $|X_n^s| \leqslant |X_n| + |X_n'| \leqslant 2M$.

De plus, si la série $\sum_{n\geqslant 1} X_n$ converge *p.s.*, $\sum_{n\geqslant 1} X_n'$ aussi, et donc $\sum_{n\geqslant 1} X_n^s$ aussi.

Le Théorème 3.12 dit alors que $\sum_{n=1}^{\infty} \mathbb{E}(X_n^{s\,2}) < +\infty$. Mais :

$$\mathbb{E}(X_n^{s\,2}) = \mathrm{Var}(X_n^s) = \mathrm{Var}(X_n - X_n') = \mathrm{Var}(X_n) + \mathrm{Var}(X_n') = 2\mathrm{Var}(X_n). \ \square$$

Nous pouvons maintenant énoncer le résultat principal.

Théorème 3.15 (Théorème des trois séries de Kolmogorov)
Si les X_n, $n \geqslant 1$, sont des v.a.r. **indépendantes**, *alors la série $\sum_{n\geqslant 1} X_n$ converge presque sûrement si et seulement si les trois séries de nombres suivantes convergent :*

$$\boxed{a)\ \sum_{n\geqslant 1} \mathbb{P}(|X_n| > M)\ ;\quad b)\ \sum_{n\geqslant 1} \mathrm{Var}(X_n^{[M]})\ ;\quad c)\ \sum_{n\geqslant 1} \mathbb{E}(X_n^{[M]})},$$

où $X_n^{[M]}$ est la v.a.r. X_n **tronquée** *en M : $X_n^{[M]} = X_n.\mathbb{1}_{\{|X_n|\leqslant M\}}$.*

On remarquera que si les conditions b) et c) sont vérifiées pour un $M > 0$, elles le sont pour tout $M > 0$: il suffit de multiplier les variables X_n par une même constante.

Preuve.

1) Supposons que $\sum_{n\geqslant 1} X_n$ converge *p.s.*. Alors :

a) $X_n \xrightarrow[n\to\infty]{p.s.} 0$; on a donc $\mathbb{P}(\overline{\lim_{n\to\infty}} |X_n| > M) = 0$. Mais

$$\left\{ \overline{\lim_{n\to\infty}} |X_n| \geqslant M \right\} \supseteq \overline{\lim_{n\to\infty}} \{|X_n| > M\}.$$

Comme les événements $A_n = \{|X_n| > M\}$ sont indépendants, le lemme de Borel-Cantelli donne $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > M) < +\infty$.

b) Les *v.a.r.* $X_n^{[M]}$ étant indépendantes et bornées par M, on peut utiliser le Théorème 3.12 (ou plutôt le Corollaire 3.14) puisque $\sum_{n\geqslant 1} X_n^{[M]}$ converge *p.s.* : en effet, on a $X_n \xrightarrow[n\to\infty]{p.s.} 0$; donc, presque sûrement, $|X_n| \leqslant M$, pour n assez grand, c'est-à-dire que, presque sûrement, $X_n^{[M]}(\omega) = X_n(\omega)$, pour n assez grand ; on obtient donc $\sum_{n\geqslant 1} \mathrm{Var}(X_n^{[M]}) < +\infty$.

c) Le Théorème de Kolmogorov appliqué à la *v.a.r.* centrée $X_n^{[M]} - \mathbb{E}(X_n^{[M]})$ donne alors la convergence *p.s.* de $\sum_{n\geqslant 1} \left(X_n^{[M]} - \mathbb{E}(X_n^{[M]})\right)$, et donc la convergence de $\sum_{n\geqslant 1} \mathbb{E}(X_n^{[M]})$ puisque $\sum_{n\geqslant 1} X_n^{[M]}$ converge *p.s.* par le *b*).

2) Réciproquement, si les trois séries convergent, la condition *b*) permet, comme ci-dessus, d'appliquer le Théorème de Kolmogorov à $X_n^{[M]} - \mathbb{E}(X_n^{[M]})$

et d'obtenir la convergence *p.s.* de $\sum_{n\geqslant 1}\big(X_n^{[M]} - \mathbb{E}\,(X_n^{[M]})\big)$. La condition *c*) entraîne alors la convergence *p.s.* de $\sum_{n\geqslant 1} X_n^{[M]}$.

Par ailleurs, la condition *a*) entraîne que $\mathbb{P}\big(\varlimsup_{n\to\infty}\{|X_n| > M\}\big) = 0$, *via* le lemme de Borel-Cantelli ; comme $\{|X_n| > M\} = \{X_n \neq X_n^{[M]}\}$, on a donc $X_n = X_n^{[M]}$ *p.s.*, pour n assez grand, et l'on obtient la convergence *p.s.* de $\sum_{n\geqslant 1} X_n$. □

Remarque. 1) Si les X_n sont *symétriques*, alors les $X_n^{[M]}$ aussi ; on a donc $\mathbb{E}\,(X_n^{[M]}) = 0$; la condition *c*) est par conséquent automatiquement vérifiée.

2) Si les X_n sont *positives*, alors :

$$\mathrm{Var}\,(X_n^{[M]}) \leqslant \mathbb{E}\,({X_n^{[M]}}^2) \leqslant \mathbb{E}\,(M.X_n^{[M]}) = M.\mathbb{E}\,(X_n^{[M]})\,,$$

et dans ce cas, la condition *b*) résulte de la condition *c*).

Donnons une application, importante, de ce théorème.

Théorème 3.16 *Soit $(a_n)_{n\geqslant 1}$ une suite de nombres complexes, et $(Z_n)_{n\geqslant 1}$ une suite i.i.d. centrée non nulle, dans L^2. Alors la série $\sum_{n\geqslant 1} a_n Z_n$ converge presque sûrement si et seulement si $\sum_{n=1}^{\infty} |a_n|^2 < +\infty$.*

Preuve. On peut supposer les a_n réels et aussi $\|Z_1\|_2 = 1$ (et donc $\|Z_n\|_2 = 1$ pour tout $n \geqslant 1$).

Si la série $\sum_{n=1}^{\infty} |a_n|^2$ converge, le Théorème de Kolmogorov donne immédiatement la convergence *p.s.* de $\sum_{n=1}^{\infty} a_n Z_n$, puisque $\|Z_n\|_2 = 1$.

Réciproquement, si la série converge *p.s.*, appliquons le Théorème des trois séries à la suite $(a_n Z_n)_{n\geqslant 1}$, avec le troncage en $M = 1$.

On a $\sum_{n=1}^{\infty} \mathbb{P}\,(|Z_1| > 1/|a_n|) < +\infty$; donc, en particulier :

$$\mathbb{P}\,(|Z_1| \leqslant 1/|a_n|) \xrightarrow[n\to\infty]{} 1\,.$$

Mais $\mathbb{E}\,(Y\mathbb{1}_A) \xrightarrow[\mathbb{P}\,(A)\to 1]{} \mathbb{E}\,(Y)$ pour toute *v.a.r.* intégrable Y ; donc :

$$\mathbb{E}\,\big[\big(Z_1^{[1/|a_n|]}\big)^2\big] \xrightarrow[n\to\infty]{} \mathbb{E}\,(Z_1^2) = \|Z_1\|_2^2 = 1\,.$$

De même $\mathbb{E}\,\big[Z_1^{[1/|a_n|]}\big] \xrightarrow[n\to\infty]{} \mathbb{E}\,(Z_1)$. Comme la *v.a.r.* Z_1 est centrée, on a $\mathbb{E}\,(Z_1) = 0$, et donc :

$$\mathrm{Var}\,(Z_1^{[1/|a_n|]}) = \mathbb{E}\,\big[\big(Z_1^{[1/|a_n|]}\big)^2\big] - \big[\mathbb{E}\,\big(Z_1^{[1/|a_n|]}\big)\big]^2 \xrightarrow[n\to\infty]{} 1\,,$$

de sorte que $a_n^2\,\mathrm{Var}\,(Z_1^{[1/|a_n|]}) \sim a_n^2$, et donc $\sum_{n=1}^{\infty} a_n^2 < +\infty$, puisque le Théorème des trois séries nous dit que $\sum_{n=1}^{\infty} a_n^2\,\mathrm{Var}\,(Z_1^{[1/|a_n|]}) < +\infty$. □

4 Lois des grands nombres

On appelle ainsi les théorèmes donnant l'existence de *limites* pour les *moyennes empiriques* :

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}\,.$$

Il y a beaucoup d'énoncés de ce type. On distingue les **lois faibles des grands nombres**, qui donnent des limites en probabilité (en fait souvent en moyenne ou en moyenne quadratique), et les **lois fortes des grands nombres**, qui donnent des limites presques sûres. Il est bien sûr plus intéressant de savoir qu'il se passe quelquechose presque sûrement, que de savoir que cela se passe en moyenne. On donne en général les deux types de lois car les lois faibles sont plus faciles à prouver. De plus, bien que, basiquement, on suppose les *v.a.r.* X_n indépendantes et de même loi, on peut affaiblir les hypothèses pour obtenir des lois faibles, alors que les lois fortes n'ont pas lieu. Il existe donc un grand nombre de types d'énoncés. Nous allons en donner quelques-uns.

4.1 Lois faibles des grands nombres

Dans le premier énoncé, dû à Tchebychev, nous supposons les *v.a.r.* dans L^2 et non corrélées.

Théorème 4.1 (Tchebychev) Loi faible des grands nombres, cas L^2
Si les v.a.r. $X_n \in L^2$ sont centrées et non corrélées, et si :

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbb{E}\,(X_k^2) \xrightarrow[n\to\infty]{} 0\,,$$

alors :

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{L^2} 0\,.$$

En particulier : $\frac{X_1+\cdots+X_n}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 0\,.$

Preuve. La preuve est très simple. Notons d'abord que, les *v.a.r.* étant centrées et non corrélées, elles sont orthogonales dans L^2 ; donc :

$$\|\overline{X}_n\|_2^2 = \frac{1}{n^2}\sum_{k=1}^{n} \|X_k\|_2^2 = \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{E}\,(X_k^2) \xrightarrow[n\to\infty]{} 0\,. \qquad \square$$

Cas particulier. Si les $X_n \in L^2$ sont *centrées*, *non corrélées*, et **de même loi**, alors $\mathbb{E}\,(X_k^2) = \sigma^2$, pour tout $k \geqslant 1$; donc :

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbb{E}\,(X_k^2) = \frac{1}{n^2}\cdot(n\sigma^2) \xrightarrow[n\to\infty]{} 0\,.$$

On obtient donc :

Corollaire 4.2 *Si les $X_n \in L^2$ sont i.i.d., de moyenne m, alors :*

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{L^2} m\,.$$

Avant de donner un énoncé valable dans L^1, rappelons que si $X \in L^1$, alors :
- $(\forall \varepsilon > 0)\ (\exists \delta > 0) \quad \mathbb{P}(A) \leqslant \delta \implies \int_A |X|\, d\mathbb{P} \leqslant \varepsilon$ (absolue continuité) ;
- $\displaystyle\lim_{a\to+\infty} \int_{\{|X|\geqslant a\}} |X|\, d\mathbb{P} = 0\,.$

Pour une partie $\mathscr{F} \subseteq L^1(\mathbb{P})$, ces propriétés peuvent être vérifiées uniformément :

Définition 4.3 *Si $\mathscr{F} \subseteq L^1(\mathbb{P})$, on dit que :*
1) *$\mathscr{F}$ est* **équicontinue** *si :*

$$(\forall \varepsilon > 0)\ (\exists \delta > 0) \qquad \mathbb{P}(A) \leqslant \delta \quad \implies \quad \left(\int_A |X|\, d\mathbb{P} \leqslant \varepsilon\,,\ \ \forall X \in \mathscr{F}\right);$$

2) *$\mathscr{F}$ est* **équi-intégrable** *si :*

$$\lim_{a\to+\infty} \left[\sup_{X\in\mathscr{F}} \int_{\{|X|\geqslant a\}} |X|\, d\mathbb{P}\right] = 0\,.$$

Exemple. S'il existe $Y \in L^1(\mathbb{P})$ telle que $|X| \leqslant Y$, pour tout $X \in \mathscr{F}$, alors $\mathscr{F}$ est équi-intégrable.

On pourra vérifier à titre d'exercice que $\mathscr{F}$ est équi-intégrable si et seulement si elle est équicontinue et *bornée* dans $L^1(\mathbb{P})$.

D'autre part, on montre (Théorème de Dunford-Pettis) que $\mathscr{F}$ est équi-intégrable si et seulement si $\mathscr{F}$ est relativement compact pour la topologie faible de $L^1(\mathbb{P})$.

On a :

Théorème 4.4 (Loi faible des grands nombres, cas L^1)
Si les v.a.r. X_n sont indépendantes, *dans L^1, et ont la* même moyenne *$m = \mathbb{E}(X_n)$, et si $\{X_n\,;\ n \geqslant 1\}$ est* équi-intégrable, *alors $\overline{X}_n \xrightarrow[n\to\infty]{\mathbb{P}} m\,.$*

Corollaire 4.5 (Khintchine) *Si les $X_n \in L^1$ sont i.i.d., alors :*

$$\overline{X}_n \xrightarrow[n\to\infty]{\mathbb{P}} m = \mathbb{E}\,(X_1)\,.$$

Preuve. Comme les $X_n \in L^1$ ont la même loi, elles ont d'une part la même espérance, et d'autre part $\displaystyle\int_{\{|X_n|\geqslant a\}} |X_n|\, d\mathbb{P}$ ne dépend pas de n, de sorte que $\{X_n\,;\ n \geqslant 1\}$ est équi-intégrable. □

Preuve du Théorème. On peut supposer $m = 0$. Soit $\varepsilon > 0$.

Par l'équi-intégrabilité, il existe $M > 0$ tel que :

$$\mathbb{E}\left(|X_n|\, \mathbb{1}_{\{|X_n|\geqslant M\}}\right) \leqslant \varepsilon^2.$$

Tronquons les X_n en M :

$$Y_n = X_n^{[M]} = X_n.\mathbb{1}_{\{|X_n|\leqslant M\}}\,,$$

et posons $Z_n = X_n - Y_n = X_n.\mathbb{1}_{\{|X_n|>M\}}$.

On a :

$$\overline{X}_n = \overline{X}_n - \mathbb{E}\,(\overline{X}_n) = [\overline{Y}_n - \mathbb{E}\,(\overline{Y}_n)] - [\overline{Z}_n - \mathbb{E}\,(\overline{Z}_n)]\,;$$

donc :

$$\mathbb{P}\,(|\overline{X}_n| \geqslant 2\varepsilon) \leqslant \mathbb{P}\,(|\overline{Y}_n - \mathbb{E}\,(\overline{Y}_n)| \geqslant \varepsilon) + \mathbb{P}\,(|\overline{Z}_n - \mathbb{E}\,(\overline{Z}_n)| \geqslant \varepsilon)\,.$$

Or :

- l'inégalité de Bienaymé-Tchebychev donne :

$$\begin{aligned}
\mathbb{P}\,(|\overline{Y}_n - \mathbb{E}\,(\overline{Y}_n)| \geqslant \varepsilon) &\leqslant \frac{1}{\varepsilon^2}\mathrm{Var}\,(\overline{Y}_n) \\
&= \frac{1}{\varepsilon^2}\sum_{k=1}^{n}\frac{\mathrm{Var}\,(Y_k)}{n^2} \quad \text{(par indépendance)} \\
&\leqslant \frac{1}{\varepsilon^2}\sum_{k=1}^{n}\frac{\mathbb{E}\,(Y_k^2)}{n^2} \leqslant \frac{1}{\varepsilon^2}\frac{1}{n^2}\times nM^2 \quad (\text{car } |Y_k| \leqslant M) \\
&= \frac{M^2}{n\varepsilon^2}\,;
\end{aligned}$$

- d'autre part, par l'inégalité de Markov :

$$\begin{aligned}
\mathbb{P}\,(|\overline{Z}_n - \mathbb{E}\,(\overline{Z}_n)| \geqslant \varepsilon) &\leqslant \frac{1}{\varepsilon}\mathbb{E}\,(|\overline{Z}_n - \mathbb{E}\,(\overline{Z}_n)|) \leqslant \frac{1}{\varepsilon}\times 2\,\mathbb{E}\,(|\overline{Z}_n|) \\
&\leqslant \frac{2}{\varepsilon}\times\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\,(|Z_k|) \\
&\leqslant \frac{2}{n\varepsilon}\sum_{k=1}^{n}\mathbb{E}\,(|X_k|\,\mathbb{1}_{\{|X_k|\geqslant M\}}) \leqslant \frac{2}{n\varepsilon}\times n\varepsilon^2 = 2\varepsilon\,.
\end{aligned}$$

Donc :

$$\mathbb{P}\left(|\overline{X}_n| \geqslant 2\varepsilon\right) \leqslant \frac{M^2}{n\varepsilon^2} + 2\varepsilon \leqslant 3\varepsilon$$

pour n assez grand. Par la Proposition 2.2, cela montre que $\overline{X}_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$. □

4.2 Lois fortes des grands nombres

Théorème 4.6 (Rajchman) Loi forte des grands nombres, cas L^2
Si les X_n sont dans L^2, sont **indépendantes** *et* centrées*, alors, pour toute suite* **croissante** *de nombres réels $u_n > 0$ tendant vers $+\infty$, on a :*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left(X_n^2\right)}{u_n^2} < +\infty \quad \Longrightarrow \quad \frac{1}{u_n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{p.s.} 0\,.$$

Le cas le plus usuel est $u_n = n$; on obtient en particulier :

Corollaire 4.7 *Si les $X_n \in L^2$ sont i.i.d., de moyenne m, alors :*

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{p.s.} m\,.$$

Preuve. En effet, on peut appliquer le théorème à $X_n - m$ car $\mathbb{E}\left((X_n - m)^2\right) = \mathrm{Var}\left(X_n\right) = \sigma^2$ ne dépend pas de n et donc :

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left((X_n - m)^2\right)}{n^2} = \sigma^2 \sum_{n=1}^{\infty} \frac{1}{n^2} < +\infty\,. \qquad \square$$

Remarque. Avec les hypothèses du corollaire et les *v.a.r.* X_n centrées, on peut même prendre $u_n = \sqrt{n}(\log n)^\alpha$ pour tout $\alpha > 1/2$; on a donc, pour tout $\alpha > 1/2$:

$$\boxed{\frac{1}{\sqrt{n}(\log n)^\alpha}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{p.s.} 0}\,. \tag{4.1}$$

Preuve du Théorème 4.6. Le Théorème de Kolmogorov, appliqué à X_n/u_n nous dit que la série $\displaystyle\sum_{n\geqslant 1} \frac{X_n}{u_n}$ converge *p.s.*. Il suffit donc d'utiliser le lemme suivant.

Lemme 4.8 (Lemme de Kronecker) *Pour toute suite* croissante *de nombres réels* $u_n > 0$ tendant vers $+\infty$, *et pour toute suite de nombres réels* x_n *telle que* $\displaystyle\sum_{n\geqslant 1} \frac{x_n}{u_n}$ *converge dans* $\mathbb{R}$, *on a :*

$$\frac{1}{u_n} \sum_{k=1}^{n} x_k \xrightarrow[n\to\infty]{} 0\,.$$

Preuve. Posons $u_0 = 0$ et, pour $n \geqslant 1$:

$$v_n = u_n - u_{n-1} \geqslant 0\ , \qquad z_n = \sum_{k=1}^{n} \frac{x_k}{u_k}\cdot$$

Alors :

$$\sum_{n=1}^{\infty} v_n = +\infty$$

et

$$z = \lim_{n\to\infty} z_n \quad \text{existe dans } \mathbb{R}.$$

On a, en posant $z_0 = 0$:

$$\begin{aligned}
\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} u_k \frac{x_k}{u_k} &= \sum_{k=1}^{n} u_k(z_k - z_{k-1}) \\
&= u_1(z_1 - z_0) + u_2(z_2 - z_1) + \cdots + u_n(z_n - z_{n-1}) \\
&= u_1 z_0 + (u_1 - u_2)z_1 + (u_2 - u_3)z_2 + \cdots + (u_{n-1} - u_n)z_{n-1} + u_n z_n \\
&= -(v_1 z_0 + v_2 z_1 + \cdots + v_n z_{n-1}) + (v_1 + \cdots + v_n)z_n \\
&= \sum_{l=1}^{n} v_l(z_n - z_{l-1}).
\end{aligned}$$

Donnons-nous $\varepsilon > 0$ et choisissons $p_\varepsilon \geqslant 1$ tel que pour $n > p \geqslant p_\varepsilon$ on ait :

$$\sup_{l>p} |z_n - z_{l-1}| \leqslant \varepsilon.$$

Alors, pour $n > p \geqslant p_\varepsilon$:

$$\begin{aligned}
\Big|\frac{1}{u_n}\sum_{k=1}^{n} x_k\Big| &\leqslant \frac{1}{u_n}\sum_{l=1}^{p} v_l\,|z_n - z_{l-1}| + \frac{1}{u_n}\sum_{l=p+1}^{n} v_l\ \sup_{l>p}|z_n - z_{l-1}| \\
&\leqslant \frac{1}{u_n}\Big(\sum_{l=1}^{p} v_l\Big)\, 2\sup_{l\geqslant 1}|z_l| + \frac{1}{u_n}\sum_{l=p+1}^{n} v_l\,\varepsilon \\
&= \frac{1}{u_n} u_p\, .2\sup_{l\geqslant 1}|z_l| + \frac{1}{u_n}(u_n - u_p)\,\varepsilon\,;
\end{aligned}$$

d'où, en faisant tendre n vers $+\infty$:

$$\overline{\lim_{n\to\infty}} \left| \frac{1}{u_n} \sum_{k=1}^{n} x_k \right| \leqslant \varepsilon,$$

ce qui prouve le lemme, puisque $\varepsilon > 0$ était arbitraire. $\square$

Théorème 4.9 (Kolmogorov) Loi forte des grands nombres, cas L^1
Si les X_n sont dans L^1 et sont i.i.d., de moyenne m, on a :

$$\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{p.s.} m\,.$$

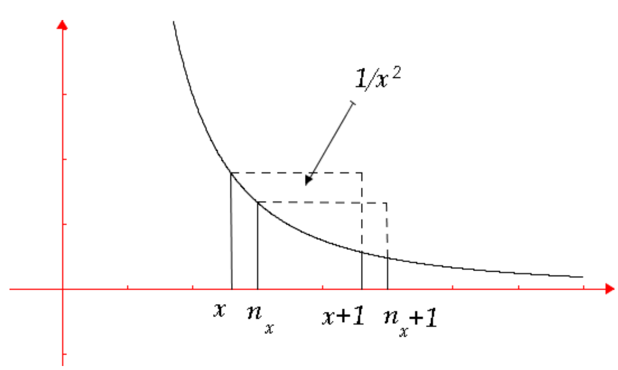
Preuve. On peut supposer, en remplaçant X_n par $X_n - \mathbb{E}\,(X_n)$ que $m = 0$. Nous allons nous ramener au cas L^2, en utilisant les *v.a.r. tronquées* :

$$Y_n = X_n\, \mathbb{1}_{\{|X_n| \leqslant n\}}\,.$$

Comme $|Y_n| \leqslant n$, on a $Y_n \in L^2$; les $Y'_n = Y_n - \mathbb{E}\,(Y_n)$ sont donc dans L^2, et elles sont indépendantes et centrées. De plus :

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mathbb{E}\,({Y'}_n^2)}{n^2} &= \sum_{n=1}^{\infty} \frac{\operatorname{Var}(Y_n)}{n^2} \leqslant \sum_{n=1}^{\infty} \frac{\mathbb{E}\,(Y_n^2)}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|\leqslant n\}}\, d\mathbb{P}_{X_1}(x) \\ &= \int_{\mathbb{R}} x^2 \Big(\sum_{n=1}^{\infty} \frac{1}{n^2} \mathbb{1}_{\{n \geqslant |x|\}} \Big)\, d\mathbb{P}_{X_1}(x) \leqslant \int_{\mathbb{R}} (1+|x|)\, d\mathbb{P}_{X_1}(x) < +\infty\,,\end{aligned}$$

car :



$$\sum_{n\geqslant |x|} \frac{1}{n^2} \leqslant \frac{1}{x^2} + \sum_{n\geqslant n_x+1} \frac{1}{n^2} \leqslant \frac{1}{x^2} + \int_{n_x}^{+\infty} \frac{dt}{t^2} \leqslant \frac{1}{x^2} + \int_{|x|}^{+\infty} \frac{dt}{t^2} = \frac{1}{x^2} + \frac{1}{|x|}\,.$$

Par conséquent, le cas L^2 de la loi forte des grands nombres dit que :

$$\frac{1}{n}\sum_{k=1}^{n} Y'_k \xrightarrow[n\to\infty]{p.s.} 0\,.$$

Mais :

$$\mathbb{E}\,(Y_n) = \mathbb{E}\,(X_n.\mathbb{1}_{\{|X_n|\leqslant n\}}) = \mathbb{E}\,(X_1.\mathbb{1}_{\{|X_1|\leqslant n\}}) \xrightarrow[n\to\infty]{} \mathbb{E}\,(X_1) = 0\,,$$

par convergence dominée. Donc :

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\,(Y_k) \xrightarrow[n\to\infty]{} 0\,,$$

et :

$$\frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{1}{n}\sum_{k=1}^{n} Y'_k + \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\,(Y_k) \xrightarrow[n\to\infty]{p.s.} 0\,.$$

Il reste à voir que $\lim_{n\to\infty}(X_n - Y_n) = 0$ presque sûrement. Mais cela résulte du lemme de Borel-Cantelli :

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbb{P}\,(X_n \neq Y_n) = \sum_{n=1}^{\infty} \mathbb{P}\,(|X_n| > n) &= \int_{\mathbb{R}} \Big(\sum_{n=1}^{\infty} \mathbb{1}_{\{|x|>n\}}\Big)\, d\mathbb{P}_{X_1}(x) \\ &\leqslant \int_{\mathbb{R}} |x|\, d\mathbb{P}_{X_1}(x) = \mathbb{E}\,(|X_1|) < +\infty\end{aligned}$$

entraîne :

$$\mathbb{P}\,(\overline{\lim_{n\to\infty}}\, |X_n - Y_n| > 0) \leqslant \mathbb{P}\,(\overline{\lim_{n\to\infty}}\, \{|X_n - Y_n| > 0\}) = \mathbb{P}\,(\overline{\lim_{n\to\infty}}\, \{X_n \neq Y_n\}) = 0\,.$$

□

5 Le Théorème-Limite Central

5.1 Introduction

Nous avons vu (voir (4.1)) que si les $X_n \in L^2$ sont *i.i.d.* et centrées, alors $S_n = X_1 + \cdots + X_n$ vérifie, pour tout $\alpha > 1/2$:

$$\frac{S_n}{\sqrt{n}(\log n)^\alpha} \xrightarrow[n\to\infty]{p.s.} 0\,.$$

Une question naturelle est de trouver l'*ordre de grandeur* de S_n. Nous venons de voir que S_n croît moins vite que $\sqrt{n}(\log n)^\alpha$ pour tout $\alpha > 1/2$.

D'un autre côté, si $X_n \sim \mathscr{N}(0,1)$ pour tout $n \geqslant 1$, alors $\dfrac{S_n}{\sqrt{n}}$ aussi. Donc pour tout $R > 0$, on a :

$$\mathbb{P}\Big(\frac{S_N}{\sqrt{N}} \geqslant R\Big) = \frac{1}{\sqrt{2\pi}} \int_R^{+\infty} \mathrm{e}^{-x^2/2}\, dx = K_R > 0\,.$$

Or :

$$\bigcap_{n\geqslant 1}\Big\{\sup_{k\geqslant n}\frac{S_k}{\sqrt{k}}\geqslant R\Big\}\subseteq\Big\{\overline{\lim_{n\to\infty}}\,\frac{S_n}{\sqrt{n}}\geqslant R\Big\};$$

donc :

$$\mathbb{P}\Big(\overline{\lim_{n\to\infty}}\,\frac{S_n}{\sqrt{n}}\geqslant R\Big)\geqslant \lim_{n\to\infty}\!\downarrow\mathbb{P}\Big(\sup_{k\geqslant n}\frac{S_k}{\sqrt{k}}\geqslant R\Big)\geqslant \underline{\lim_{n\to\infty}}\,\mathbb{P}\Big(\frac{S_n}{\sqrt{n}}\geqslant R\Big)=K_R>0.$$

Comme $\overline{\lim\limits_{n\to\infty}}\,\dfrac{S_n}{\sqrt{n}}$ est *p.s.* constante par la loi du 0–1, cette constante ne peut être que $+\infty$; donc $\overline{\lim\limits_{n\to\infty}}\,\dfrac{S_n}{\sqrt{n}}\overset{p.s.}{=}+\infty$, de sorte que, dans ce cas, S_n croît plus vite que $\sqrt{n}$.

L'ordre de grandeur doit donc se trouver entre $\sqrt{n}$ et $\sqrt{n}(\log n)^\alpha$, $\alpha>1/2$. Il revient à Khintchine d'avoir trouvé cet ordre de grandeur pour les *v.a.r.* de Bernoulli (1924) ; le cas général est dû à Hartman et Wintner (1941).

Théorème 5.1 (Loi du logarithme itéré) *Si les v.a.r. $X_n\in L^2$ sont i.i.d. et centrées, d'écart-type σ, alors, presque sûrement, l'ensemble des valeurs d'adhérence de $\dfrac{S_n}{\sqrt{n\log(\log n)}}$ est tout l'intervalle $[-\sigma\sqrt{2},+\sigma\sqrt{2}]$.*

En particulier $\overline{\lim\limits_{n\to\infty}}\,\dfrac{S_n}{\sqrt{n\log(\log n)}}\overset{p.s.}{=}\sigma\sqrt{2}$ et $\underline{\lim\limits_{n\to\infty}}\,\dfrac{S_n}{\sqrt{n\log(\log n)}}\overset{p.s.}{=}-\sigma\sqrt{2}$.

La preuve est difficile, et on ne la donnera pas. Il faut en retenir qu'il n'y a **pas de limite** *p.s.*, et pas de “bon” ordre de grandeur pour avoir une limite presque sûre.

Pour trouver un ordre de grandeur, nous allons introduire une notion de convergence plus faible : la **convergence en loi**.

5.2 Convergence en loi

Définition 5.2 *On dit que la suite de v.a.r. $(X_n)_{n\geqslant 1}$* **converge en loi** *(ou aussi* converge en distribution*) vers la v.a.r. X si :*

$$\boxed{\mathbb{E}\big[f(X_n)\big]\xrightarrow[n\to\infty]{}\mathbb{E}\big[f(X)\big]\ ,\quad \forall f\in\mathscr{C}_b(\mathbb{R})\,.}$$

On note $X_n\xrightarrow[n\to\infty]{\mathscr{L}}X$.

$\mathscr{C}_b(\mathbb{R})$ est l'ensemble des fonctions continues bornées sur $\mathbb{R}$. $f(X)$ et $f(X_n)$ sont des notations pour $f \circ X$ et $f \circ X_n$.

Remarque importante. La condition s'écrit aussi :

$$\int_{\mathbb{R}} f(x)\, d\mathbb{P}_{X_n}(x) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(x)\, d\mathbb{P}_X(x) \ , \quad \forall f \in \mathscr{C}_b(\mathbb{R})\,;$$

on voit donc que cette convergence *ne fait intervenir que la* **loi** *des v.a.r.* X_n et X. C'est donc une convergence concernant les **lois** des *v.a.r.* et pas les *v.a.r.* elles-mêmes. En particulier, et contrairement aux autres convergences, les différentes *v.a.r.* X_n et X peuvent très bien ne même pas être définies sur le même espace de probabilité.

Il en résulte des comportements d'apparence inhabituels pour cette convergence.

Exemple. Soit deux *v.a.r.* différentes X et Y, mais ayant la même loi. Alors, si l'on pose $X_n = X$, pour tout $n \geqslant 1$, on a $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$ et $X_n \xrightarrow[n\to\infty]{\mathscr{L}} Y$, bien que $X \neq Y$.

Supposons maintenant que X soit symétrique, *i.e.* ait la même loi que $(-X)$ (par exemple $X \sim \mathscr{N}(0,1)$, ou X *v.a.r.* telle que $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = 1/2$), alors, en prenant $X_n = X$ pour tout $n \geqslant 1$, et $Y = -X$, on a $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$ et $X_n \xrightarrow[n\to\infty]{\mathscr{L}} (-X)$, bien que $2X = X_n + X_n \ \not\xrightarrow[n\to\infty]{\mathscr{L}} X + (-X) = 0$, et $X^2 = (X_n)(X_n) \ \not\xrightarrow[n\to\infty]{\mathscr{L}} (X)(-X) = -X^2$.

Ces précautions étant prises, on a :

Proposition 5.3 *Si* $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$*, alors* $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$.

Nous utiliserons le lemme suivant, disant que, dans la définition de la convergence en loi, il suffit de tester sur les fonctions continues à support compact, au lieu de toutes les fonctions continues bornées sur $\mathbb{R}$.

Lemme 5.4 *Pour que* $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$*, il suffit (et il faut) que :*

$$\mathbb{E}\left[f(X_n)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[f(X)\right], \quad \forall f \in \mathscr{K}(\mathbb{R}) = \mathscr{D}^0(\mathbb{R})\,.$$

Preuve. Soit $f \in \mathscr{C}_b(\mathbb{R})$ arbitraire.

Prenons une suite croissante vers $\mathbb{1}$ de fonctions positives $\varphi_k \in \mathscr{K}(\mathbb{R})$. Alors $f\varphi_k \in \mathscr{K}(\mathbb{R})$, et l'on a :

$$\begin{aligned}|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\leqslant |\mathbb{E}[(f\varphi_k)(X_n)] - \mathbb{E}[(f\varphi_k)(X)]| \\ &\quad + \|f\|_\infty \left[\int_{\mathbb{R}} (\mathbb{1} - \varphi_k)\, d\mathbb{P}_{X_n} + \int_{\mathbb{R}} (\mathbb{1} - \varphi_k)\, d\mathbb{P}_X\right].\end{aligned}$$

Mais, pour k fixé, on a : $\mathbb{E}[(f\varphi_k)(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[(f\varphi_k)(X)]$ et donc :

$$\int_{\mathbb{R}} (\mathbb{1} - \varphi_k)\, d\mathbb{P}_{X_n} = 1 - \int_{\mathbb{R}} \varphi_k \, d\mathbb{P}_{X_n} \xrightarrow[n\to\infty]{} 1 - \int_{\mathbb{R}} \varphi_k \, d\mathbb{P}_X = \int_{\mathbb{R}} (\mathbb{1} - \varphi_k)\, d\mathbb{P}_X \,.$$

(On notera qu'il est essentiel que $\mathbb{P}_X$ et $\mathbb{P}_{X_n}$ soient des probabilités, de sorte que $\int_{\mathbb{R}} d\mathbb{P}_{X_n} = \int_{\mathbb{R}} d\mathbb{P}_X = 1$: il n'y a pas de perte de masse).

Donc :

$$\overline{\lim_{n\to\infty}} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leqslant 2\, \|f\|_\infty \int_{\mathbb{R}} (\mathbb{1} - \varphi_k)\, d\mathbb{P}_X \,.$$

Il reste à faire tendre k vers l'infini, et à utiliser le Théorème de convergence monotone pour obtenir :

$$\overline{\lim_{n\to\infty}} \big|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]\big| = 0\,. \qquad \square$$

Preuve de la Proposition 5.3.

Si $f \in \mathscr{K}(\mathbb{R})$, alors f est *uniformément continue.*

Soit alors $\varepsilon > 0$, et choisissons $\delta > 0$ tel que :

$$|x - x'| \leqslant \delta \quad \Longrightarrow \quad |f(x) - f(x')| \leqslant \varepsilon\,.$$

Choisissons ensuite $N \geqslant 1$ tel que :

$$n \geqslant N \quad \Longrightarrow \quad \mathbb{P}(|X_n - X| > \delta) \leqslant \varepsilon\,.$$

Alors, pour $n \geqslant N$:

$$\begin{aligned}|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\leqslant \int_{\{|X_n - X| \leqslant \delta\}} |f(X_n) - f(X)|\, d\mathbb{P} + \int_{\{|X_n - X| > \delta\}} |f(X_n) - f(X)|\, d\mathbb{P} \\ &\leqslant \varepsilon + 2\|f\|_\infty \mathbb{P}(|X_n - X| > \delta) \\ &\leqslant (1 + 2\|f\|_\infty)\, \varepsilon\,. \qquad \square\end{aligned}$$

En général, la convergence en loi n'entraîne pas la convergence en probabilité, puisque l'on peut avoir $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$ et $Y_n \xrightarrow[n\to\infty]{\mathscr{L}} Y$ sans que $X_n + Y_n$ ne converge en loi vers $X + Y$. Toutefois, dans certains cas, c'est cependant vrai. Voyons-en un tout de suite.

Proposition 5.5 *Si les X_n sont définies sur le même espace de probabilité, et si $X_n \xrightarrow[n\to\infty]{\mathscr{L}} a\mathbb{1}$, la limite étant une v.a.r.* **constante**, *alors $X_n \xrightarrow[n\to\infty]{\mathbb{P}} a\mathbb{1}$.*

Preuve. En effet, soit $\varepsilon > 0$ et φ_ε continue sur $\mathbb{R}$, nulle hors de $[a-\varepsilon, a+\varepsilon]$ telle que $0 \leqslant \varphi_\varepsilon \leqslant 1$ et $\varphi_\varepsilon(a) = 1$. Alors :

$$X_n \xrightarrow[n\to\infty]{\mathscr{L}} a\mathbb{1} \quad \Longrightarrow \quad \mathbb{E}\left[\varphi_\varepsilon(X_n)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[\varphi_\varepsilon(a\mathbb{1})\right] = \int_{\mathbb{R}} \varphi_\varepsilon \, d\delta_a = \varphi_\varepsilon(a) = 1\,.$$

Donc, pour tout $\varepsilon > 0$:

$$\begin{aligned}\mathbb{P}\left(|X_n - a\mathbb{1}| > \varepsilon\right) &= 1 - \mathbb{P}\left(X_n \in [a-\varepsilon, a+\varepsilon]\right) = 1 - \int_{\mathbb{R}} \mathbb{1}_{[a-\varepsilon,a+\varepsilon]} \, d\mathbb{P}_{X_n} \\ &\leqslant 1 - \int_{\mathbb{R}} \varphi_\varepsilon \, d\mathbb{P}_{X_n} = 1 - \mathbb{E}\left[\varphi_\varepsilon(X_n)\right] \xrightarrow[n\to\infty]{} 0\,.\end{aligned}$$ □

Le critère suivant est d'un usage constant.

Théorème 5.6 $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$ *si et seulement si :*

$$\Phi_{X_n}(t) \xrightarrow[n\to\infty]{} \Phi_X(t)\,, \quad \forall t \in \mathbb{R}\,.$$

Preuve. La condition nécessaire est évidente car $x \mapsto \mathrm{e}^{ixt}$ est continue et bornée sur $\mathbb{R}$.

Inversement, si $\Phi_{X_n}(t) \xrightarrow[n\to\infty]{} \Phi_X(t)$ pour tout $t \in \mathbb{R}$, alors, pour toute $f \in L^1(\mathbb{R})$, on a, en utilisant les Théorèmes de Fubini et de convergence dominée :

$$\begin{aligned}\int_{\mathbb{R}} \hat{f}(x) \, d\mathbb{P}_{X_n}(x) &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(t)\, \mathrm{e}^{-2\pi ixt} \, dt\right] d\mathbb{P}_{X_n}(x) \\ &= \int_{\mathbb{R}} f(t)\Phi_{X_n}(-2\pi t) \, dt \\ &\xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(t)\Phi_X(-2\pi t) \, dt = \int_{\mathbb{R}} \hat{f}(x) \, d\mathbb{P}_X(x)\,.\end{aligned}$$

Comme $\mathscr{F}L^1(\mathbb{R}) = \{\hat{f}\,;\ f \in L^1(\mathbb{R})\}$ est dense dans $\mathscr{C}_0(\mathbb{R})$, on obtient, pour toute $g \in \mathscr{C}_0(\mathbb{R}) \supseteq \mathscr{K}(\mathbb{R})$:

$$\int_{\mathbb{R}} g(x) \, d\mathbb{P}_{X_n}(x) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} g(x) \, d\mathbb{P}_X(x)\,.$$ □

On peut alors démontrer le *Théorème-Limite central* (TLC), parfois appelé Théorème de la limite centrée.

Théorème 5.7 (Paul Lévy) Théorème-Limite Central
Soit $(X_n)_{n\geqslant 1}$ *une suite i.i.d. de v.a.r.* $X_n \in L^2(\mathbb{P})$*, de moyenne* m *et de variance* σ^2*. Alors :*

$$\frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathscr{L}} Z_0 \sim \mathscr{N}(0,1)\,.$$

Remarque. Posons $S_n = X_1 + \cdots + X_n$; le théorème s'écrit :

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathscr{L}} Z_0 \sim \mathscr{N}(0,1)\,.$$

On notera que la loi forte des grands nombres dit que :

$$\frac{S_n - nm}{\sigma\sqrt{n}(\log n)^\alpha} \xrightarrow[n\to\infty]{p.s.} 0\,, \quad \forall \alpha > \frac{1}{2}\,,$$

et que la loi du logarithme itéré dit que :

$$\varlimsup_{n\to\infty} \frac{S_n - nm}{\sigma\sqrt{n\log(\log n)}} \overset{p.s.}{=} \sqrt{2}\,; \quad \varliminf_{n\to\infty} \frac{S_n - nm}{\sigma\sqrt{n\log(\log n)}} \overset{p.s.}{=} -\sqrt{2}\,.$$

Le bon ordre de grandeur de $(S_n - nm)/\sigma$ pour la convergence presque sûre est $\sqrt{n\log(\log n)}$, mais il n'y a pas de limite ; lorsque l'on remplace cette valeur par la valeur plus petite $\sqrt{n}$, on a une convergence, en loi, vers $\mathscr{N}(0,1)$.

Preuve. Soit Φ la *f.c.* commune des *v.a.r. centrées* $X'_k = X_k - m$. Comme $X'_k \in L^2$, Φ est deux fois dérivable et a donc un développement limité d'ordre 2 au voisinage de 0 :

$$\Phi(u) = 1 - \frac{\sigma^2}{2}\,u^2 + o\,(u^2),$$

(puisque $\Phi'(0) = i\mathbb{E}\,(X'_k) = 0$ et que $\Phi''(0) = -\mathbb{E}\,({X'_k}^2) = -\sigma^2$).

Alors, si :

$$Y_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - m)\,,$$

l'indépendance donne :

$$\Phi_{Y_n}(t) = \Big[\Phi\Big(\frac{t}{\sigma\sqrt{n}}\Big)\Big]^n = \Big[1 - \frac{t^2}{2n} + o_n\Big(\frac{t^2}{n}\Big)\Big]^n \xrightarrow[n\to\infty]{} \mathrm{e}^{-t^2/2}\,. \qquad \square$$

On notera que le $o_n\left(\frac{t^2}{n}\right)$ apparaissant dans la formule ci-dessus est un terme complexe. Toutefois, on peut définir $\log(1+u) = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k} u^k$ pour $u \in \mathbb{C}$ de module $|u| < 1$; le calcul habituel reste donc valable.

Il y a eu beaucoup de perfectionnements du *TLC*. Il est souhaitable, par exemple, de supprimer la condition, trop forte, que les *v.a.r.* X_n possèdent la même loi. Dans cette direction, on citera le résultat suivant.

Théorème 5.8 (TLC de Lindeberg, 1922) *Soit $X_n \in L^2(\mathbb{P})$ des v.a.r. centrées et indépendantes, et notons σ_{S_n} l'écart-type de $S_n = X_1 + \cdots + X_n$.*

Si l'on a la condition de Lindeberg *:*

$$(\forall \varepsilon > 0) \qquad g_n(\varepsilon) = \frac{1}{\sigma_{S_n}^2} \sum_{k=1}^{n} \int_{\{|X_k| \geqslant \varepsilon \sigma_{S_n}\}} X_k^2 \, d\mathbb{P} \xrightarrow[n\to\infty]{} 0\,, \tag{L}$$

alors :

$$\frac{S_n}{\sigma_{S_n}} \xrightarrow[n\to\infty]{\mathscr{L}} Z_0 \sim \mathscr{N}(0,1)\,.$$

La preuve est un peu calculatoire.

Il faut noter que la condition de Lindeberg (L) est vérifiée si les X_n ont la même loi (ce qui donne le TLC), ou encore si elles sont bornées par une même constante. Elle peut paraître artificielle, mais Feller a montré (1935) qu'elle était nécessaire, à condition de demander que :

$$\max_{k \leqslant n} \frac{\mathrm{Var}\,(X_k)}{\mathrm{Var}\,(S_n)} \xrightarrow[n\to\infty]{} 0\,,$$

condition qui est vérifiée lorsque la condition de Lindeberg l'est :

$$\mathrm{Var}\,(X_k) = \int_{|x| < \varepsilon \sigma_{S_n}} x^2 \, d\mathbb{P}_{X_k} + \int_{|x| \geqslant \varepsilon \sigma_{S_n}} x^2 \, d\mathbb{P}_{X_k} \leqslant \varepsilon^2 \sigma_{S_n}^2 + \sum_{j=1}^{n} \int_{|x| \geqslant \varepsilon \sigma_{S_n}} x^2 \, d\mathbb{P}_{X_k}\,.$$

5.3 Convergence en loi et fonction de répartition

Proposition 5.9 *On a $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$ si et seulement si $F_{X_n}(x) \xrightarrow[n\to\infty]{} F_X(x)$ pour tout $x \in \mathbb{R}$* **tel que** *F_X soit* **continue en** *x.*

Bien sûr, s'il y a convergence des fonctions de répartition pour *tout* $x \in \mathbb{R}$, il y aura convergence en loi ; mais la convegence en loi n'entraîne pas la convergence partout des fonctions de répartition.

Exemple. Si $\mathbb{P}(X_n = -1/n) = \mathbb{P}(X_n = 1/n) = 1/2$, on a $X_n \sim \frac{1}{2}(\delta_{-1/n} + \delta_{1/n})$, et :

$$F_{X_n}(x) = \begin{cases} 0 & \text{si} \quad x < -1/n \\ 1/2 & \text{si} \quad -1/n \leqslant x < 1/n \\ 1 & \text{si} \quad x \geqslant 1/n \end{cases} \xrightarrow[n\to\infty]{} F^*(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1/2 & \text{si} \quad x = 0 \\ 1 & \text{si} \quad x > 0\,. \end{cases}$$

F^* n'est pas une fonction de répartition (elle n'est pas continue à droite en 0), mais, pour tout $x \neq 0$, on a $F^*(x) = \mathbb{1}_{[0,+\infty[}(x)$, qui est la fonction de répartition de $X = 0$. Donc $X_n \xrightarrow[n\to\infty]{\mathscr{L}} 0\,$.

En fait, $X_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$ car, lorsque $n > 1/\varepsilon$, on a $\mathbb{P}\left(|X_n| \geqslant \varepsilon\right) = 0$.

Notons que si la *v.a.r.* limite X a une densité, alors sa fonction de répartition F_X est continue sur $\mathbb{R}$; dans ce cas, la convergence en loi équivaut à la convergence des fonctions de répartitions en tout point. En particulier, si $Z_0 \sim \mathscr{N}(0,1)$, F_{Z_0} est continue et le TLC signifie que pour tous les $a < b$, on a :

$$\mathbb{P}\Big(a \leqslant \frac{S_n - nm}{\sigma\sqrt{n}} \leqslant b\Big) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-x^2/2}\, dx\,.$$

Preuve de la Proposition.

1) Supposons que $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$.

Soit, pour chaque $k \geqslant 1$, $\varphi_k \colon \mathbb{R} \to \mathbb{R}$ définie par :

$$\varphi_k(t) = \begin{cases} 1 & \text{si} \quad t \leqslant x \\ \text{affine} & \text{sur} \quad [x, x+1/k] \\ 0 & \text{si} \quad t \geqslant x + 1/k\,. \end{cases}$$

Alors φ_k est continue et bornée sur $\mathbb{R}$, $0 \leqslant \varphi_k \leqslant \mathbb{1}$, et la suite $(\varphi_k)_{k\geqslant 1}$ décroît vers $\mathbb{1}_{]-\infty, x]}$.

Pour k fixé :

$$F_{X_n}(x) \leqslant \mathbb{E}\left[\varphi_k(X_n)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[\varphi_k(X)\right] \leqslant F_X(x + 1/k)\,.$$

Donc $\overline{\lim}_{n\to\infty} F_{X_n}(x) \leqslant F_X(x)$, en utilisant la continuité à droite de F_X.

De même, si $\psi_k \colon \mathbb{R} \to \mathbb{R}$ est définie par :

$$\psi_k(t) = \begin{cases} 1 & \text{si} \quad t \leqslant x - 1/k \\ \text{affine} & \text{sur} \quad [x - 1/k] \\ 0 & \text{si} \quad t \geqslant x\,, \end{cases}$$

ψ_k est continue et bornée sur $\mathbb{R}$, $0 \leqslant \psi_k \leqslant \mathbb{1}$, et la suite $(\psi_k)_{k\geqslant 1}$ croît vers $\mathbb{1}_{]-\infty, x[}$; et, puisque, pour k fixé :

$$F_{X_n}(x) \geqslant \mathbb{E}\left[\psi_k(X_n)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[\psi_k(X)\right] \geqslant F_X(x - 1/k)\,,$$

on obtient, en utilisant l'*hypothèse sur la continuité de F_X en x* :

$$\underline{\lim}_{n\to\infty} F_{X_n}(x) \geqslant F_X(x)\,.$$

2) Inversement, soit Δ l'ensemble des points de discontinuité de F_X. Pour $a, b \notin \Delta$, on a, par hypothèse :

$$\mathbb{P}_{X_n}(]a, b]) = F_{X_n}(b) - F_{X_n}(a) \xrightarrow[n\to\infty]{} F_X(b) - F_X(a) = \mathbb{P}_X(]a, b])\,.$$

Soit $f \in \mathscr{K}(\mathbb{R}) = \mathscr{D}^0(\mathbb{R})$ continue à support compact, et soit $\varepsilon > 0$.

Comme Δ est **dénombrable** (car F_X est croissante), on peut trouver, grâce à la continuité uniforme de f sur $\mathbb{R}$, une fonction en escalier :

$$\varphi = \sum_{j=1}^{k} c_j \mathbb{1}_{]a_j, b_j]} \,,$$

avec $a_j, b_j \notin \Delta$, pour tout $j \leqslant k$, telle que $\|f - \varphi\|_\infty \leqslant \varepsilon$. Alors :

$$\int_{\mathbb{R}} \varphi \, d\mathbb{P}_{X_n} = \sum_{j=1}^{k} c_j \mathbb{P}_{X_n}(]a_j, b_j]) \xrightarrow[n\to\infty]{} \sum_{j=1}^{k} c_j \mathbb{P}_X(]a_j, b_j]) = \int_{\mathbb{R}} \varphi \, d\mathbb{P}_X \,;$$

donc :

$$\begin{aligned} \overline{\lim_{n\to\infty}} \left| \int_{\mathbb{R}} f \, d\mathbb{P}_{X_n} - \int_{\mathbb{R}} f \, d\mathbb{P}_X \right| &\leqslant \overline{\lim_{n\to\infty}} \Big(\int_{\mathbb{R}} |f - \varphi| \, d\mathbb{P}_{X_n} + \int_{\mathbb{R}} |f - \varphi| \, d\mathbb{P}_X \Big) \\ &\leqslant 2\varepsilon \,. \end{aligned}$$

□

6 Quelques utilisations de la convergence en loi

6.1 Espaces gaussiens

Théorème 6.1 *Toute limite en loi de v.a.r. gaussiennes est encore gaussienne.*

Rappelons que l'on considère les constantes comme des *v.a.r.* gaussiennes (dégénérées).

Preuve. Soit $X_n \sim \mathscr{N}(m_n, \sigma_n^2)$ telles que $X_n \xrightarrow[n\to\infty]{\mathscr{L}} X$. On a :

$$\Phi_X(t) = \lim_{n\to\infty} \Phi_{X_n}(t) = \lim_{n\to\infty} \exp\Big(i\, m_n t - \frac{1}{2}\sigma_n^2 t^2 \Big) \,, \quad \forall t \in \mathbb{R} \,.$$

Alors :
a) *la suite $(\sigma_n)_{n\geqslant 1}$ converge.*
En effet, on a :

$$|\Phi_X(t)| = \lim_{n\to\infty} \mathrm{e}^{-\frac{1}{2}\sigma_n^2 t^2} \,, \qquad \forall t \in \mathbb{R} \,.$$

Or $\Phi_X(0) = 1$ et Φ_X est continue ; donc pour $|t| \leqslant \delta$, on a $|\Phi_X(t)| \geqslant 1/2 > 0$, de sorte que l'on peut écrire :

$$\log |\Phi_X(t)| = \lim_{n\to\infty} \Big(-\frac{1}{2}\sigma_n^2 t^2 \Big),$$

et cela entraîne la convergence de $(\sigma_n)_{n\geqslant 1}$.

b) *la suite $(m_n)_{n\geqslant 1}$ converge.*

Notons pour cela $\sigma = \lim_{n\to\infty}\sigma_n$. On a donc :

$$\lim_{n\to\infty} \mathrm{e}^{i\,m_n t} = \Phi_X(t)\mathrm{e}^{\frac{1}{2}\sigma^2 t^2} = l(t)\,, \quad \forall t\in\mathbb{R}\,.$$

Nous allons alors montrer que :

i) la suite $(m_n)_{n\geqslant 1}$ est bornée.

En effet, comme $|l(t)| = 1$, il existe un borélien B *borné* tel que $\displaystyle\int_B l(t)\,dt \neq 0$ (sinon, on aurait $l = 0$ presque partout) ; on a donc, par convergence dominée :

$$\lim_{n\to\infty}\int_B \mathrm{e}^{i\,m_n t}\,dt = \int_B l(t)\,dt \neq 0\,.$$

Si $(m_n)_{n\geqslant 1}$ n'étaitpas bornée, elle aurait une sous-suite $(m_{n_k})_{k\geqslant 1}$ telle que $|m_{n_k}| \xrightarrow[k\to\infty]{} +\infty$. Mais, cela contredirait le fait que la transformée de Fourier d'une fonction intégrable tend vers 0 à l'infini :

$$\lim_{|x|\to+\infty}\int_B \mathrm{e}^{ixt}\,dt = \lim_{|x|\to+\infty}\widehat{\mathbb{1}_B}\Big(-\frac{x}{2\pi}\Big) = 0\,.$$

ii) la suite $(m_n)_{n\geqslant 1}$ n'a qu'une seule valeur d'adhérence.

En effet, si m et m' sont deux valeurs d'adhérence, on doit avoir :

$$l(t) = \mathrm{e}^{imt} = \mathrm{e}^{im't}$$

pour tout $t\in\mathbb{R}$, d'où $m = m'$ en dérivant par rapport à t, puis en prenant $t = 0$.

c) Il résulte de ce qui précède que $\Phi_X(t) = \exp\big(imt - \frac{1}{2}\sigma^2 t^2\big)$, et par conséquent $X \sim \mathscr{N}(m,\sigma^2)$ (si $\sigma = 0$, $\Phi_X(t) = \mathrm{e}^{imt}$, et X est presque sûrement constante, égale à m). □

Conséquence importante. Rappelons d'abord que si X est une gaussienne, alors $X\in L^r(\mathbb{P})$ pour tout $r\geqslant 1$. Si maintenant $\mathscr{G}\subseteq L^r(\mathbb{P})$ est une partie de $L^r(\mathbb{P})$ formée de gaussiennes, son *adhérence est encore formée de gaussiennes*, puisque la convergence dans $L^r(\mathbb{P})$ entraîne la convergence en probabilité, donc la convergence en loi. Dans la définition suivante, la condition que $\mathscr{G}$ soit fermé est donc peu importante, car son adhérence sera aussi composée de gaussiennes.

Définition 6.2 *On appelle* **sous-espace gaussien** *tout* sous-espace vectoriel *fermé $\mathscr{G}$ de $L^2(\mathbb{P})$ qui est entièrement formé de v.a.r. gaussiennes.*

Remarque importante. Par contre, le fait que $\mathscr{G}$ soit un sous-espace vectoriel a la conséquence essentielle suivante : pour tout choix de $X_1,\ldots,X_n\in\mathscr{G}$, on a $a_1X_1 + \cdots + a_nX_n \in \mathscr{G}$; donc $a_1X_1 + \cdots + a_nX_n$ est une gaussienne, et ceci pour tous $a_1, \ldots, a_n \in \mathbb{R}$. Cela signifie que le vecteur aléatoire $(X_1, \ldots, X_n)$ est un **vecteur gaussien**.

Il en résulte en particulier que si une suite $(X_n)_{n\geqslant 1}$ de *v.a.r.* centrées est orthogonale dans $L^2(\mathbb{P})$ (en particulier, si c'est une base orthonormée de $L^2(\mathbb{P})$), alors cette suite est **indépendante**.

On a :

Théorème 6.3 *Si $\mathscr{G}$ est un sous-espace gaussien de $L^2(\mathbb{P})$, il est fermé dans tous les espaces $L^r(\mathbb{P})$ pour $1 \leqslant r < +\infty$, et toutes les normes $\| \ \|_r$ sont équivalentes sur $\mathscr{G}$.*

En particulier, $\mathscr{G}$ est un sous-espace hilbertien (donc réflexif) de $L^1(\mathbb{P})$.

Preuve. C'est un principe général pour tout sous-espace vectoriel fermé de L^2 qui est contenu dans tous les L^r. En effet :

a) Si $r \geqslant 2$: et si des $X_n \in \mathscr{G}$ vérifient $\|X_n - X\|_r \xrightarrow[n\to\infty]{} 0$, alors, puisque $\|X_n - X\|_2 \leqslant \|X_n - X\|_r$, on a $\|X_n - X\|_2 \xrightarrow[n\to\infty]{} 0$, et donc $X \in \mathscr{G}$. $\mathscr{G}$ est donc fermé dans L^r, et il résulte du Théorème des isomorphismes de Banach qu'il existe $C_r > 0$ telle que : $\|X\|_r \leqslant C_r\|X\|_2$, pour tout $X \in \mathscr{G}$.

b) Si $1 \leqslant r \leqslant 2$, comme $\|X\|_1 \leqslant \|X\|_r \leqslant \|X\|_2$, il suffit de montrer que $\| \ \|_1$ et $\| \ \|_2$ sont équivalentes sur $\mathscr{G}$. Mais l'inégalité de Cauchy-Schwarz donne :

$$\begin{aligned}\|X\|_2^2 = \int_\Omega X^2\, d\mathbb{P} &= \int_\Omega |X|^{1/2}|X|^{3/2}\, d\mathbb{P} \leqslant \Big(\int_\Omega |X|\, d\mathbb{P}\Big)^{1/2}\Big(\int_\Omega |X|^3\, d\mathbb{P}\Big)^{1/2} \\ &= \|X\|_1^{1/2}\|X\|_3^{3/2} \leqslant \|X\|_1^{1/2}\big(C_3\|X\|_2\big)^{3/2},\end{aligned}$$

d'où $\|X\|_2 \leqslant C_3^3\|X\|_1$. □

6.2 Séries de v.a.r. indépendantes

Théorème 6.4 (Paul Lévy) *Si des v.a.r. X_n, $n \geqslant 1$, sont* **indépendantes***, les convergences presque sûre, en probabilité, et en loi de la* **série** *$\sum_{n\geqslant 1} X_n$ sont équivalentes.*

Pour montrer cela, il suffit de montrer, puisque la convergence *p.s.* entraîne toujours la convergence en probabilité, qui elle-même entraîne toujours la convergence en loi, que cette dernière entraîne la convergence *p.s.* pour les séries de *v.a.r.* indépendantes.

Il sera pratique d'utiliser le résultat suivant, sorte de "critère de Cauchy" pour la convergence en loi.

Théorème 6.5 (Théorème de continuité de Paul Lévy) *Soit* $(X_n)_{n\geqslant 1}$ *une suite de variables aléatoires réelles. Si les fonctions caractéristiques* Φ_{X_n}*,* $n \geqslant 1$*, convergent ponctuellement vers une fonction* Φ **continue** *en* 0*, alors* $(X_n)_{n\geqslant 1}$ *converge en loi vers une v.a.r.* X *et* Φ *est la fonction caractéristique de* X*.*

Preuve. Posons $\mu_n = \mathbb{P}_{X_n}$. Comme l'ensemble des probabilités sur $\mathbb{R}$ est une partie de la boule unité du dual de l'espace de Banach séparable $\mathscr{C}_0(\mathbb{R})$, on peut extraire une sous-suite w^*-convergente $(\mu_{n_k})_{k\geqslant 1}$, vers une mesure μ de norme $\leqslant 1$. μ est clairement une mesure positive.

Il s'agit maintenant de montrer que μ est une probabilité, c'est-à-dire qu'il n'y a pas de perte de masse, et donc que $\mu(\mathbb{R}) = 1$. Cela terminera la preuve puisque si X est une *v.a.r.* de loi $\mathbb{P}_X = \mu$, les $\Phi_k = \Phi_{X_{n_k}}$, $k \geqslant 1$, convergeront donc ponctuellement vers Φ_X, et ainsi on aura $\Phi = \Phi_X$. Ensuite, comme la suite entière $(\Phi_{X_n})_{n\geqslant 1}$, converge vers $\Phi = \Phi_X$, on a bien convergence en loi de $(X_n)_{n\geqslant 1}$ vers X.

Pour la conservation de la masse, notons d'abord que, puisque $\Phi_{X_n}(0) = 1$, on a $\Phi(0) = 1$. Considérons le noyau de Cauchy $K(x) = \frac{1}{\pi(1+x^2)}$ et l'unité approchée :

$$K_n(x) = \frac{n}{\pi(1+n^2x^2)}\cdot$$

Comme $(K_n)_{n\geqslant 1}$ est une unité approchée, et comme Φ est continue en 0, on a :

$$\int_{\mathbb{R}} K_n(x)\Phi(x)\,dx = (K_n * \Phi)(0) \xrightarrow[n\to\infty]{} \Phi(0) = 1.$$

Mais, pour tout $n \geqslant 1$, le Théorème de convergence dominée (puisque $|\Phi_k(x)| \leqslant 1$ et $K_n \in L^1(\mathbb{R})$) et le Théorème de Fubini donnent :

$$\begin{aligned}\int_{\mathbb{R}} K_n(x)\Phi(x)\,dx &= \lim_{k\to\infty}\int_{\mathbb{R}} K_n(x)\Phi_k(x)\,dx \\ &= \lim_{k\to\infty}\int_{\mathbb{R}} \hat{K}_n\Big(-\frac{t}{2\pi}\Big)\,d\mu_{n_k}(t) = \int_{\mathbb{R}} \hat{K}_n\Big(-\frac{t}{2\pi}\Big)\,d\mu(t)\,,\end{aligned}$$

puisque $\hat{K}_n \in \mathscr{C}_0(\mathbb{R})$. Il reste à remarquer que $\hat{K}_n(-t/2\pi) = \mathrm{e}^{-|t|/n} \xrightarrow[n\to\infty]{} 1$, de sorte que :

$$\int_{\mathbb{R}} \hat{K}_n(-t/2\pi)\,d\mu(t) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} d\mu = \mu(\mathbb{R})\,.$$

On a donc $\mu(\mathbb{R}) = \Phi(0) = 1$. □

Remarque. On n'a utilisé que la continuité de Φ en 0, mais elle entraîne la continuité sur $\mathbb{R}$; en effet, l'inégalité de Cauchy-Schwarz donne $|\Phi_n(t+h) - \Phi_n(t)|^2 \leqslant 2\big(1 - \mathrm{Re}\,\Phi_n(h)\big)$, et donc, en passant à la limite : $|\Phi(t+h) - \Phi(t)|^2 \leqslant 2\big(1 - \mathrm{Re}\,\Phi(h)\big)$.

Maintenant, pour prouver le Théorème 6.4, allons utiliser les inégalités suivantes, qui permettrons d'appliquer le Théorème des trois séries.

Proposition 6.6 (inégalités de troncature) *Pour toute v.a.r. X, on a, pour tout $t > 0$:*

1) $\displaystyle\int_{-1/t}^{1/t} x^2 \, d\mathbb{P}_X(x) = \int_\Omega X^2 \mathbb{1}_{\{|X| \leqslant 1/t\}} \, d\mathbb{P} \leqslant \frac{3}{t^2}\big(1 - \operatorname{Re}\Phi_X(t)\big)$.

2) $\displaystyle\mathbb{P}\big(|X| \geqslant 1/t\big) \leqslant \frac{7}{t}\int_0^t \big(1 - \operatorname{Re}\Phi_X(u)\big)\, du$.

Preuve.

1) Utilisons l'inégalité :

$$1 - \cos u \geqslant \frac{u^2}{2} - \frac{u^4}{24} ,$$

valable pour tout $u \in \mathbb{R}$. Elle donne :

$$\begin{aligned} 1 - \operatorname{Re}\Phi_X(t) &= \int_\mathbb{R} \big(1 - \cos(tx)\big)\, d\mathbb{P}_X(x) \geqslant \int_\mathbb{R} \frac{1}{2} t^2 x^2 \Big(1 - \frac{t^2x^2}{12}\Big)\, d\mathbb{P}_X(x) \\ &\geqslant \int_{|x|<1/t} \frac{1}{2}t^2x^2 \times \frac{11}{12}\, d\mathbb{P}_X(x) = \frac{11}{24} t^2 \int_{|x|<1/t} x^2 \, d\mathbb{P}_X(x) . \end{aligned}$$

2) Utilisons le théorème de Fubini :

$$\begin{aligned} \frac{1}{t}\int_0^t \big(1 - \operatorname{Re}\Phi_X(u)\big)\, du &= \frac{1}{t}\int_0^t \Big(\int_\mathbb{R} \big(1 - \cos(ux)\big)\, d\mathbb{P}_X(x)\Big)\, du \\ &= \int_\mathbb{R} \Big(1 - \frac{\sin(tx)}{tx}\Big)\, d\mathbb{P}_X(x) \\ &\geqslant \int_{|x|\geqslant 1/t} \Big(1 - \frac{\sin(tx)}{tx}\Big)\, d\mathbb{P}_X(x) \\ &\geqslant \frac{1}{7}\, \mathbb{P}\big(|X| \geqslant 1/t\big) , \end{aligned}$$

car :

$$|v| \geqslant 1 \quad \Longrightarrow \quad \frac{\sin v}{v} \leqslant \frac{6}{7} ,$$

qui est évident si $v \geqslant 7/6$, et vient de :

$$\frac{\sin v}{v} \leqslant 1 - \frac{v^2}{6} + \frac{v^4}{120} = \varphi(v) \leqslant \varphi(1) = \frac{101}{120} < \frac{7}{6}$$

pour $1 \leqslant v \leqslant 7/6$. □

Preuve du Théorème 6.4. Posons $S_n = X_1 + \cdots + X_n$. Par hypothèse $S_n \xrightarrow[n\to\infty]{\mathcal{L}} X$.

Pour tout $t \in \mathbb{R}$, on a :

$$\Phi_X(t) = \lim_{n\to\infty} \Phi_{X_n}(t) = \lim_{n\to\infty} \Phi_{X_1}(t)\dots\Phi_{X_n}(t) = \prod_{k=1}^{\infty} \Phi_{X_k}(t)\,.$$

Comme on l'a dit, les inégalités de troncature vont servir à vérifier les conditions du Théorème des trois séries. Comme ces inégalités ne font intervenir que les parties réelles des fonctions caractéristiques, nous allons nous ramener à des *v.a.r.* dont la fonction caractéristique est réelle, grâce au lemme suivant, immédiat à vérifier.

Lemme 6.7 *Pour toute v.a.r. X, la fonction caractéristique de sa symétrisée X^s vaut $\Phi_{X^s}(t) = |\Phi_X(t)|^2$, pour tout $t \in \mathbb{R}$.*

Il sera pratique ici de définir la symétrisée X^s, non pas sur Ω, mais sur $\Omega\times\Omega$. On posera donc :

$$X^s\colon \quad \begin{array}{ccc} \Omega\times\Omega & \longrightarrow & \mathbb{R} \\ (\omega,\omega') & \longmapsto & X(\omega) - X(\omega')\,. \end{array}$$

Ecrivons alors :

$$\Phi_{X^s}(t) = |\Phi_X(t)|^2 = \prod_{k=1}^{\infty} |\Phi_{X_k}(t)|^2 = \prod_{k=1}^{\infty} \Phi_{X_k^s}(t)\,.$$

Comme $\Phi_{X^s}(0) = 1$, et comme Φ_{X^s} est continue, il existe $\delta > 0$ tel que $\Phi_{X^s}(t) \geqslant 1/2$ pour $|t| \leqslant \delta$. Pour ces valeurs, on a :

$$-\log\big(\Phi_{X^s}(t)\big) = \sum_{k=1}^{\infty} \big[-\log\big(\Phi_{X_k^s}(t)\big)\big] \geqslant \sum_{k=1}^{\infty} \big[1 - \Phi_{X_k^s}(t)\big]\,.$$

Vérifions alors les hypothèses du Théorème des trois séries pour $\sum_{k\geqslant 1} X_k^s$:
(a) la deuxième inégalité de troncature s'écrit :

$$\mathbb{P}\big(|X_k^s| \geqslant 1/\delta\big) \leqslant \frac{7}{\delta}\int_0^{\delta} \big(1 - \Phi_{X_k^s}(t)\big)\,dt\,,$$

d'où, puisque $1 - \Phi_{X_k^s}(t) \geqslant 0$:

$$\begin{aligned}\sum_{k=1}^{\infty} \mathbb{P}\big(|X_k^s| \geqslant 1/\delta\big) &\leqslant \frac{7}{\delta}\int_0^{\delta} \Big[\sum_{k=1}^{\infty} \big(1 - \Phi_{X_k^s}(t)\big)\Big]\,dt \\ &\leqslant \frac{7}{\delta}\int_0^{\delta} \Big[-\log\big(\Phi_{X^s}(t)\big)\Big]\,dt \leqslant \frac{7}{\delta}\times\delta\log 2 = 7\log 2 < +\infty\,;\end{aligned}$$

(b) les *v.a.r.* X_k^s étant symétriques, les variables tronquées en $1/\delta$:

$$\big(X_k^s\big)^{[1/\delta]} = X_k^s.\mathbb{1}_{\{|X_k^s|\leqslant 1/\delta\}}$$

sont aussi symétriques ; donc $\mathbb{E}\left[\left(X_k^s\right)^{[1/\delta]}\right] = 0$, et ainsi la convergence de la série $\sum_{k\geqslant 1}\mathbb{E}\left[\left(X_k^s\right)^{[1/\delta]}\right]$ est évidente ;

(c) puisque $\mathbb{E}\left[\left(X_k^s\right)^{[1/\delta]}\right] = 0$, on a $\operatorname{Var}\left[\left(X_k^s\right)^{[1/\delta]}\right] = \mathbb{E}\left(\left[\left(X_k^s\right)^{[1/\delta]}\right]^2\right)$; la première inégalité de troncature permet alors d'écrire :

$$\begin{aligned}\sum_{k=1}^{\infty}\operatorname{Var}\left[\left(X_k^s\right)^{[1/\delta]}\right] &= \sum_{k=1}^{\infty}\int_{-1/\delta}^{1/\delta} x^2\, d\mathbb{P}_{X_k^s}(x)\\ &\leqslant \frac{3}{\delta^2}\sum_{k=1}^{\infty}\left[1-\Phi_{X_k^s}(\delta)\right] \leqslant \frac{3}{\delta^2}\left[-\log\left(\Phi_{X^s}(\delta)\right)\right] < +\infty\,.\end{aligned}$$

Le Théorème des trois séries nous dit alors que la série $\sum_{k\geqslant 1} X_k^s$ converge presque sûrement, c'est-à-dire que pour presque tout $(\omega,\omega')\in\Omega\times\Omega$, la série $\sum_{k\geqslant 1}\left[X_k(\omega)-X_k(\omega')\right]$ converge. Par le Théorème de Fubini, cela signifie aussi que, pour presque tout $\omega'\in\Omega$, la série $\sum_{k\geqslant 1}\left[X_k(\omega)-X_k(\omega')\right]$ converge pour presque tout $\omega\in\Omega$. Fixons un tel $\omega'\in\Omega$, et posons $a_k = -X_k(\omega')$. La série $\sum_{k\geqslant 1}(X_k+a_k)$ converge donc *p.s.*. Elle converge donc en loi.

Mais, par hypothèse, $\sum_{k\geqslant 1} X_k$ converge en loi. Bien qu'en général on ne puisse pas conclure sur la convergence en loi de la différence, ici les a_k sont des constantes, et cela va permettre de terminer.

Posons $Z=\sum_{k=1}^{\infty}(X_k+a_k)$. Puisque $\Phi_X(0)=\Phi_Z(0)=1$, il existe $\delta>0$ tel que $|\Phi_Z(t)|\geqslant 1/2$ et $|\Phi_X(t)|\geqslant 1/2$ pour $|t|\leqslant\delta$. Comme :

$$\Phi_Z(t)=\lim_{n\to\infty}\prod_{k=1}^{n}\Phi_{(X_k+a_k)}(t)=\lim_{n\to\infty}\prod_{k=1}^{n}\left[\mathrm{e}^{ia_kt}\Phi_{X_k}(t)\right],$$

on obtient : $\displaystyle\lim_{n\to\infty}\prod_{k=1}^{n}\mathrm{e}^{ia_kt} = \frac{\Phi_Z(t)}{\Phi_X(t)}\cdot$ Notons $s_n=\sum_{k=1}^{n}a_k$; d'après ce qui précède, la suite $(\mathrm{e}^{its_n})_{n\geqslant 1}$ converge pour $|t|\leqslant\delta$. Mais l'ensemble des $t\in\mathbb{R}$ pour lesquels cette suite converge est un sous-groupe de $\mathbb{R}$ (car si $\mathrm{e}^{its_n}\xrightarrow[n\to\infty]{}L(t)$ et $\mathrm{e}^{it's_n}\xrightarrow[n\to\infty]{}L(t')$, alors $\mathrm{e}^{i(t-t')s_n}\xrightarrow[n\to\infty]{}L(t)\overline{L(t')}$). Comme il contient l'intervalle $[-\delta,\delta]$, c'est donc $\mathbb{R}$ tout entier. Ainsi $\mathrm{e}^{its_n}\xrightarrow[n\to\infty]{}L(t)$ pour tout $t\in\mathbb{R}$, et la fonction L est continue en 0, puisqu'elle est égale à $\frac{\Phi_Z(t)}{\Phi_X(t)}$ sur $[-\delta,\delta]$. Comme $\mathrm{e}^{its_n}=\Phi_{s_n\mathbb{1}}(t)$ est la fonction caractéristique de la *v.a.r.* constante $s_n\mathbb{1}$, cela prouve, grâce au Théorème de continuité de Paul Lévy, que $(s_n\mathbb{1})_{n\geqslant 1}$ converge en loi. Mais la loi de $s_n\mathbb{1}$ est la masse de Dirac δ_{s_n}, et il est maintenant facile de voir que si $(\delta_{s_n})_{n\geqslant 1}$ converge vers une mesure μ non nulle pour la topologie $*$-faible $\sigma\big(\mathscr{M}(\mathbb{R}),\mathscr{C}_0(\mathbb{R})\big)$, la suite $(s_n)_{n\geqslant 1}$ converge dans $\mathbb{R}$. En effet, d'abord, la suite $(s_n)_{n\geqslant 1}$ est bornée, car si elle ne l'était pas, on pourait en extraire une sous-suite $(s_{n_k})_{k\geqslant 1}$ tendant vers $-\infty$ ou vers $+\infty$; mais alors, pour toute $f\in\mathscr{C}_0(\mathbb{R})$, on aurait $f(s_{n_k})\xrightarrow[k\to\infty]{}0$, et donc $\mu(f)=\lim_{k\to\infty}f(s_{n_k})=0$, c'est-à-dire $\mu=0$, ce qui a été exclu. Ensuite, si $(s_n)_{n\geqslant 1}$ est bornée par M, en prenant une fonction $f_0 \in \mathscr{C}_0(\mathbb{R})$ telle que $f_0(x) = x$ pour $|x| \leqslant M$, on a $s_n = f_0(s_n) = \delta_{s_n}(f_0) \xrightarrow[n\to\infty]{} \mu(f_0)$.

Ainsi donc, la série $\sum\limits_{k\geqslant 1} a_k$ est convergente.

Il ne reste plus qu'à dire que la convergence de $\sum_{k\geqslant 1} a_k$ et la convergence *p.s.* de $\sum_{k\geqslant 1}(X_k + a_k)$ entraînent la converge *p.s.* de la série $\sum_{k\geqslant 1} X_k$. $\square$