ENS Rennes

Université de Rennes 1

Mémoire de Master 2 préparation à l'agrégation.

Leçon 262 - Modes de convergence d'une suite de variables aléatoires. Exemples et applications

17 février 2018

Auteur :
Corentin Kilque

Encadrant :
Dr. Ismael Bailleul

Table des matières

Introduction

L'étude de suites de variables aléatoires ainsi que de leur convergence est une question très naturelle en probabilité. On est souvent amené par exemple à approcher certaines variables aléatoires par des suites de variables aléatoires plus simple ou que l'on sait simuler. On peut citer le fait que la façon effective de simuler une variable aléatoire sur ordinateur se fait à partir d'une loi uniforme sur $[0,1]$, que l'on approche par une suite de variables aléatoires de Bernoulli, au moyen de la décomposition en base deux d'un nombre réel.

Cette leçon s'intéresse alors à préciser la notion de convergence, et les propriétés qui en découlent. On commence par les modes de convergence trajectoriels, puis on donne la convergence en loi. On peut aussi justifier l'ordre de l'introduction des modes de convergence par le soucis d'aller de la notion de convergence la plus intuitive, à celle la moins intuitive. On s'attache, pour chaque mode de convergence donné, à donner ses propriétés, à caractériser la convergence d'une suite, et à donner le lien avec les notions de convergence déjà introduites. Un diagramme donné en annexe résume ces liens.

La première partie commence donc par donner la notion de convergence presque sûre, analogue en théorie de la mesure de la convergence simple. On énonce notamment la traduction en terme de convergence presque sûre des lemmes de Borel-Cantelli qui fournissent des critères très puissants. A la fin de cette partie est donné comme premier développement la preuve qu'une marche aléatoire symétrique sur $\mathbb{Z}^d$ pour $d \geqslant 3$ tend presque sûrement vers l'infini. On introduit ensuite la convergence en probabilité, proche de la convergence presque sûre dans sa définition, et avec laquelle on précise les liens. La métrisabilité de cette convergence est également précisée. Pour conclure cette partie on énonce les lois des grands nombres, qui sont des convergence de suites particulières en probabilité et presque sûrement, et qui trouvent notamment des applications en analyse.

On introduit dans une deuxième partie la convergence en norme $||.||_p$ héritée de la théorie de la mesure, et on étudie notamment les propriétés des espaces L^p. Il est ensuite précisé les liens avec la convergence presque sûre et en probabilité, en introduisant pour cette dernière la notion d'uniforme intégrabilité.

La dernière partie est consacrée à la convergence en loi, qui n'est pas trajectorielle, et une notion de convergence faible puisque toutes les autres notions de convergence implique celle-ci, quitte à extraire. Elle présente néanmoins l'avantage d'être équivalente à la convergence des fonctions caractéristiques, des fonctions génératrices dans le cas discret ainsi que des fonctions de répartition. L'équivalence avec la convergence des fonctions caractéristiques est utilisée en particulier pour montrer le théorème central limite qui constitue un raffinement des lois des grands nombre et fait l'objet du second développement.

On termine ce rapport en apportant la réponse aux questions qui avaient été posées par le jury lors de l'oral.

Janvier 2018.

Dans tout le rapport, on note $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et on considère $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ l'espace $\mathbb{R}^d$ muni de la tribu borélienne. Les variables aléatoires considérées seront des fonctions mesurables $X : \Omega \to \mathbb{R}^d$.

1. Convergence presque sûre et convergence en probabilités

1.1. Convergence presque sûre.

Définition 1 ([PB07, définition V.1.1.])**.** Si $(X_n)_n$ et X sont des variables aléatoires à valeurs dans $\mathbb{R}^d$ on dit que $(X_n)_n$ convergence presque sûrement vers X si

$$\mathbb{P}\Big(\omega \in \Omega, \lim_{n\to+\infty} X_n(\omega) = X(\omega)\Big) = 1.$$

On note alors $\lim_{n\to+\infty} X_n = X$ p.s. ou $X_n \longrightarrow X$ p.s. lorsque $n \to +\infty$.

Proposition 2 ([PB07, remarque V.1.1.])**.** *La suite $(X_n)_n$ converge presque sûrement vers X si et seulement si*

$$\forall \varepsilon > 0, \quad \mathbb{P}\Big(\bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{|X_n - X| < \varepsilon\}\Big) = 1.$$

De la complétude de $\mathbb{R}$ on déduit le critère de Cauchy de convergence presque sûre suivant, très utile puisqu'il n'exige pas de connaitre la limite.

Proposition 3 ([PB07, remarque V.1.1.] critère de Cauchy)**.** *La suite $(X_n)_n$ converge presque sûrement si et seulement si*

$$\forall \varepsilon > 0, \quad \mathbb{P}\Big(\bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{|X_n - X_m| < \varepsilon\}\Big) = 1.$$

Exemple 4 ([PB07, exemple V.1.3.]). On considère $(X_k)_k$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre $p \in [0,1]$. Alors la suite $U_n := \sum_{k=1}^n 2^{-k} X_k$ converge presque sûrement d'après le critère de Cauchy. En effet, si $m < n$,

$$|U_n - U_m| \leqslant \sum_{k=m+1}^{n} 2^{-k} \leqslant 2^{-m}$$

et donc

$$\bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{\omega, |U_n(\omega) - U_m(\omega)| < \varepsilon\} \supset \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{\omega, 2^{-m} < \varepsilon\} = \bigcup_{m\in\mathbb{N}} \{\omega, 2^{-m} < \varepsilon\} = \Omega \quad \forall \varepsilon > 0.$$

Proposition 5. *Si $(X_n)_n$ converge en presque sûrement vers X et que $f : \mathbb{R}^d \to \mathbb{R}^n$ est continue alors $f(X_n)$ converge presque sûrement vers $f(X)$. D'autre part, si pour $1 \leqslant i \leqslant n$, $(X_k^i)_k$ est une suite de variables aléatoires réelles, alors $((X_k^1, \dots, X_k^n))_k$ converge presque sûrement vers $(X^1, \dots, X^n)$ si et seulement si pour tout $1 \leqslant i \leqslant n$, $(X_n^i)_n$ converge presque sûrement vers X^i.*

On déduit immédiatement des lemmes de Borel-Cantelli et de 2 le critère et la caractérisation de convergence presque sûre suivante.

Proposition 6 ([PB07, proposition V.1.2.]). *Soit $(X_n)_n$ une suite de variables aléatoires, alors*

(i) Si pour tout $\varepsilon > 0$, $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| \geqslant \varepsilon) < +\infty$, alors $(X_n)_n$ converge presque sûrement vers X.

(ii) Si les $(X_n)_n$ sont mutuellement indépendantes, alors $(X_n)_n$ converge presque sûrement vers 0 si et seulement si pour tout $\varepsilon > 0$, $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n| \geqslant \varepsilon) < +\infty$.

Contre-exemple 7 ([Ouv09, remarque 10.3.]). Le point (i) du théorème est suffisant mais pas nécessaire comme le montre l'exemple suivant, qui montre également que l'hypothèse d'indépendance dans le deuxième point est bien justifiée. Soit $X_n = \mathbb{1}_{[0,1/n[}$ définie sur $([0,1], \mathcal{B}([0,1]), \mathbb{P})$ où $\mathbb{P}$ est la mesure de Lebesgue. Alors la suite $(X_n)_n$ converge presque sûrement vers 0 alors que pour $\varepsilon > 0$ et $n \in \mathbb{N}^*$, $\mathbb{P}(|X_n| > \varepsilon) = 1/n$ et donc $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n| \geqslant \varepsilon) = +\infty$.

Un résultat sur les séries dont le terme général est majoré par le terme général d'une série convergente nous donne le résultat suivant.

Proposition 8 ([Ouv09, théorème 10.2.]). *Si il existe $(\varepsilon_n)_n$ une suite de réels positifs qui soient le terme général d'une série convergente et que $(X_n)_n$ est une suite de variables aléatoires telles que*

$$\sum_{n=0}^{+\infty} \mathbb{P}(|X_{n+1} - X_n| > \varepsilon_n) < +\infty$$

alors la suite $(X_n)_n$ converge presque sûrement.

Le théorème qui suit ainsi que sa démonstration constitue le premier développement de la leçon.

Théorème 9. *Soit $(e_i)_{1\leqslant i\leqslant d}$ la base canonique dans $\mathbb{R}^d$, et $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs dans $\mathbb{R}^d$ telles que*

$$\mathbb{P}(X_1 = e_i) = \mathbb{P}(X_1 = -ei) = \frac{1}{2d}$$

pour tout $1 \leqslant i \leqslant d$. On note $S_n := \sum_{i=1}^{n} X_i$, $S_0 = 0$. Alors pour $d \geqslant 3$,

$$\mathbb{P}(|S_n| \to +\infty) = 1$$

c'est à dire que $(S_n)_n$ diverge presque sûrement vers $+\infty$.

Démonstration. On cherche à montrer que $\sum \mathbb{P}(S_n = 0)$ converge. En effet, si on note $R = \mathbb{E}[\sum_{n\geqslant 0} \mathbb{1}_{S_n=0}]$ l'espérance du nombre de retours en 0, on a par Fubini-Tonelli, $R = \sum_{n\geqslant 0} \mathbb{P}(S_n = 0)$, et l'on pourra aboutir à une conclusion.

On cherche alors à calculer $\mathbb{P}(S_n = 0)$. On pose, pour $t = (t_1, \cdots, t_d) \in \mathbb{R}^d$,

$$\varphi(t) := \varphi_{X_1}(t) = \mathbb{E}[e^{iX_1 \cdot t}] = \sum_{j=1}^{d} \frac{1}{2d}(e^{it\cdot e_j} + e^{-it\cdot e_j}) = \frac{1}{d}\sum_{j=1}^{d}\cos(t_j).$$

De plus, par indépendance, on a $\varphi_{S_n}(t) = (\varphi(t))^n$.

D'autre part, avec $\varphi_{S_n}(t) = \sum_{k\in\mathbb{Z}^d} \mathbb{P}(S_n = k)e^{ik\cdot t}$, si on pose $T := [-\pi, \pi]$, on a :

$$\frac{1}{(2\pi)^d}\int_{T^d} \varphi_{S_n}(t)dt = \frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d}\int_{T^d} \mathbb{P}(S_n = k)e^{ik\cdot t}dt = \mathbb{P}(S_n = 0)$$

par Fubini puisque $\frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d}\int_{T^d} |\mathbb{P}(S_n = k)e^{ik\cdot t}|dt = \sum_{k\in\mathbb{Z}^d}\mathbb{P}(S_n = k) = 1$ et puisque $\frac{1}{(2\pi)^d}\int_{T^d} e^{ik\cdot t}dt = \frac{1}{(2\pi)^d}\int_{T^d}\prod_{j=1}^{d} e^{ik_jt_j}dt = \delta_0^k$.

Puisqu'il faut un nombre pair d'étape à $(S_n)_n$ pour revenir en 0, on sait que pour n impair, $\mathbb{P}(S_n = 0) = 0$. On remarque que $|\varphi| \leqslant 1$ sur T^d et $|\varphi(t)| = 1$ si et seulement si $t = (0, \ldots, 0), \pm(\pi, \ldots, \pi)$. On obtient alors :

$$\begin{aligned}\sum_{n\geqslant 0}\mathbb{P}(S_n = 0) &= \frac{1}{(2\pi)^d}\sum_{n\geqslant 0}\int_{T^d}\varphi_{S_{2n}}(t)dt = \frac{1}{(2\pi)^d}\sum_{n\geqslant 0}\int_{T^d}\varphi(t)^{2n}dt \\ &= \frac{1}{(2\pi)^d}\int_{T^d}\sum_{n\geqslant 0}\varphi(t)^{2n}dt = \frac{1}{(2\pi)^d}\int_{T^d}\frac{1}{1-\varphi(t)^2}dt\end{aligned}$$

par Fubini-Tonelli puisqu'on a justifié que $\varphi(t)$ est réel et puisque $|\varphi| < 1$ presque partout sur T^d.

Puisque $\frac{1}{1-\varphi^2}$ est continue sur $T^d \backslash \{(0, \ldots, 0), \pm(\pi, \ldots, \pi)\}$, il nous reste maintenant à justifier l'intégrabilité de la fonction $\frac{1}{1-\varphi^2}$ en $(0, \ldots, 0), \pm(\pi, \ldots, \pi)$. On se limite au point $(0, \ldots, 0)$ puisque pour tout $y \in \mathbb{R}$, $\cos(y \pm \pi) = -\cos(y)$.

Or, en 0,

$$\varphi(t) = \frac{1}{d}\sum_{j=1}^{d}\left(1 - \frac{t_j^2}{2} + o(t_j^2)\right) = 1 - \frac{||t||^2}{2d} + o(||t||^2)$$

et ainsi

$$1 - \varphi(t)^2 = \frac{||t||^2}{d} + o(||t||^2)$$

donc

$$\frac{1}{1-\varphi(t)^2} \sim \frac{d}{||t||^2}$$

qui est intégrable en 0 dès que $d \geqslant 3$.

Soit maintenant $k \in \mathbb{Z}^d$, on montre que $\sum_{n\geqslant 0} \mathbb{P}(S_n = k)$ converge en se ramenant en 0 : on pose $l = |k|$, on a, pour $n \geqslant l+1$

$$\begin{aligned}\mathbb{P}(S_n = 0) &\geqslant \mathbb{P}(S_n = 0 \text{ et } S_l = -k)\\ &\geqslant \mathbb{P}(X_1 + \cdots + X_l = -k \text{ et } X_{l+1} + \cdots + X_n = k)\\ &= \mathbb{P}(S_l = -k)\mathbb{P}(S_{n-l} = k)\end{aligned}$$

car les X_i sont i.i.d. Or, puisque $l = |k|$, $\mathbb{P}(S_l = -k) \geqslant \frac{1}{(2d)^l} > 0$ donc

$$\sum_{n\geqslant 0} \mathbb{P}(S_n = k) = \sum_{n\geqslant 1} \mathbb{P}(S_n = k) = \sum_{n\geqslant l+1} \mathbb{P}(S_{n-l} = k) \leqslant \frac{1}{\mathbb{P}(S_l = -k)} \sum_{n\geqslant l+1} \mathbb{P}(S_n = 0) < \infty.$$

On peut alors conclure : pour $k \in \mathbb{Z}^d$, on note $N_k := \sum_{n\geqslant 0} \mathbb{1}_{S_n = k}$ le nombre de retours de S_n en k, et puisque $\mathbb{E}[N_k] = \sum_{n\geqslant 0} \mathbb{P}(S_n = k)$, N_k est fini presque sûrement. On a alors

$$\begin{aligned}\mathbb{P}(|S_n| \to +\infty) &= \mathbb{P}(\forall A \in \mathbb{N}^*,\ \exists n_0,\ \forall n \geqslant n_0,\ |S_n| \geqslant A)\\ &= \lim_{A\to+\infty} \mathbb{P}(\exists n_0,\ \forall n \geqslant n_0,\ |S_n| \geqslant A)\\ &= \lim_{A\to+\infty} \mathbb{P}(\forall k,\ |k| \leqslant A,\ N_k \text{ est fini})\\ &= 1\end{aligned}$$

puisque $(\{\exists n_0,\ \forall n \geqslant n_0,\ |S_n| \geqslant A\})_{A\in\mathbb{N}^*}$ est décroissante, que $\{\exists n_0,\ \forall n \geqslant n_0,\ |S_n| \geqslant A\} = \{\forall k,\ |k| \leqslant A,\ N_k \text{ est fini}\}$ et puisque l'union dénombrable d'ensembles de mesure nulle est de mesure nulle. □

1.2. **Convergence en probabilité.**

Définition 10 ([PB07, définition V.2.1.])**.** Si $(X_n)_n$ et X sont des variables aléatoires on dit que $(X_n)_n$ converge en probabilité vers X si

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} \mathbb{P}(|X_n - X| \geqslant \varepsilon) = 0.$$

On note alors $\lim_{n\to+\infty} X_n = X$ en probabilité ou $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$.

Proposition 11 ([Ouv09, remarque 10.1.])**.** *La limite en probabilité d'une suite de variables aléatoires est presque sûrement unique, c'est à dire que si $(X_n)_n$ est une suite de variables aléatoires telle que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ et $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} Y$ alors $X = Y$ presque sûrement.*

Démonstration. On a, par inégalité triangulaire, pour tout $\varepsilon > 0$ et tout $n \in \mathbb{N}$

$$(|X - Y| > \varepsilon) \subset (|X - X_n| > \varepsilon/2) \cup (|Y - X_n| > \varepsilon/2)$$

ainsi

$$\mathbb{P}(|X - Y| > \varepsilon) \leqslant \mathbb{P}(|X - X_n| > \varepsilon/2) + \mathbb{P}(|Y - X_n| > \varepsilon/2)$$

et donc, par convergence en probabilité,

$$\mathbb{P}(|X - Y| > \varepsilon) = 0.$$

Or

$$(X \neq Y) = \bigcup_{n \in \mathbb{N}^*} \Big(|X - Y| > 1/n \Big)$$

et on conclut avec la σ-sous additivité de la mesure $\mathbb{P}$. □

Proposition 12 ([PB07, proposition V.2.5.]). *Soit $(X_n)_n$ et $(Y_n)_n$ deux suites de variables aléatoires qui convergent en probabilité respectivement vers X et Y, alors*

(i) Si $f : \mathbb{R}^d \to \mathbb{R}^n$ est une application continue, alors $f(X_n) \xrightarrow[n \to +\infty]{\mathbb{P}} f(X)$

(ii) Pour tous $\alpha, \beta \in \mathbb{R}$, $\alpha X_n + \beta Y_n \xrightarrow[n \to +\infty]{\mathbb{P}} \alpha X + \beta Y$

(iii) $X_n Y_n \xrightarrow[n \to +\infty]{\mathbb{P}} XY$.

On a un premier résultat sur le lien entre les modes de convergence.

Proposition 13 ([Ouv09, théorème 10.4.]). *Si $(X_n)_n$ converge presque sûrement vers X alors $(X_n)_n$ converge en probabilité vers X.*

La réciproque de ce théorème est fausse comme le montre le contre-exemple suivant.

Contre-exemple 14 ([Ouv09, théorème 10.4.]). On considère une suite de variables aléatoires réelles indépendantes $(X_n)_n$ de loi de Bernoulli $\mathcal{B}(1/n)$. Alors puisque pour $0 < \varepsilon < 1$,

$$\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n = 1) = \frac{1}{n}$$

la suite $(X_n)_n$ converge en probabilité vers 0. On utilise alors le point (ii) de la proposition 6 pour montrer que $(X_n)_n$ ne converge pas presque sûrement vers 0. En effet,

$$\sum_{n \in \mathbb{N}^*} \mathbb{P}(|X_n| \geqslant \varepsilon) = \sum_{n \in \mathbb{N}^*} \frac{1}{n} = +\infty.$$

Définition 15 ([PB07, lemme V.2.3.]). On pose pour X, Y deux variables aléatoires définies sur le même espace probabilisé,

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

dont on vérifie facilement que cela définit une distance.

La distance que l'on vient d'introduire est la distance de convergence en probabilité (il n'y a pas unicité) comme le montre la proposition suivante.

Proposition 16 ([PB07, lemme V.2.3.]). *La suite $(X_n)_n$ converge en probabilité vers X si et seulement si $d(X_n, X) \xrightarrow[n \to +\infty]{} 0$.*

On peut déduire de la métrisabilité de la convergence en probabilité la caractérisation suivante, dont le condition nécessaire est très souvent utilisé.

Théorème 17 ([PB07, théorème V.2.4.]). *Soit $(X_n)_n, X$ des variables aléatoires. La suite $(X_n)_n$ converge en probabilité vers X si et seulement de toute suite croissante d'entiers $(n_k)_k$ on peut extraire une sous-suite $(n_{k_i})_i$ telle que $(X_{n_{k_i}})_i$ converge presque sûrement vers X.*

Démonstration. On montre que l'on peut extraire une sous-suite de $(X_n)_n$ qui converge presque sûrement vers X ce qui montrera le sens direct en l'appliquant à toute sous-suite. On construit une suite extraite ainsi : $n_0 = 1$ et pour $k \geqslant 1$ on pose n_k le plus petit entier n tel que
$$\mathbb{P}\Big(|X_n - X| \geqslant \frac{1}{k}\Big) \leqslant 2^{-k}$$
qui existe par convergence en probabilité et propriété de $\mathbb{N}$. On a alors
$$\sum_{k\in\mathbb{N}^*} \mathbb{P}(|X_{n_k} - X| \geqslant \frac{1}{k}) < +\infty$$
et donc pour $\varepsilon > 0$,
$$\sum_{k\in\mathbb{N}^*} \mathbb{P}(|X_{n_k} - X| \geqslant \varepsilon) < +\infty$$
et on conclut par le point (i) de la proposition 6.

Réciproquement, si de toute suite croissante d'entiers $(n_k)_k$ on peut extraire une sous-suite $(n_{k_i})_i$ telle que $(X_{n_{k_i}})_i$ converge presque sûrement vers X, en particulier de toute suite croissante d'entiers $(n_k)_k$ on peut extraire une sous-suite $(n_{k_i})_i$ telle que $(X_{n_{k_i}})_i$ converge en probabilité vers X et donc puisque la convergence est métrisable, $(X_n)_n$ converge en probabilité vers X. □

La métrisabilité permet également de montrer le critère de Cauchy suivant.

Proposition 18 ([PB07, théorème V.2.6.]). *Si la suite $(X_n)_n$ est de Cauchy en probabilité, c'est à dire si*
$$\forall \varepsilon > 0, \exists n \geqslant 0, \forall p, q \geqslant n, \quad \mathbb{P}(|X_p - X_q| \geqslant \varepsilon) \leqslant \varepsilon$$
alors la suite $(X_n)_n$ converge en probabilité.

1.3. **Lois des grands nombres.** On donne maintenant des lois les grands nombres : des résultats de convergence en probabilité et presque sûre de suites particulières qui trouvent de nombreuses applications.

La démonstration du théorème qui suit est élémentaire, tandis que celle du suivant, qui a des hypothèses plus faibles, nécessite plus de travail.

Théorème 19 (Loi faible des grands nombres, [Ouv09, théorème 10.18.]). *Soit $(X_n)_n$ une suite de variables aléatoires indépendantes et identiquement distribuées admettant un moment d'ordre deux. On note $m := \mathbb{E}[X_1]$ et $\sigma^2 := \mathrm{Var}(X_1)$. Alors si on pose*
$$S_n = \sum_{k=1}^{n} X_n,$$

$\left(\frac{S_n}{n}\right)_n$ *converge en probabilité vers* m.

Démonstration. On remarque que, pour tout $n \geqslant 1$, par linéarité $\mathbb{E}[S_n] = nm$ et par indépendance $\text{Var}(S_n) = n\sigma^2$. On a alors, par inégalité de Tchebychev, pour $\varepsilon > 0$ et $n \geqslant 1$,

$$\mathbb{P}\Big(\Big|\frac{S_n}{n} - m\Big| > \varepsilon\Big) = \mathbb{P}\big(|S_n - nm| > n\varepsilon\big) \leqslant \frac{Var(S_n)}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

d'où le résultat. □

Théorème 20 (Loi faible des grands nombres, [Ouv09, théorème 10.19.])**.** *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes et identiquement distribuées admettant une moyenne que l'on note* $m := \mathbb{E}[X_1]$. *Alors si on pose*

$$S_n = \sum_{k=1}^{n} X_n,$$

$\left(\frac{S_n}{n}\right)_n$ *converge en probabilité vers* m.

On donne alors le théorème de la loi forte des grands nombres, qui admet les même hypothèses que précédemment mais dont la conclusion est plus forte. Ici encore, si l'on suppose l'existence d'une moment d'ordre quatre pour la loi commune la démonstration est plus aisée, mais on ne la détaillera pas ici.

Théorème 21 (Loi forte des grands nombres, [PB07, théorème V.5.2.])**.** *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes et identiquement distribuées admettant une moyenne que l'on note* $m := \mathbb{E}[X_1]$. *Alors si on pose*

$$S_n = \sum_{k=1}^{n} X_n,$$

$\left(\frac{S_n}{n}\right)_n$ *converge presque sûrement vers* m.

Exemple 22. Si $(X_n)_n$ admet pour loi commune une loi de Bernoulli de paramètre p, alors on sait que S_n la somme des $(X_i)_{1\leqslant i\leqslant n}$ suit une loi binomiale $\mathcal{B}(n,p)$.

Application 23 (Méthode de Monte-Carlo, [Ouv09, exercice 10.8.])**.** *Soit* $f : [0,1] \to \mathbb{R}$ *mesurable et* $(U_n)_n$ *une suite de variables aléatoires indépendantes de loi uniforme sur* $[0,1]$. *Alors si* $X_n = f(U_n)$, *d'après la loi des grands nombres,*

$$\frac{\sum_{i=1}^{n} X_n}{n} \xrightarrow[n\to+\infty]{p.s.} \mathbb{E}[X_1] = \int_0^1 f(x)dx.$$

Une application de la loi des grands nombre en analyse nous est donnée par le théorème suivant, qui fournit une suite explicite de polynômes qui approchent uniformément une fonction continue sur $[0,1]$ donnée, ce qui précise le théorème d'approximation de Weierstrass.

Théorème 24 ([HQ02, théorème XIII.II.3.]). *Soit $f : [0,1] \to \mathbb{C}$ une fonction continue et soit $B_n(f,x)$ le polynôme $B_n(f,x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n})$. Alors B_n converge uniformément vers f sur $[0,1]$.*

Démonstration. On note $||.||_\infty$ la norme sup sur $[0,1]$ et $\omega(\delta)$ le module d'uniforme continuité de f, c'est à dire

$$\omega(\delta) := \sup\{|f(x) - f(y)|, |x-y| \leqslant \delta\}.$$

On remarque que en reprenant les notations de l'exemple 22 et en notant $x = p \in [0,1]$, $B_n(f,x) = \mathbb{E}[f(S_n(x)/n)]$. Ainsi, on a, pour $x \in [0,1]$, et $\delta \in [0,1]$,

$$\begin{aligned}|f(x) - B_n(f,x)| &= \left|\mathbb{E}\left[f(x) - f\left(\frac{S_n(x)}{n}\right)\right]\right| \leqslant \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n(x)}{n}\right)\right|\right] \\ &= \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n(x)}{n}\right)\right| \mathbb{1}_{\left|x - \frac{S_n(x)}{n}\right| \leqslant \delta}\right] + \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n(x)}{n}\right)\right| \mathbb{1}_{\left|x - \frac{S_n(x)}{n}\right| > \delta}\right] \\ &\leqslant \omega(\delta) + 2||f||_\infty \mathbb{P}\left(\left|x - \frac{S_n(x)}{n}\right| > \delta\right) \leqslant \omega(\delta) + \frac{2||f||_\infty \operatorname{Var}\left(\frac{S_n(x)}{n}\right)}{\delta^2}\end{aligned}$$

or $\operatorname{Var}\left(\frac{S_n(x)}{n}\right) = \frac{x(1-x)}{n} \leqslant \frac{1}{2n}$ et on peut conclure. □

2. CONVERGENCE DANS $L^p(\mathbb{P})$

Dans cette partie, les variables aléatoires considérées sont réelles.

2.1. Définition et première propriétés.

Définition 25 ([MB06, définition 9.5.]). On définit, pour $1 \leqslant p < +\infty$ l'espace vectoriel

$$L^p(\mathbb{P}) = \mathcal{L}^p(\mathbb{P}) / \sim$$

où

$$\mathcal{L}^p(\mathbb{P}) = \left\{X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \text{ mesurable}, \int_\Omega |X|^p d\mathbb{P} < +\infty\right\}$$

et où $\sim$ est la relation d'équivalence donnée par l'égalité presque partout.

Définition 26 ([MB06, théorème 9.1.]). Pour $1 \leqslant p < +\infty$, si $X \in L^p(\mathbb{P})$, on définit

$$||X||_p = \left(\int_\Omega |X|^p d\mathbb{P}\right)^{1/p}$$

appelé *norme p* de X.

Proposition 27 (Inégalité de Hölder, [MB06, théorème 9.1.]). *Si $1 < p, q < +\infty$ sont tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit qu'ils sont* conjugués*), alors si $X \in L^p$ et $Y \in L^q$, $XY \in L^1$ et*

$$||XY||_1 \leqslant ||X||_p ||Y||_q.$$

Proposition 28 (Inégalité de Minkowski, [MB06, théorème 9.2.]). *Si* $1 \leqslant p < +\infty$ *et* $X, Y \in L^p$, *alors*

$$||X + Y||_p \leqslant ||X||_p + ||Y||_p.$$

Théorème 29 (Riesz-Fischer, [MB06, théorème 9.3.]). *Pour* $1 \leqslant p < +\infty$, *d'après l'inégalité de Minkowski,* $(L^p, ||.||_p)$ *est un espace vectoriel normé, il est de plus complet.*

Proposition 30 ([MB06, corollaire 9.2.]). *Si* $1 \leqslant p \leqslant q < +\infty$, *alors si* $X \in L^q$, $X \in L^p$, *c'est à dire que* $L^q \subset L^p$ *et de plus l'injection est continue, au sens où*

$$||X||_p \leqslant ||X||_q.$$

Corollaire 31. *La convergence en norme* $||.||_q$ *implique la convergence en norme* $||.||_p$.

2.2. **Liens avec la convergence presque sûre.** Il n'existe pas d'implication sans hypothèses supplémentaires entre convergence presque sûre et convergence en norme $||.||_p$ comme le montre les contre-exemples qui suivent. Il existe cependant des liens, que l'on précise maintenant.

Proposition 32 ([MB06, théorème 9.3.]). *Si* $(X_n)_n$ *converge dans* L^p *vers* X $(1 \leqslant p < +\infty)$ *alors* $(X_n)_n$ *converge vers* X *presque sûrement à une sous-suite près.*

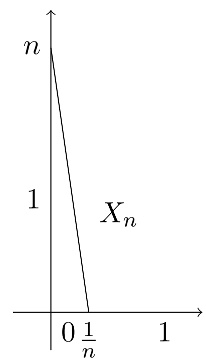
La convergence presque sûre de toute la suite n'est pas vraie en général comme le montre le contre-exemple suivant.

Contre-exemple 33. On définit la suite $(X_n)_n$ sur $([0,1], \mathcal{B}([0,1]), \lambda)$ de la manière suivante : $X_0 = \mathbb{1}_{[0,1]}$, $X_1 = \mathbb{1}_{[0,1/2]}$, $X_2 = \mathbb{1}_{[1/2,1]}$, $X_3 = \mathbb{1}_{[0,1/3]}$, $X_4 = \mathbb{1}_{[1/3,2/3]}$, $X_5 = \mathbb{1}_{[2/3,1]}$, et on continue par récurrence. Ainsi construite, on a $||X_n||_1 \underset{n\to+\infty}{\longrightarrow} 0$ mais $(X_n)_n$ ne converge vers 0 en aucune point de $[0,1]$. Pour obtenir une convergence presque sûre on pourrait par exemple considérer la suite extraite $(X_0, X_1, X_3, X_6, \dots)$ qui converge vers 0 sur $]0,1]$.

Théorème 34 (Beppo-Levi, [MB06, théorème 7.1.]). *Soit* $(X_n)_n$ *une suite croissante de variables aléatoires réelles positives. Alors* $\mathbb{E}[\lim_{n\to+\infty} X_n] = \lim_{n\to+\infty} \mathbb{E}[X_n]$.

Théorème 35 (Convergence dominée, [MB06, proposition 9.6.]). *Soit* $1 \leqslant p < +\infty$ *et soit* $(X_n)_n$ *une suite de* L^p *qui converge presque sûrement vers* X. *S'il existe* $Y \in L^p$ *telle que* $|X_n| \leqslant Y$ *pour tout* $n \in \mathbb{N}$ *et* $\mathbb{P}$*-presque partout, alors* $||X_n - X||_p \underset{n\to+\infty}{\longrightarrow} 0$.

Contre-exemple 36. On définit la suite $(X_n)_n$ sur $([0,1], \mathcal{B}([0,1]), \lambda)$ de la manière suivante : $X_n = (n - n^2 t)\mathbb{1}_{[0,1/n]}$. Alors $(X_n)_n$ converge presque sûrement vers 0 mais $||X_n||_1 = 1/2$ qui ne converge pas vers 0.



On donne ici un autre exemple, plus probabiliste.

Contre-exemple 37 ([PB07, exemple V.3.2.]). Soit $p > 1$ et soit $(X_n)_n$ la suite de variables aléatoires définies sur $\mathbb{R}$ de lois $\mathbb{P}(X_n = n) = n^{-p} = 1 - \mathbb{P}(X_n = 0)$. Si $\varepsilon > 0$, alors pour tout $n \geqslant 1$ tel que $n \geqslant \varepsilon$, $\mathbb{P}(|X_n| > \varepsilon) = n^{-p}$ et donc par le lemme de Borel-Cantelli (proposition 6), puisque $p > 1$, $(X_n)_n$ converge presque sûrement vers 0. Or pour tout $n \geqslant 1$, $||X_n||_p = 1$.

On donne maintenant un exemple qui sera important pour la notion d'uniforme intégrabilité.

Exemple 38. Si $X \in L^1$, alors $(X\mathbb{1}_{|X|>n})_n$ converge vers 0 dans L^1 par le théorème de convergence dominée puisque la suite converge presque sûrement vers 0 et que l'on peut la dominer partout par $|X| \in L^1$.

2.3. **Uniforme intégrabilité et convergence en probabilité.** On donne un premier lien entre la convergence en probabilité et la convergence L^p.

Proposition 39 ([PB07, théorème V.3.5.]). *Si $(X_n)_n$ converge dans L^p, $p \geqslant 1$ vers X, alors $(X_n)_n$ converge en probabilité vers X.*

Démonstration. On va utiliser le théorème 17 pour montrer la convergence en probabilité. Soit $(n_k)_k$ une suite croissante d'entiers. Alors $(X_{n_k})_k$ converge encore dans L^p vers X, et donc d'après la proposition 32 il existe une sous suite de $(X_{n_k})_k$ qui converge presque sûrement, on peut donc conclure. □

La réciproque de ce théorème est fausse comme le montre le contre-exemple suivant.

Contre-exemple 40 ([PB07, exemple V.3.2.]). Pour $\alpha > 0$, on définit sur $([0,1], \mathcal{B}([0,1]), \mathbb{P})$ où $\mathbb{P} = \lambda$ la suite de variables aléatoires réelles $(X_n)_n$ par, pour $\omega \in [0,1]$,

$$X_n(\omega) = \omega^{-\alpha}\mathbb{1}_{]0,1/n]}(\omega).$$

Alors pour $\varepsilon \in]0,1[$, $\mathbb{P}(|X_n| > \varepsilon) = 1/n$ et donc $(X_n)_n$ converge en probabilité vers 0, mais si $p \geqslant 1$ et que α est tel que $\alpha p \geqslant 1$, alors

$$\int_{[0,1]} |X_n|^p d\mathbb{P} = \int_0^{1/n} \omega^{-p\alpha} d\omega = +\infty$$

et donc en particulier $(X_n)_n$ ne converge pas vers 0 dans L^p.

On peut néanmoins obtenir une réciproque partielle en introduisant la notion d'uniforme intégrabilité d'une famille de variables aléatoires.

Définition 41 (Uniforme intégrabilité, [PB07, définition V.3.3.])**.** Soit $(X_i)_{i\in I}$ une famille de variables aléatoires de L^1. On dit que $(X_i)_i$ est uniformément intégrable si

$$\lim_{c\to+\infty}\sup_{i\in I}\int_{\{|X_i|>c\}}|X_i|d\mathbb{P}=0.$$

Exemple 42 ([PB07, définition V.3.3.]). D'après l'exemple 38, une variable aléatoire intégrable, et donc une famille finie de variables aléatoires intégrables est uniformément intégrable. On montre également grâce à cet exemple que si la famille $(X_i)_{i\in I}$ est elle qu'il existe $Y\in L^1$ telle que $|X_i|\leqslant Y$ presque partout alors la famille est uniformément intégrable.

On donne maintenant un caractérisation de l'uniforme intégrabilité qui nous sera utile dans la preuve qui suit.

Proposition 43 ([PB07, proposition V.3.4.])**.** *Une famille de variables aléatoires $(X_i)_{i\in I}$ est uniformément intégrable si et seulement si les deux points suivants sont vérifiés :*

(i) $\sup_{i\in I}\mathbb{E}[|X_i|]<+\infty$

(ii) Pour tout $\varepsilon>0$, il existe $\eta>0$ tel que,

$$\forall A\in\mathcal{A},\mathbb{P}(A)\leqslant\eta\quad\Longrightarrow\quad\forall i\in I,\quad\int_A|X_i|d\mathbb{P}\leqslant\varepsilon.$$

On peut maintenant montrer une réciproque partielle de la proposition 39.

Théorème 44 ([PB07, théorème V.3.5.])**.** *Soit $(X_n)_n$ une suite de variables aléatoires de L^1, et X une variable aléatoire. Alors il y a équivalence entre :*

(i) $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ *et $(X_n)_n$ est uniformément intégrable*

(ii) X est intégrable et $||X_n-X||_1 \xrightarrow[n\to+\infty]{} 0$.

Démonstration. On suppose (i) vérifié. Le théorème 17 nous assure qu'il existe une sous-suite de $(X_n)_n$ qui converge presque sûrement vers X. Alors par le lemme de Fatou,

$$\mathbb{E}[|X|]\leqslant\liminf_k\mathbb{E}[|X_{n_k}|]\leqslant\sup_{n\in\mathbb{N}}\mathbb{E}[|X_n|]<+\infty$$

par la proposition précédente. Ainsi $X\in L^1$. D'autre part, pour $\varepsilon>0$,

$$\begin{aligned}\mathbb{E}[|X_n-X|]\leqslant\int_{\{|X_n-X|\leqslant\varepsilon\}}|X_n-X|d\mathbb{P}+\int_{\{|X_n-X|>\varepsilon\}}|X_n|d\mathbb{P}+\int_{\{|X_n-X|>\varepsilon\}}|X|d\mathbb{P}\\ \leqslant\varepsilon+\int_{\{|X_n-X|>\varepsilon\}}|X_n|d\mathbb{P}+\int_{\{|X_n-X|>\varepsilon\}}|X|d\mathbb{P}.\end{aligned}$$

Puisque $X \in L^1$, la famille $((X_n)_n, X)$ est encore uniformément intégrable et on peut donc appliquer la proposition précédente avec ε, qui nous fournit un $\eta > 0$. Par hypothèse de convergence en probabilité, à partir d'un certain rang, $\mathbb{P}(|X_n - X| > \varepsilon) \leqslant \eta$ et donc d'après la proposition précédente, à partir d'un certain rang,

$$\mathbb{E}[|X_n - X|] < 3\varepsilon$$

d'où la convergence L^1.

Réciproquement, si (ii) est vérifié, on a déjà montré la convergence en probabilité, il nous donc reste à prouver l'uniforme intégrabilité, ce que l'on va faire grâce à la proposition précédente. Soit $\varepsilon > 0$ et $N \geqslant 0$ tel que pour tout $n \geqslant N$, $||X_n - X||_1 \leqslant \varepsilon$. Puisque toutes la variables aléatoires mises en jeu sont intégrables, la famille $(X, X_0, \ldots, X_N)$ est uniformément intégrable et donc d'après la proposition précédente, il existe $\eta > 0$ tel que si $\mathbb{P}(A) \leqslant \eta$, pour $0 \leqslant n \leqslant N$,

$$\int_A |X| d\mathbb{P} \leqslant \varepsilon \quad \text{et} \quad \int_A |X_n| d\mathbb{P} \leqslant \varepsilon.$$

D'autre part pour $n \geqslant N$, par inégalité triangulaire,

$$\int_A |X_n| d\mathbb{P} \leqslant \int_A |X| d\mathbb{P} + ||X - X_n||_1 \leqslant 2\varepsilon.$$

On a donc montré le deuxième point de la caractérisation de la proposition précédente et le premier point s'obtient par inégalité triangulaire. □

On a un résultat analogue pour la convergence L^p.

Corollaire 45 ([PB07, corollaire V.3.6.]). *Soit $p > 1$ et $(X_n)_n$ une suite de variables aléatoires telles que $\sup_{n \in \mathbb{N}} \mathbb{E}[|X_n|^p] < +\infty$. Si $X_n \xrightarrow[n \to +\infty]{\mathbb{P}} X$ alors pour tout $1 \leqslant q < p$, $||X_n - X||_q \xrightarrow[n \to +\infty]{} 0$.*

3. Convergence en loi

3.1. Définition et premières propriétés.

Définition 46 ([Ouv09, définition 14.9.]). Soit $(X_n)_n$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$, définies sur des espaces probabilisés $(\Omega_n, \mathcal{A}_n, \mathbb{P}^n)$ et X une variable aléatoire à valeurs dans $\mathbb{R}^d$ définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que $(X_n)_n$ converge en loi vers X et on note $X_n \xrightarrow[n \to +\infty]{\mathcal{L}}$ si pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}^d$ continue bornée (on note $\mathcal{C}_b(\mathbb{R}^d)$ l'ensemble de telles fonctions),

$$\mathbb{E}[f(X_n)] \xrightarrow[n \to +\infty]{} \mathbb{E}[f(X)].$$

Dans toute la suite on considérera des variables aléatoires définies sur des espaces de probabilités à possiblement différents, comme dans la définition. On obtient immédiatement la proposition suivante.

Proposition 47 ([Ouv09, proposition 14.12.]). *Si $f : \mathbb{R}^d \to \mathbb{R}^k$ est une fonction continue et que la suite $(X_n)_n$ converge en loi vers X, alors $(f(X_n))_n$ converge en loi vers $f(X)$.*

On a le théorème suivant, très utile pour les preuves qui vont suivre.

Théorème 48 (Théorème du porte manteau, [Ouv09, proposition 14.5.]). *Si $(X_n)_n, X$ sont des variables aléatoires, on a l'équivalence entre*

(i) $(X_n)_n$ *converge en loi vers* X

(ii) Pour tout fermé F, $\limsup_{n\to+\infty} \mathbb{P}_{X_n}(F) \leqslant \mathbb{P}_X(F)$

(iii) Pour tout ouvert O, $\liminf_{n\to+\infty} \mathbb{P}_{X_n}(O) \geqslant \mathbb{P}_X(O)$

(iv) Pour tout borélien A *tel que* $\mathbb{P}(\partial A) = 0$, $\lim_{n\to+\infty} \mathbb{P}_{X_n}(A) = \mathbb{P}_X(A)$.

Proposition 49 ([Ouv09, proposition 14.17.]). *Soit $(X_n)_n, X$ des variables aléatoires de fonctions de répartition $(F_{X_n})_n, F_X$. Alors la suite $(X_n)_n$ converge en loi vers X si et seulement si $(F_{X_n}(x))_n$ converge vers $F_X(x)$ pour tout x point de continuité de F_X.*

Exemple 50. Soit $(m_n)_n$ et $(\sigma_n)_n$ deux suites de réelles convergeant respectivement vers m et σ. Si $(X_n)_n$ est une suite de variables aléatoires de loi $\mathcal{N}(m_n, \sigma_n^2)$ et X est une variable aléatoire de loi $\mathcal{N}(m, \sigma^2)$, alors $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$.

On donne maintenant deux critères de convergence en loi pour les variables aléatoires discrètes.

Proposition 51 ([Ouv09, proposition 14.16.]). *Soit $(X_n)_n, X$ des variables aléatoires à valeurs dans $\mathbb{Z}$. On a l'équivalence*

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X \iff \forall k \in \mathbb{Z}, \quad \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k).$$

On en déduit la proposition suivante.

Proposition 52 ([Fel68, theorem XI.6.]). *Soit $(X_n)_n, X$ des variables aléatoires à valeurs dans $\mathbb{Z}$ et $(G_n)_n, G$ leurs séries génératrice. Alors on a l'équivalence*

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X \iff \forall t \in]0,1[, \quad \lim_{n\to+\infty} G_n(t) = G(t).$$

3.2. **Convergence en loi et fonction caractéristique.**

Théorème 53 (Lévy, [Ouv09, théorème 14.11.]). *Soit $(X_n)_n$ une suite de variables aléatoires de fonctions caractéristiques φ_n. On a,*

(i) Si $(X_n)_n$ converge en loi vers X, alors (φ_n) converge uniformément sur tout compact vers φ_X.

(ii) Si (φ_n) converge simplement vers une fonction φ continue en 0, alors il existe une variable aléatoire X telle que φ soit la fonction caractéristique de X et $(X_n)_n$ converge en loi vers X.

Ce théorème est très puissant et l'on va en donner deux applications.

Théorème 54 (Poisson, [Ouv09, théorème 14.19.]). *Soit pour $n \in \mathbb{N}^*$ X_n une variable aléatoire de loi binomiale $\mathcal{B}(n, p_n)$. Si $\lim_{n\to+\infty} np_n = \lambda > 0$, alors $(X_n)_n$ converge en loi vers une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$.*

On énonce et démontre maintenant le théorème central limite, qui est un raffinement de la loi des grands nombres. On commence par donner un lemme utilisé dans la preuve. Le lemme, le théorème central limite ainsi que l'application qui suit constitue le second développement.

Lemme 55 ([HQ02, exercice XIII.7.]). *Soit $(z_n)_n$ une suite de nombres complexes de limite $z \in \mathbb{C}$, alors*

$$\left(1 + \frac{z_n}{n}\right)^n \underset{n\to+\infty}{\longrightarrow} e^z.$$

Démonstration. Pour $n \geqslant 0$, on a,

$$\exp(z_n) - \left(1 + \frac{z_n}{n}\right)^n = \sum_{k=0}^{+\infty} \frac{z_n^k}{k!} - \sum_{k=0}^{n} \binom{n}{k} \left(\frac{z_n}{n}\right)^k = \sum_{k=0}^{+\infty} a_k^{(n)} z_n^k$$

où

$$a_k^{(n)} = \begin{cases} \frac{1}{k!} & k \geqslant n+1 \\ \frac{1}{k!}\left(1 - \frac{n(n-1)\cdots(n-k+1)}{n^k}\right) & k \leqslant n \end{cases}$$

ainsi $a_k^{(n)} \geqslant 0$ pour tout n, k, et donc,

$$\left|\exp(z_n) - \left(1 + \frac{z_n}{n}\right)^n\right| \leqslant \sum_{k=0}^{+\infty} a_k^{(n)} |z_n|^k = \exp(|z_n|) - \left(1 + \frac{|z_n|}{n}\right)^n.$$

Enfin, avec pour $x \geqslant 0$, $\ln(1+x) \geqslant x - \frac{x^2}{2}$ et $1 - e^{-x} \leqslant x$,

$$\begin{aligned} \left|\exp(z_n) - \left(1 + \frac{z_n}{n}\right)^n\right| &\leqslant \exp(|z_n|) - \exp\left(n \ln\left(1 + \frac{|z_n|}{n}\right)\right) \\ &\leqslant \exp(|z_n|) - \exp\left(n\left(\frac{|z_n|}{n} - \frac{|z_n|^2}{2n^2}\right)\right) \\ &= \exp(|z_n|)\left(1 - \exp\left(-\frac{|z_n|^2}{2n}\right)\right) \\ &\leqslant \exp(|z_n|)\frac{|z_n|^2}{2n}. \end{aligned}$$

On peut alors conclure :

$$\begin{aligned} \left|\exp(z) - \left(1 + \frac{z_n}{n}\right)^n\right| &\leqslant |\exp(z) - \exp(z_n)| + \left|\exp(z_n) - \left(1 + \frac{z_n}{n}\right)^n\right| \\ &\leqslant |\exp(z) - \exp(z_n)| + \exp(|z_n|)\frac{|z_n|^2}{2n} \underset{n\to+\infty}{\longrightarrow} 0. \end{aligned}$$

□

Théorème 56 (Central limite, [HQ02, théorème XIII.II.22.]). *Soit $(X_n)_n$ une suite de variables aléatoires indépendantes et identiquement distribuées admettant un moment d'ordre 2. On note $m = \mathbb{E}[X_1]$ et $\sigma^2 = \operatorname{Var}(X_1)$. Alors si $S_n = \sum_{k=1}^n X_n$,*

$$\frac{S_n - nm}{\sqrt{n}\sigma} \xrightarrow{\mathcal{L}} X$$

où $X \sim \mathcal{N}(0,1)$.

Démonstration. On peut sans perte de généralité se ramener au cas $m = 0$ et $\sigma = 1$. On utilise le théorème de Lévy et on cherche donc à montrer que $\varphi_{\frac{S_n}{\sqrt{n}}}(t) \to e^{-t^2/2}$ pour tout $t \in \mathbb{R}$.

On note $\varphi := \varphi_{X_1}$. Puisque X_1 admet un moment d'ordre 2, φ est de classe $\mathcal{C}^2$ et vérifie : $\varphi'(0) = \mathbb{E}[iX_1] = 0$ et $\varphi''(0) = [-X^2] = -1$. On a donc, par développement de Taylor à l'ordre 2 en 0, pour tout $t \in \mathbb{R}$,

$$\varphi(t) = 1 - \frac{t^2}{2} + o(t^2).$$

On a d'autre part, par indépendance et identique distribution,

$$\begin{aligned} \varphi_{\frac{S_n}{\sqrt{n}}}(t) &= \mathbb{E}\left[\prod_{k=1}^n e^{itX_k/\sqrt{n}}\right] = \prod_{k=1}^n \mathbb{E}\left[e^{itX_k/\sqrt{n}}\right] \\ &= \mathbb{E}\left[e^{itX_1/\sqrt{n}}\right]^n = \varphi\left(\frac{t}{\sqrt{n}}\right)^n. \end{aligned}$$

Ainsi, d'après le développement limité de φ en 0,

$$\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + o(\frac{1}{n})\right)^n.$$

et on conclut avec le lemme. □

Application 57. *On a,*

$$e^{-n} \sum_{k=0}^{+\infty} \frac{n^k}{k!} = \frac{1}{2}.$$

Démonstration. Soit (X_i) une suite de variables aléatoires i.i.d. de loi de Poisson $\mathcal{P}(1)$, et $S_n = X_1 + \cdots + X_n$ qui suit donc une loi de Poisson $\mathcal{P}(n)$. On a $\mathbb{E}[X_1] = \operatorname{Var}(X_1) = 1$ et on applique le théorème central limite pour obtenir :

$$Z_n = \frac{S_n - n}{\sqrt{n}} \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0,1).$$

Or on remarque que

$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\sum_{k=0}^{n}\mathbb{P}(S_n=k)=\mathbb{P}(S_n\leqslant n)=\mathbb{P}(Z_n\leqslant 0)$$

et donc par convergence des fonctions de répartitions on obtient

$$e^{-n}\sum_{k=0}^{+\infty}\frac{n^k}{k!}=\mathbb{P}(Z\leqslant 0)=\frac{1}{2}$$

puisque Z est symétrique. □

Application 58 ([PB07, exemple V.5.5.])**.** *Si $(X_n)_n$ admet pour loi commune une loi de Bernoulli de paramètre p, alors on sait que S_n la moyenne arithmétique des $(X_i)_{1\leqslant i\leqslant n}$ suit une loi binomiale $\mathcal{B}(n,p)$. Le théorème central limite nous donne, pour $a<b$,*

$$\lim_{n\to+\infty}\mathbb{P}\Big(a\leqslant\frac{S_n-np}{\sqrt{np(1-p)}}\leqslant b\Big)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt.$$

Remarque 59 ([PB07, exemple V.5.5.])*.* On reprend les notations de l'exemple précédent. Nous verrons que la convergence en probabilité implique la convergence en loi, ainsi, d'après la loi des grands nombres, S_n/n converge en loi vers p. D'après l'exemple précédent, on a alors

$$\lim_{n\to+\infty}\mathbb{P}\Big(a\leqslant\frac{S_n-np}{\sqrt{n\frac{S_n}{n}\left(1-\frac{S_n}{n}\right)}}\leqslant b\Big)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt$$

et donc pour n assez grand, l'intervalle

$$\left[\frac{S_n}{n}-\frac{b}{\sqrt{n}}\sqrt{\frac{S_n}{n}\Big(1-\frac{S_n}{n}\Big)},\frac{S_n}{n}-\frac{a}{\sqrt{n}}\sqrt{\frac{S_n}{n}\Big(1-\frac{S_n}{n}\Big)}\right]$$

contient p avec une probabilité proche de $\frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt$. En statistiques, la paramètre p est souvent inconnu et ce résultat permet donc d'encadrer sa valeur en observant des réalisation de X_i.

3.3. **Liens avec les autres notions de convergence.**

Proposition 60 ([Ouv09, proposition 14.13.])**.** *Soit $(X_n)_n$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$ à valeurs dans $\mathbb{R}^d$ et qui converge en probabilité vers X (définie sur $(\Omega,\mathcal{A},\mathbb{P})$ et à valeurs dans $\mathbb{R}^d$). Alors $(X_n)_n$ converge en loi vers X.*

Remarque 61*.* Puisque la convergence presque sûre implique la convergence en probabilité, on obtient immédiatement avec la proposition précédente que la convergence presque sûre implique la convergence en loi.

La réciproque de la proposition est fausse, comme le montre l'exemple suivant, très simple puisque la suite considérée est constante.

Contre-exemple 62 ([Ouv09, proposition 14.13.]). Soit X de loi de Bernoulli $\mathcal{B}(p)$ et $(X_n)_n$ la suite constante égale à X. Si on pose $Y = 1 - X$, Y est de même loi que X et donc $(X_n)_n$ converge en loi vers Y. Cependant $|X - Y| = |2X - 1| = 1$ et donc pour $\varepsilon \in]0, 1[$, $\mathbb{P}(|X_n - Y| > \varepsilon) = 1$ et donc $(X_n)_n$ ne converge pas en probabilité vers Y.

On a cependant une réciproque partielle.

Proposition 63 ([Ouv09, proposition 14.14.]). *Soit $(X_n)_n$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$ et qui converge en loi vers une variable aléatoire a $\mathbb{P}$-p.s. constante. Alors $(X_n)_n$ converge en probabilité vers a.*

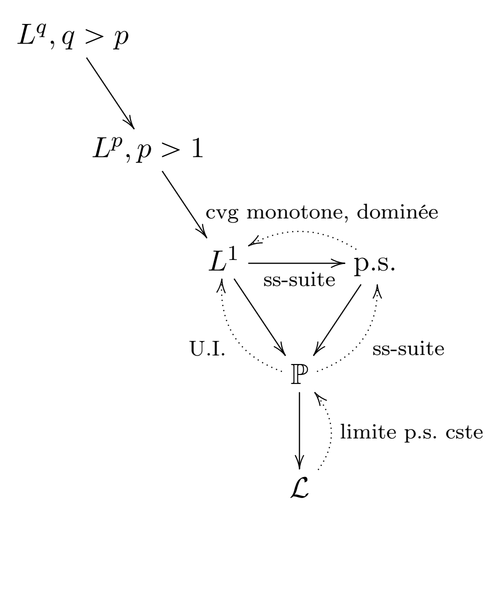
Contrairement aux autres notions de convergence, la convergence en loi d'un vecteur n'est pas assurée par la convergence en loi de chacune de ses composantes. Cependant la proposition précédente nous permet d'avoir un résultat de ce type sous certaines hypothèses.

Proposition 64 (lemme de Slutsky, [Ouv09, exercice 14.8.]). *Soit $(X_n)_n$ et $(Y_n)_n$ deux suites de variables aléatoires réelles définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et qui convergent en loi respectivement vers une variable aléatoire X et une constante y_0. Alors $(X_n, Y_n)_n$ converge en loi vers (X, y_0).*

En particulier, $X_n + Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X + y_0$.

4. Annexe

On résume ici les liens entre les différentes notions de convergence, une flèche de A à B indiquant qu'une convergence au sens de A implique une convergence au sens de B.



5. Réponses aux questions du jury

Question 1. *Concernant le théorème 9 sur la divergence presque sûre d'une marche aléatoire symétrique sur $\mathbb{Z}^d$, que peut-on dire si l'on considère une marche aléatoire symétrique ($\mathbb{P}(X = x) = \mathbb{P}(X = -x)$) sur $\mathbb{Z}^d$ mais pas forcément à valeurs dans la base canonique ?*

Réponse. Dans la preuve donnée on montre la divergence en prouvant la sommabilité de la série $\sum \mathbb{P}(S_n = 0)$. Pour cela on se sert d'un équivalent en 0 de φ la fonction caractéristique des variables aléatoires considérées. Or dans le cas à valeurs dans la base canonique, cette fonction φ est une somme de sinus et est donc facile à développer asymptotiquement et le développement obtenu permet de conclure. Maintenant si X n'est plus à valeurs dans la base canonique l'on perd cette forme de φ et donc la possibilité de conclure de la même manière. □

Question 2. *En quoi le théorème central limite est-il un raffinement de la loi des grands nombres ? Peut-on itérer le processus ?*

Réponse. La loi des grands nombres nous dit que $\frac{S_n}{n}$ converge vers m, mais on veut savoir à quelle vitesse. On soustrait alors m, et dans la même idée qu'un équivalent, on divise par une quantité qui nous semble être la vitesse de convergence de $\frac{S_n}{n}$ vers m :

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - m\right) = \frac{S_n - nm}{\sqrt{n}\sigma} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Il n'existe pas d'itération de ce processus. □

Question 3. *Donner un contre-exemple au lemme de Slutsky.*

Réponse. On considère X qui suit une loi de Rademacher de paramètre $1/2$ et $(X_n)_n, (Y_n)_n$ les suites constantes égales respectivement à X et $-X$. Alors puisque X est symétrique, $X = -X$ en loi et donc $(X_n)_n$ et $(Y_n)_n$ convergent en loi vers X. Mais $X_n + Y_n = 0$ ne converge pas en loi vers $2X$. □

Question 4. *Montrer que si une suite $(X_n)_n$ de variables aléatoires de lois $\mathcal{N}(m_n, \sigma_n^2)$ converge en loi vers X, alors $X \sim \mathcal{N}(m, \sigma^2)$ où m, σ sont les limites des suites $(m_n)_n, (\sigma_n)_n$ quitte à extraire.*

Réponse. On a déjà montré dans l'exemple 50 que si les suites $(m_n)_n, (\sigma_n)_n$ convergent vers m, σ alors la suite $(X_n)_n$ converge en loi vers une variable aléatoire de loi $\mathcal{N}(m, \sigma^2)$. On va utiliser ce résultat ici.

On suppose alors la convergence en loi, et on montre que quitte à extraire, les suites $(m_n)_n, (\sigma_n)_n$ convergent ce qui nous permettra de conclure d'après la remarque préliminaire. On a, d'après le théorème de Lévy, pour tout $t \in \mathbb{R}$,

$$e^{itm_n - \sigma_n^2 t^2/2} \underset{n\to+\infty}{\longrightarrow} \varphi_X(t).$$

On peut alors facilement en déduire la convergence de la suite $(\sigma_n)_n$: on sait que

$$|e^{itm_n-\sigma_n^2t^2/2}| = e^{-\sigma_n^2t^2/2} \underset{n\to+\infty}{\longrightarrow} |\varphi_X(t)|$$

et donc

$$\sigma_n = \sqrt{\frac{-2\log|\varphi_{X_n}(t)|}{t^2}} \underset{n\to+\infty}{\longrightarrow} \sqrt{\frac{-2\log|\varphi_X(t)|}{t^2}} =: \sigma$$

pour t dans un voisinage épointé de 0.

On cherche alors à montrer que $(m_n)_n$ converge. On va montrer pour cela que la suite est bornée. En effet, si c'est le cas, on pourra extraire une sous-suite convergeant vers m qui sera alors uniquement déterminé par $e^{itm} = \varphi_X(t)e^{\sigma^2t^2/2}$ pour tout $t \in \mathbb{R}$.

On commence par montrer que $(X_n)_n$ vérifie

$$\forall\varepsilon > 0, \exists M > 0, \quad \sup_{n\in\mathbb{N}} \mathbb{P}(|X_n| > M) < \varepsilon.$$

On considérant la suite $(|X_n|)_n$ qui converge en loi vers $|X|$. Soit alors $\varepsilon > 0$. On sait que la suite $(F_n)_n$ des fonctions de répartitions des $(|X_n|)_n$ converge vers $F := F_{|X|}$ en tout point de continuité de cette dernière. Soit donc $t \in \mathbb{R}$ un point de continuité de F tel que $F(t) > 1 - \varepsilon$ (qui existe puisque les points de discontinuité de F sont au plus dénombrable), et soit n_0 tel que pour tout $n \geqslant n_0$, $F_n(t) > 1 - \varepsilon$, alors pour tout $n \geqslant n_0$, $\mathbb{P}(|X_n| > t) = 1 - F_n(t) < \varepsilon$. D'autre part, pour $0 \leqslant n \leqslant n_0 - 1$, il existe M_n tel que $\mathbb{P}(|X_n| > M_n) < \varepsilon$, on peut alors conclure en posant $M = \max(M_n, t)$.

On suppose enfin par l'absurde que $(m_n)_n$ n'est pas bornée. On sait que pour tout n, $\mathbb{P}(|X_n| > |m_n|) \geqslant 1/2$. Par hypothèse, pour tout $M > 0$, il existe $n \in \mathbb{N}$ tel que $|m_n| > M$, pour un tel n, on a alors $\mathbb{P}(|X_n| > M) \geqslant 1/2$, et donc

$$\sup_{n\in\mathbb{N}} \mathbb{P}(|X_n| > M) \geqslant 1/2$$

ce qui contredit la propriété que l'on a montré plus haut, et conclut la preuve □

Question 5. *Que peut-on dire de la méthode de Monte-Carlo ?*

Réponse. La méthode de Monte-Carlo permet d'approcher une intégrale mais converge plus lentement que des méthodes classiques d'analyse numérique, cependant elle ne nécessite aucune hypothèse de régularité sur la fonction. On peut préciser la vitesse de convergence sous une hypothèse d'intégrabilité de f : si $\int_0^1 |f(t)|^2 dt \leqslant C$, alors par l'inégalité de Tchebychev,

$$\mathbb{P}(|S_n - I| > \varepsilon) \leqslant \frac{\mathrm{Var}(S_n)}{\varepsilon^2} \leqslant \frac{C}{n\varepsilon^2}$$

en notant $I = \int_0^1 f(t)dt$ et $S_n = \sum_{k=1}^n X_n$ en reprenant les notations de l'application 23. On voit de plus, grâce à l'inégalité de Markov, que plus la fonction f est intégrable à une puissance élevée et plus la méthode converge vite. □

Références

[Fel68] W. Feller. *An introduction to probability theory and its applications*, volume Volume 1. Wiley, 3 edition, 1968.

[HQ02] Claude Zuily Hervé Queffélec. *Analyse pour l'agrégation*. Dunod, 2 edition, 2002.

[MB06] Gilles Pagès Marc Briane. *Theéorie de l'inteégration*. Vuibert, 4ed. edition, 2006.

[Ouv09] Jean-Yves Ouvrard. *Probabilités Tome 2*. Cassini, 2009.

[PB07] M. Ledoux Ph. Barbe. *Probabilité*. EDP SCIENCES, 2007.