TD 12 – Lemmes de Borel–Cantelli et convergences

1 – Petites questions

Soit $X, X_0, X_1, \ldots$ des variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $\mu, \mu_0, \mu_1, \ldots$ des mesures finies sur $\mathbb{R}$.

1. Supposons que $(X_n)_{n\geq 0}$ converge en loi vers X. Soit $f : \mathbb{R} \to \mathbb{R}$ continue. A-t-on $f(X_n) \to f(X)$ en loi ?
2. Supposons que pour tout $f : \mathbb{R} \to \mathbb{R}$ continue à support compact, $\int f \, d\mu_n \to \int f \, d\mu$. Y a-t-il convergence étroite de $(\mu_n)_{n\geq 0}$ vers μ ?
3. Supposons que pour tout $f : \mathbb{R} \to \mathbb{R}$ continue à support compact, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$. Y a-t-il convergence en loi de $(X_n)_{n\geq 0}$ vers X ?
4. Supposons que $(X_n)_{n\geq 0}$ converge en loi vers X. A-t-on $\mathbb{E}[X_n] \to \mathbb{E}[X]$?

Corrigé.

1. Vrai : si $g : \mathbb{R} \to \mathbb{R}$ est continue bornée, alors $g \circ f$ est continue bornée et donc $\mathbb{E}[g(f(X_n))] \to \mathbb{E}[g(f(X))]$.
2. Faux : prendre $\mu_n = \delta_n$.
3. Vrai : cela a été vu dans le cours (en terme de mesures) : si une suite de mesures de probabilité $(\mu_n)_{n\in\mathbb{N}}$ converge vaguement vers une mesure de probabilité μ, alors $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement vers μ
   Remontrons le, en travaillant dans le cadre plus général d'un espace polonais E (métrique complet séparable). Soit $\varepsilon > 0$. On sait que la loi de X est tendue, c'est-à-dire il existe K compact tel que $\mathbb{P}(X \in K) \geq 1 - \varepsilon$ (d'après le cours toute mesure borélienne finie sur un polonais est tendue). Soit $g : E \to \mathbb{R}$ continue à support compact telle que $\mathbb{1}_K \leq g \leq 1$. Soit $f : E \to \mathbb{R}$ continue bornée. On a
$$\begin{aligned} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\leq |\mathbb{E}[fg(X_n)] - \mathbb{E}[fg(X)]| + |\mathbb{E}[f(1-g)(X_n)] - \mathbb{E}[f(1-g)(X)]| \\ &\leq |\mathbb{E}[fg(X_n)] - \mathbb{E}[fg(X)]| + \|f\|_\infty (\mathbb{E}[(1-g)(X_n)] + \mathbb{E}[(1-g)(X)]), \end{aligned}$$
   car $1 - g \geq 0$. On a $\mathbb{E}[(1-g)(X_n)] = 1 - \mathbb{E}[g(X_n)]$ et, en appliquant l'hypothèse à fg et g qui sont continues à support compact, on obtient
$$\limsup_{n\to\infty} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq 0 + \|f\|_\infty (1 - \mathbb{E}[g(X)] + \mathbb{E}[(1-g)(X)]) \leq 2\varepsilon \|f\|_\infty,$$
   car $\mathbb{E}[(1-g)(X)] \leq \mathbb{P}(X \notin K) \geq \varepsilon$. Cela conclut avec $\varepsilon > 0$.
4. Faux : on prend $(\Omega, \mathcal{A}, \mathbb{P}) = ([0,1], \mathcal{B}(\mathbb{R}), \lambda)$ et $X_n(t)$ la fonction tente telle que $X_n(0) = 0, X_n(1/n) = n, X_n(2/n) = 0$. Alors X_n converge p.s. vers 0 mais $\mathbb{E}[X_n] = 1 \neq \mathbb{E}[0]$.
   Un autre exemple davantage probabiliste : soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi telles que $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$. On note $Z_n = X_1 + X_2 + \cdots + X_n$ et soit $T = \inf\{n \geq 1 ; Z_n = 1\}$. On pose finalement :
$$W_n = Z_{\min(n,T)}.$$

Pour toute question, n'hésitez pas à m'envoyer un mail à michel.pain@ens.fr, ou bien à venir me voir au bureau V2.

Ainsi, (W_n) est la marche aléatoire issue de 0 qui fait des sauts ± 1 qui reste en 1 une fois qu'elle l'a atteint. Il est possible de vérifier que $T < \infty$ p.s. de sorte que (W_n) converge presque sûrement vers 1. Or il est facile de vérifier que pour tout $n \geq 1$, $\mathbb{E}[W_n] = 0$, de sorte que $\mathbb{E}[W_n]$ ne converge pas vers $\mathbb{E}[1]$. Avec le langage du second semestre, cela fournit l'exemple d'une martingale qui converge p.s. mais pas dans L^1.

2 – Lemmes de Borel–Cantelli et loi du 0–1 de Kolmogorov

Exercice 1. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ positives, indépendantes et de même loi.

1. Montrer que pour tout $\alpha > 0$ on a l'équivalence suivante :
$$\mathbb{E}[X_0] < \infty \Leftrightarrow \sum_{n\geq 0} \mathbb{P}(X_0 \geq \alpha n) < \infty.$$

2. En déduire la dichotomie suivante :
$$\limsup_{n\to\infty} \frac{X_n}{n} = \begin{cases} 0 & \text{si } \mathbb{E}[X_1] < \infty, \\ \infty & \text{si } \mathbb{E}[X_1] = \infty, \end{cases}$$
presque sûrement.

Corrigé.

1. Soit X_0 une v.a. positive et $\alpha > 0$, il est facile de vérifier les encadrements de X_0 suivants :
$$\sum_{n\geq 0} \alpha n \mathbb{1}_{\alpha n \leq X_0 < \alpha(n+1)} \leq X_0 \leq \sum_{n\geq 0} \alpha(n+1) \mathbb{1}_{\alpha n \leq X_0 < \alpha(n+1)}.$$
En prenant l'espérance de la ligne précédente on obtient
$$\alpha \sum_{n\geq 0} \mathbb{P}(X_0 \geq \alpha n) \leq \mathbb{E}[X_0] < \alpha \sum_{n\geq 0} \mathbb{P}(X_0 \geq \alpha n) + \alpha,$$
dont il est facile de déduire que
$$\mathbb{E}[X_0] < \infty \Leftrightarrow \sum_{n\geq 0} \mathbb{P}(X_0 \geq \alpha n) < \infty.$$

2. Soit $\alpha > 0$, $\limsup \frac{X_n}{n} \geq \alpha$ ssi il existe une infinité de n tels que $X_n \geq \alpha n$. D'après la question précédente,
$$\sum_{n\geq 0} \mathbb{P}(X_n \geq \alpha n) \begin{cases} < \infty & \text{si } \mathbb{E}[X_1] < \infty \\ = \infty & \text{si } \mathbb{E}[X_1] = \infty \end{cases},$$
donc, par Borel–Cantelli (en utilisant dans le cas de la série divergente que les événements $\{X_n \geq \alpha n\}$ sont indépendants),
$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{n} \geq \alpha\right) = \begin{cases} 0 & \text{si } \mathbb{E}[X_1] < \infty \\ 1 & \text{si } \mathbb{E}[X_1] = \infty \end{cases}$$
et par conséquent
$$\limsup_{n\to\infty} \frac{X_n}{n} = \begin{cases} 0 & \text{si } \mathbb{E}[X_1] < \infty, \\ \infty & \text{si } \mathbb{E}[X_1] = \infty, \end{cases}$$
presque sûrement.

Exercice 2. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes et de loi de Bernoulli de paramètre $1/2$. Pour $n \geq 1$, on définit

$$L_n := \max\{k \geq 1 \mid \exists i \leq n-k : X_{i+1} = \cdots = X_{i+k} = 1\},$$

la longueur de la plus grande série de 1 successifs obtenue avant l'instant n. Montrer que

$$\frac{L_n}{\ln n} \xrightarrow[n\to\infty]{} \frac{1}{\ln 2},$$

presque sûrement.

Corrigé. Soit $\varepsilon > 0$.
Limite supérieure. Pour $j \geq 1$, on a

$$\mathbb{P}(L_n \geq j) = \mathbb{P}\left(\bigcup_{i=0}^{n-j}\{X_{i+1} = \cdots = X_{i+j} = 1\}\right) \leq \sum_{i=0}^{n-j}\mathbb{P}(\{X_{i+1} = \cdots = X_{i+j} = 1\}) = \sum_{i=0}^{n-j}\frac{1}{2^j} = \frac{n-j+1}{2^j} \leq \frac{n}{2^j}.$$

Soit $j_n := \lfloor (1+\varepsilon)\ln(n)/\ln(2)\rfloor$. Alors $\mathbb{P}(L_n \geq j_n) \leq 2/n^{\varepsilon}$. Posons $n = n_k = \lfloor k^{2/\varepsilon}\rfloor$ de sorte que $\sum_k \mathbb{P}(L_{n_k} \geq j_{n_k}) < \infty$. D'après le lemme de Borel–Cantelli, p.s., pour tout k suffisamment grand on a

$$L_{n_k} < j_{n_k} \leq (1+\varepsilon)\frac{\ln(n_k)}{\ln(2)}.$$

Pour $n \in [n_{k-1}, n_k[$ avec k suffisamment grand, on a

$$L_n \leq L_{n_k} < (1+\varepsilon)\frac{\ln(n_k)}{\ln(2)} \leq (1+2\varepsilon)\frac{\ln(n_{k-1})}{\ln(2)} \leq (1+2\varepsilon)\frac{\ln(n)}{\ln(2)},$$

donc on a presque sûrement

$$\limsup_{n\to\infty}\frac{L_n}{\ln n} \leq \frac{1+2\varepsilon}{\ln 2}.$$

Limite inférieure. Posons $k_n := \lfloor (1-\varepsilon)\ln(n)/\ln(2)\rfloor$ et $N_n := \lfloor n/k_n\rfloor$. Pour $1 \leq i \leq N_n$, on définit

$$A_i := \{X_{(i-1)k_n+1} = X_{(i-1)k_n+2} = \cdots = X_{ik_n} = 1\}.$$

Alors $\bigcup_{i=1}^{N_n} A_i \subset \{L_n \geq k_n\}$. Puisque les événements $A_i, 1 \leq i \leq N_n$ sont indépendants, ceci entraîne que

$$\mathbb{P}(L_n < k_n) \leq \mathbb{P}\left(\bigcap_{i=1}^{N_n} A_i^c\right) = \prod_{i=1}^{N_n}\mathbb{P}(A_i^c) = \mathbb{P}(A_1^c)^{N_n} = \left(1 - \frac{1}{2^{k_n}}\right)^{N_n} = \exp\left(\left\lfloor\frac{n}{k_n}\right\rfloor \ln\left(1 - \frac{1}{2^{k_n}}\right)\right),$$

qui est sommable en n, car

$$\left\lfloor\frac{n}{k_n}\right\rfloor \ln\left(1 - \frac{1}{2^{k_n}}\right) \sim -\frac{n\ln(2)}{(1-\varepsilon)\ln(n)}\frac{1}{2^{k_n}} \leq -\frac{n\ln(2)}{(1-\varepsilon)\ln(n)}\frac{1}{2n^{1-\varepsilon}} = -\frac{\ln(2)}{2(1-\varepsilon)}\frac{n^{\varepsilon}}{\ln(n)}$$

D'après le lemme de Borel–Cantelli, p.s., pour tout n suffisamment grand,

$$L_n \geq k_n \geq (1-\varepsilon)\frac{\ln(n)}{\ln(2)} - 1,$$

donc on a presque sûrement

$$\liminf_{n\to\infty}\frac{L_n}{\ln n} \geq \frac{1-\varepsilon}{\ln 2}.$$

Conclusion. En faisant tendre $\varepsilon \to 0$ on a : presque sûrement $\lim_{n\to\infty} L_n/\ln(n) = 1/\ln(2)$.

Exercice 3. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes.

1. Montrer que la série $\sum_{n\geq 0} X_n$ converge presque sûrement ou diverge presque sûrement.
2. Supposons à présent que les variables aléatoires X_n sont positives et identiquement distribuées. Montrer que $\sum_{n\geq 0} X_n = \infty$ presque sûrement, sauf dans un cas à préciser.

Corrigé.

1. Notons $\mathcal{F}_N = \sigma(X_N, X_{N+1}, \ldots)$. La série $\sum_{n\geq 0} X_n$ converge si et seulement si la série $\sum_{n\geq N} X_n$ converge pour tout $N \geq 0$. Or

$$\left\{\omega \in \Omega : \sum_{n\geq N} X_n(\omega) \text{ converge}\right\} \in \mathcal{F}_N. \tag{1}$$

On a donc

$$\left\{\omega \in \Omega : \sum_{n\geq 0} X_n(\omega) \text{ converge}\right\} \in \bigcap_{N\geq 0} \mathcal{F}_N.$$

On conclut en utilisant la loi du 0–1 de Kolmogorov.

Remarque. Dans un but pédagogique, expliquons d'où vient (1) en détail. D'après le critère de Cauchy,

$$\left\{\omega \in \Omega : \sum_{n\geq N} X_n(\omega) \text{ converge}\right\} = \bigcap_{k\geq 1} \bigcup_{k\geq N} \bigcap_{p,q\geq k} \left\{\omega \in \Omega : \left|\sum_{n=p}^{q} X_n(\omega)\right| < \frac{1}{k}\right\}.$$

Or, pour $p, q \geq N$,

$$\left\{\omega \in \Omega : \left|\sum_{n=p}^{q} X_n(\omega)\right| < \frac{1}{k}\right\} \in \sigma(X_p, X_{p+1}, \ldots, X_q) \subset \mathcal{F}_N,$$

ce qui établit (1) car $\mathcal{F}_N$ est une tribu.

2. Tout d'abord, si X_0 est presque sûrement nulle, alors $\sum_n X_n = 0$ p.s. Supposons que X_0 n'est pas nulle, alors on peut trouver $\varepsilon > 0$ tel que $\mathbb{P}(X_1 > \varepsilon) > 0$. On a trivialement que $\sum_n \mathbb{P}(X_n > \varepsilon) = \infty$, et les X_n sont indépendants, donc par Borel–Cantelli, p.s. les X_n sont supérieurs à ε une infinité de fois, donc $\sum_n X_n = \infty$.

3 – Loi des grands nombres

Exercice 4. (*Théorème de Bernstein-Weierstrass*) Soit $f : [0,1] \to \mathbb{C}$ continue. Le n^{e} *polynôme de Bernstein* de f est défini par

$$B_n(x) := \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right), \quad x \in \mathbb{R}.$$

1. Montrer sans calcul que B_n converge simplement vers f.
2. Montrer que B_n converge uniformément vers f.

Corrigé.

1. On remarque que

$$B_n(x) = \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right],$$

avec S_n qui suit une loi binomiale de paramètre (n, x). On peut prendre $S_n = X_1 + \cdots + X_n$ avec $(X_k)_{k\geq 1}$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre p. Alors, d'après la loi forte des grands nombres, S_n/n converge p.s. vers $\mathbb{E}[X_1] = x$. Donc, par continuité de f et par convergence dominée, $B_n(x) \to f(x)$.

2. Soit $\varepsilon > 0$, soit $\eta > 0$ le module d'uniforme continuité de f associé à ε. Alors on a

$$\begin{aligned}|f(x) - B_n(x)| &= \mathbb{E}[|f(x) - f(S_n(x))|] \\ &\leq \mathbb{E}[|f(x) - f(S_n(x))|\mathbf{1}_{|S_n(x)-x|\leq\eta}] + \mathbb{E}[|f(x) - f(S_n(x))|\mathbf{1}_{|S_n(x)-x|\geq\eta}] \\ &\leq \varepsilon + 2\|f\|_\infty \mathbb{P}(|S_n(x) - x| \geq \eta)\end{aligned}$$

Pour évaluer $\mathbb{P}(|S_n(x) - x| \geq \eta)$, on utilise l'inégalité de Markov :

$$\mathbb{P}(|S_n(x) - x| \geq \eta) \leq \frac{\mathrm{Var}(|S_n(x)|)}{\eta^2} = \frac{x(1-x)}{n\eta^2} \leq \frac{1}{2n\eta^2}.$$

La majoration ci-dessus est uniforme en x, ce qui permet de conclure.

4 – Convergence en loi

Exercice 5. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes et de loi uniforme sur $[0,1]$. On pose $M_n := \max(X_1, \dots, X_n)$. Montrer que la suite $(n(1-M_n))_{n\geq 1}$ converge en loi et expliciter la loi limite.

Corrigé. Pour tout $n \geq 1$, la variable aléatoire $n(1 - M_n)$ est à valeurs dans $[0, n]$. On a donc, pour tout $t < 0$, $\mathbb{P}(n(1 - M_n) \leq t) = 0$. Soit $t \geq 0$ fixé. Pour tout $n \geq t$, on a

$$\mathbb{P}(n(1 - M_n) \leq t) = \mathbb{P}\left(M_n \geq 1 - \frac{t}{n}\right) = 1 - \mathbb{P}\left(M_n < 1 - \frac{t}{n}\right) = 1 - \left(1 - \frac{t}{n}\right)^n.$$

Donc $\mathbb{P}(n(1 - M_n) \leq t) \to (1 - \mathrm{e}^{-t})\mathbb{1}_{\{t\geq 0\}}$, et la fonction $t \mapsto (1 - \mathrm{e}^{-t})\mathbb{1}_{\{t\geq 0\}}$ est la fonction de répartition de la loi exponentielle de paramètre 1. Ainsi, la suite $(n(1-M_n))_{n\geq 1}$ converge en loi vers une variable aléatoire exponentielle de paramètre 1.

Exercice 6. Soient $(X_n)_{n\geq 1}$ une suite de variables aléatoires constantes, respectivement égales p.s. à $x_n \in \mathbb{R}$, et X une variable aléatoire réelle. Montrer que $X_n \to X$ en loi quand $n \to \infty$ si et seulement s'il existe $x \in \mathbb{R}$ tel que X est de loi δ_x et $x_n \to x$ quand $n \to \infty$.

Corrigé. Si $x_n \to x$ et si X est de loi δ_x alors pour toute fonction continue $g\colon \mathbb{R} \to \mathbb{R}$ on a $\mathbb{E}[g(X_n)] = g(x_n) \to g(x) = \mathbb{E}[g(X)]$, ce qui signifie que $X_n \to X$ en loi.
Si $X_n \to X$ en loi alors $F_{X_n}(t) \to F_X(t)$ pour tout $t \in D$ où D est l'ensemble des points de continuité de F (D est dense car le complémentaire d'un ensemble dénombrable). Pour tout $t \in \mathbb{R}$, on a $F_{X_n}(t) \in \{0,1\}$ et donc $F_X(t) \in \{0,1\}$ pour tout $t \in D$. Comme F_X est continue à droite, on a $F_X(t) \in \{0,1\}$ pour tout $t \in \mathbb{R}$ Comme F_X est croissante, il existe $x \in \mathbb{R}$ tel que $F_X = \mathbb{1}_{[x,\infty[}$. Donc X est de loi δ_x. En outre, pour $t > x$ (on a $t \in D$), on a $\mathbb{1}_{x_n \leq t} = F_{X_n}(t) \to F_X(t) = 1$. Donc $x_n \leq t$ à partir d'un certain rang, donc $\limsup_{n\to\infty} x_n \leq t$ et avec $t \downarrow x$ on obtient $\limsup_{n\to\infty} x_n \leq x$. On procède de même pour $\liminf_{n\to\infty} x_n \geq x$.

5 – Compléments (hors TD)

Exercice 7. Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, respectivement de loi exponentielle de paramètre n.

1. Montrer que Z_n converge presque sûrement vers 0 lorsque $n \to \infty$.
2. Montrer que presque sûrement, à partir d'un certain rang, $Z_n < Z_1$.
3. On suppose ici que les variables aléatoires $(Z_n)_{n\geq 1}$ sont indépendantes. Calculer $\sum_{n\geq 1} \mathbb{P}(Z_n \geq Z_1)$. Commenter.

Corrigé.

1. Soit $\varepsilon > 0$. On a $\mathbb{P}(Z_n > \varepsilon) = e^{-n\varepsilon}$. Donc
$$\sum_{n\geq 0} \mathbb{P}(Z_n > \varepsilon) < \infty.$$
D'après le lemme Borel–Cantelli, pour tout $\varepsilon > 0$, presque sûrement, à partir d'un certain rang on a $Z_n \leq \varepsilon$. Donc presque sûrement, pour entier $k \geq 1$, à partir d'un certain rang $0 \leq Z_n \leq 1/k$. On en déduit que Z_n converge presque sûrement vers 0.
2. Soit $A := \{\omega \in \Omega : Z_n(\omega) \to 0 \text{ et } Z_1 > 0\}$. Soit $\omega \in A$. Alors à partir d'un certain rang $Z_n(\omega) < Z_1(\omega)$. Comme $\mathbb{P}(A) = 1$, ceci conclut.
3. On a $\mathbb{P}(Z_n > Z_1) = 1/(n+1)$ pour $n \geq 2$. Ainsi
$$\sum_{n\geq 1} \mathbb{P}(Z_n > Z_1) = \infty.$$
Ici le lemme Borel–Cantelli (version série divergente) ne s'applique pas car les événements $\{Z_n > Z_1\}$ ne sont pas indépendants.

Exercice 8. (*Loi des grands nombres, cas non intégrable*) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes et de même loi. On pose, pour tout $n \geq 1$, $S_n := X_1 + \ldots + X_n$.

1. Montrer que
$$\mathbb{E}[|X_1|] \leq 1 + \sum_{n\geq 1} \mathbb{P}(|X_n| \geq n).$$
2. En déduire que, si X_1 n'est pas intégrable, alors la suite $(S_n/n)_{n\geq 1}$ diverge presque sûrement.

Corrigé.

1. On utilise la relation
$$\mathbb{E}[|X_1|] = \int_0^\infty \mathbb{P}(|X_1| > t)\,\mathrm{d}t,$$
pour obtenir $\mathbb{E}(|X_1|) \leq 1 + \sum_{n\geq 1} \mathbb{P}(|X_1| \geq n)$. Puis on a $\mathbb{P}(|X_1| \geq n) = \mathbb{P}(|X_n| \geq n)$.
2. Si X_1 n'est pas intégrable alors $\sum_{n\geq 1} \mathbb{P}(|X_n| \geq n) = +\infty$. Les événements $\{|X_n| \geq n\}$ étant indépendants, on a $\mathbb{P}(\limsup\{|X_n| \geq n\}) = 1$ d'après le lemme de Borel–Cantelli. On remarque que
$$\{(S_n/n)_{n\geq 1} \text{ converge dans } \mathbb{R}\} \subset \left\{\frac{S_n}{n} - \frac{S_{n+1}}{n+1} \to 0\right\} \bigcap \left\{\frac{S_n}{n(n+1)} \to 0\right\}.$$
Donc
$$\{(S_n/n)_{n\geq 1} \text{ converge dans } \mathbb{R}\} \subset \left\{\frac{X_n}{n} \to 0\right\} \subset (\limsup\{|X_n| \geq n\})^c.$$
Donc S_n/n diverge p.s.

Exercice 9. (*Mesure empirique*) Soit μ une mesure de probabilité sur $\mathbb{R}$. Soit $(X_k)_{k\geq 1}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes de loi μ. Pour $n \in \mathbb{N}$, on définit
$$\mu_n := \frac{1}{n}\sum_{k=1}^{n} \delta_{X_k},$$
qui est une mesure aléatoire appelée *mesure empirique* associée à l'échantillon $(X_1, \ldots, X_n)$. L'objectif de cet exercice est de montrer que la mesure empirique converge vers la vraie mesure, c'est-à-dire qu'à partir des observations X_k, on arrive à retrouver la loi inconnue μ.

1. Montrer que $\mathcal{C}_c(\mathbb{R})$ est séparable pour la norme infini.
2. Soit $(\nu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilité sur $\mathbb{R}$ et ν une mesure de probabilité sur $\mathbb{R}$. Soit H une partie dense de $\mathcal{C}_c(\mathbb{R})$ pour la norme infini. Montrer que ν_n converge étroitement vers ν si et seulement si
$$\forall h \in H, \qquad \int_{\mathbb{R}} h \,\mathrm{d}\nu_n \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} h \,\mathrm{d}\nu.$$
3. Montrer que presque sûrement, μ_n converge étroitement vers μ quand $n \to \infty$.

Corrigé.

1. Soit $f \in \mathcal{C}_c(\mathbb{R})$. Pour $k, n \in \mathbb{N}$, on fixe $f_{n,k} \in \mathbb{Q}$ tel que
$$\begin{cases} |f_{n,k} - f(k/n)| \le 1/n & \text{si } f(k/n) \notin \mathbb{Q}, \\ f_{n,k} = f(k/n) & \text{si } f(k/n) \in \mathbb{Q}. \end{cases}$$
On définit pour $n \in \mathbb{N}$ la fonction f_n comme étant la fonction affine par morceaux interpolant les $f_{n,k}$ pour $k \in \mathbb{Z}$. Alors $f_n \in \mathcal{C}_c(\mathbb{R})$ et on vérifie, en utilisant l'uniforme continuité de f, que
$$\|f - f_n\|_\infty \xrightarrow[n\to\infty]{} 0.$$
Donc l'ensemble des fonctions affines par morceaux à support compact interpolant un nombre fini de points à coordonnées rationnelles est dense dans $\mathcal{C}_c(\mathbb{R})$ pour la norme infini et il est aussi dénombrable.
2. Le sens direct est clair, montrons la réciproque : on suppose que
$$\forall h \in H, \qquad \int_{\mathbb{R}} h \,\mathrm{d}\nu_n \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} h \,\mathrm{d}\nu.$$
Soit $f \in \mathcal{C}_c(\mathbb{R})$ et $\varepsilon > 0$. Il existe $h \in H$ telle que $\|h - f\|_\infty \le \varepsilon$. Alors on a
$$\begin{aligned} \left| \int_{\mathbb{R}} f \,\mathrm{d}\nu_n - \int_{\mathbb{R}} f \,\mathrm{d}\nu \right| &\le \left| \int_{\mathbb{R}} f \,\mathrm{d}\nu_n - \int_{\mathbb{R}} h \,\mathrm{d}\nu_n \right| + \left| \int_{\mathbb{R}} h \,\mathrm{d}\nu_n - \int_{\mathbb{R}} h \,\mathrm{d}\nu \right| + \left| \int_{\mathbb{R}} h \,\mathrm{d}\nu - \int_{\mathbb{R}} f \,\mathrm{d}\nu \right| \\ &\le \|h - f\|_\infty + \left| \int_{\mathbb{R}} h \,\mathrm{d}\nu_n - \int_{\mathbb{R}} h \,\mathrm{d}\nu \right| + \|h - f\|_\infty. \end{aligned}$$
Par hypothèse, on a donc
$$\limsup_{n\to\infty} \left| \int_{\mathbb{R}} f \,\mathrm{d}\nu_n - \int_{\mathbb{R}} f \,\mathrm{d}\nu \right| \le 2\varepsilon,$$
et avec $\varepsilon \to 0$, on en déduit
$$\forall f \in \mathcal{C}_c(\mathbb{R}), \qquad \int_{\mathbb{R}} f \,\mathrm{d}\nu_n \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f \,\mathrm{d}\nu,$$
c'est-à-dire que ν_n converge vaguement vers ν. Comme ν et les ν_n sont des mesures de probabilités, cela implique la convergence étroite.
3. Soit H une partie dénombrable dense de $\mathcal{C}_c(\mathbb{R})$ pour la norme infini, qui existe par la question 1. Soit $h \in H$. On a
$$\left| \int_{\mathbb{R}} h \,\mathrm{d}\mu_n - \int_{\mathbb{R}} h \,\mathrm{d}\mu \right| = \left| \frac{1}{n} \sum_{k=1}^{n} h(X_k) - \mathbb{E}[h(X_1)] \right| \xrightarrow[n\to\infty]{} 0, \quad \text{presque sûrement,}$$
par la loi forte des grands nombres, car $\mathbb{E}[h(X_1)] < \infty$. Comme H est dénombrable, on peut échanger le "$\forall h \in H$" et le "presque sûrement" : on a donc presque sûrement
$$\forall h \in H, \qquad \int_{\mathbb{R}} h \,\mathrm{d}\mu_n \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} h \,\mathrm{d}\mu.$$
Par la question 2., on en déduit que presque sûrement, μ_n converge étroitement vers μ quand $n \to \infty$.

Exercice 10. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes et identiquement distribuées. On note F la fonction de répartition de X_1 et on suppose que X_1 n'est pas p.s. constante. On pose $\alpha := \inf\{x \in \mathbb{R} : F(x) > 0\}$ et $\beta := \sup\{x \in \mathbb{R} : F(x) < 1\}$.

1. Montrer que $\alpha < \beta$, que $\alpha \neq +\infty$ et que $\beta \neq -\infty$.
2. Montrer que
$$\liminf_{n\to\infty} X_n = \alpha \quad \text{et} \quad \limsup_{n\to\infty} X_n = \beta,$$
presque sûrement.

Corrigé.

1. On a $\alpha \leq \beta$ car F est croissante. Supposons que $\alpha = \beta$. Alors par continuité à droite $F(\alpha) = 1$ et F est la fonction de répartition de la mesure δ_α ce qui contredit l'hypothèse de l'énoncé. Si $\beta = +\infty$ cela signifie que $F \equiv 1$ ce qui n'est pas possible car $F(x) \to 0$ quand $x \to -\infty$. De même $\alpha \neq +\infty$.
2. Montrons tout d'abord que, pour $a \in \mathbb{R}$ tel que $0 < F(a) < 1$, on a p.s.
$$\liminf_{n\to\infty} X_n \leq a \leq \limsup_{n\to\infty} X_n.$$
On a $\limsup\{X_n > a\} \subset \{\limsup_{n\to\infty} X_n \geq a\}$ et $\mathbb{P}(X_n > a) = 1 - F(a) \in]0,1[$. Donc $\sum_{n\geq 1} \mathbb{P}(X_n > a) = +\infty$ et d'après le lemme de Borel–Cantelli (les événements $\{X_n > a\}$ étant indépendants), on a
$$\mathbb{P}(\limsup\{X_n > a\}) = 1.$$
Ainsi p.s., $\limsup_{n\to\infty} X_n \geq a$. De même, $\limsup\{X_n \leq a\} \subset \{\liminf_{n\to\infty} X_n \leq a\}$, et on montre comme précedemment que p.s. $\liminf_{n\to\infty} X_n \leq a$.
Soit $(\beta_k)_{k\geq 1}$ une suite croissante qui tend vers β telle que $\alpha < \beta_k < \beta$. D'après le petit résultat montré juste avant, on a pour tout $k \geq 1$, p.s., $\limsup_{n\to\infty} X_n \geq \beta_k$. Ainsi, p.s., $\limsup_{n\to\infty} X_n \geq \beta$. Supposons $\beta < +\infty$. Soit $k \geq 1$. On a $\{\limsup_{n\to\infty} X_n > \beta + 1/k\} \subset \limsup\{X_n > \beta + 1/k\}$. Or $F(\beta + 1/k) = 1$, donc $\mathbb{P}(X_n > \beta + 1/k) = 0$
$$\mathbb{P}(\limsup\{X_n > \beta + 1/k\}) = 0.$$
Ainsi pour tout $k \geq 1$, p.s., $\limsup_{n\to\infty} X_n \leq \beta + 1/k$ puis, p.s., $\limsup_{n\to\infty} X_n \leq \beta$.
Pour l'autre limite, ça marche pareil. Soit $(\alpha_k)_{k\geq 1}$ une suite décroissante qui tend vers α telle que $\alpha < \beta_k < \beta$. D'après le petit résultat montré juste avant, on a pour tout $k \geq 1$, p.s., $\limsup_{n\to\infty} X_n \leq \alpha_k$. Ainsi, p.s., $\limsup_{n\to\infty} X_n \leq \alpha$. Supposons $\alpha > -\infty$. Soit $k \geq 1$. On a $\{\liminf_{n\to\infty} X_n < \alpha - 1/k\} \subset \limsup\{X_n < \alpha - 1/k\}$. Comme $F(\alpha - 1/k) = 0$ pour tout $k \geq 1$, on conclut comme précedemment que, p.s., $\liminf_{n\to\infty} X_n \geq \alpha$.

Exercice 11. Montrer qu'il n'existe pas de mesure de probabilité sur $\mathbb{N}^*$ telle que, pour tout $n \geq 1$, la probabilité de l'ensemble des multiples de n soit égale à $1/n$.

Corrigé. Supposons qu'il existe une telle probabilité $\mathbb{P}$. Soit $p_1, \ldots, p_k$ des nombres premiers, on a
$$\mathbb{P}(p_1\mathbb{N} \cap \cdots \cap p_k\mathbb{N}) = \mathbb{P}((p_1 \cdots p_k)\mathbb{N}) = \frac{1}{p_1 \cdots p_k} = \prod_{i=1}^{k} \mathbb{P}(p_i\mathbb{N}),$$
donc $p_1\mathbb{N}, \ldots, p_k\mathbb{N}$ sont des événements indépendants. En outre,
$$\sum_{p \text{ premier}} \mathbb{P}(p\mathbb{N}) = \sum_{p \text{ premier}} \frac{1}{p} = \infty,$$

où l'on utilise que la série des inverses des nombres premiers diverge (voir sur Wikipédia pour une preuve). Donc, par le lemme de Borel–Cantelli, presque tout entier $n \in \mathbb{N}$ (pour $\mathbb{P}$) appartient à une infinité de $p\mathbb{N}$, c'est-à-dire est multiple d'une infinité de nombres premiers distincts. Mais, il n'y a que 0 qui vérifie cette propriété. Donc $\mathbb{P}(\mathbb{N}^*) = 0$ et ainsi $\mathbb{P}(\{0\}) = 1$. Cela contredit la définition de $\mathbb{P}$: $\forall n \in \mathbb{N}$, $0 \in n\mathbb{N}$ donc $\mathbb{P}(n\mathbb{N}) = 1 \neq 1/n$.

Exercice 12. (*Retours en 0 de la marche aléatoire simple asymétrique*) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ indépendantes de loi de Bernoulli de paramètre $p \in]0,1[$ tel que $p \neq 1/2$. On considère $S_0 := 0$ et $S_n := X_1 + \cdots + X_n$ pour $n \geq 1$. Montrer que presque sûrement le nombre de passages en 0 de la marche $(S_n)_{n\in\mathbb{N}}$ est fini.

Corrigé. On pose $A_n := \{S_n = 0\}$, de sorte que $\limsup_{n\to\infty} A_n$ soit l'événement "$(S_n)_{n\in\mathbb{N}}$ passe une infinité de fois en 0". D'après le lemme de Borel–Cantelli, il suffit de montrer que $\sum \mathbb{P}(A_n)$ converge. Or $\mathbb{P}(A_n)$ est nul si n est impair et

$$\mathbb{P}(A_{2n}) = \mathbb{P}(S_{2n} = 0) = \binom{2n}{n} p^n (1-p)^n.$$

Or $\mathbb{P}(A_{2n+2})/\mathbb{P}(A_{2n}) = \frac{(2n+1)(2n+1)}{(n+1)^2} p(1-p) \to 4p(1-p)$ quand $n \to \infty$. Or $4p(1-p) < 1$ puisque $p \neq 1/2$, d'où la convergence de la série.

Exercice 13.

1. Soit μ une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Pour tout $n \geq 1$, on définit une mesure de probabilité μ_n sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ par :
$$\mu_n := \sum_{k\in\mathbb{Z}} \mu([k/n, (k+1)/n[)\delta_{k/n}.$$
Montrer que μ_n converge étroitement vers μ quand $n \to \infty$.
2. En déduire que si $(X_n)_{n\geq 1}$ est une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, respectivement de loi géométrique de paramètre $\mathrm{e}^{-1/n}$, alors la suite $(X_n/n)_{n\geq 1}$ converge en loi vers une variable aléatoire exponentielle de paramètre 1.

Corrigé.

1. Soit X une variable aléatoire réelle définie sur un espace $(\Omega, \mathcal{A}, \mathbb{P})$ de loi μ. Alors on voit que pour tout $n \geq 1$, la variable aléatoire $Y_n = \lfloor nX \rfloor / n$ ($\lfloor x \rfloor$ désigne la partie entière de x) suit la loi μ_n. Et $Y_n \to X$ p.s. Donc $Y_n \to X$ en loi, ce qui signifie que $\mu_n \to \mu$ étroitement.
2. On pose $m(\mathrm{d}x) = \mathrm{e}^{-x}\mathbb{1}_{\{x>0\}}\mathrm{d}x$. Alors on vérifie que pour tout $n \geq 1$, m_n est la loi de la variable aléatoire X_n/n. En effet,
$$\mathbb{P}(X_n = k) = \mathrm{e}^{-k/n} - \mathrm{e}^{-(k+1)/n} = \int_{k/n}^{(k+1)/n} \mathrm{e}^{-x}\,\mathrm{d}x = m_n(\{k/n\}).$$

Exercice 14.

1. Soit $f : [0,1] \to \mathbb{R}$ continue. Déterminer la limite de
$$\int_{[0,1]^n} f\left(\frac{x_1 + \ldots + x_n}{n}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$
quand $n \to \infty$.

2. Soit $f : \mathbb{R}_+ \to \mathbb{R}$ continue bornée et $\lambda > 0$. Déterminer la limite de

$$\sum_{k=0}^{\infty} \mathrm{e}^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right)$$

quand $n \to \infty$.

Corrigé.

1. On a

$$\int_{[0,1]^n} f\left(\frac{x_1 + \ldots + x_n}{n}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n = \mathbb{E}\left[f\left(\frac{U_1 + \ldots + U_n}{n}\right)\right]$$

où $U_1, \ldots, U_n$ sont des variables aléatoires indépendantes et uniformes sur $[0,1]$. D'après la loi forte des grands nombres, $n^{-1}(U_1 + \ldots + U_n)$ converge p.s. et donc en loi vers $1/2$. Ainsi,

$$\lim_{n\to\infty} \int_{[0,1]^n} \mathrm{d}x_1, \ldots \mathrm{d}x_n f\left(\frac{x_1 + \ldots + x_n}{n}\right) = f\left(\frac{1}{2}\right).$$

2. On a

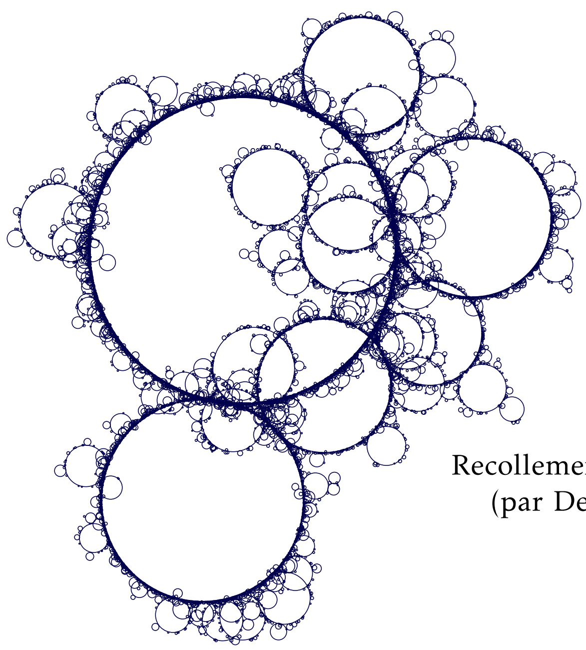
$$\sum_{k=0}^{\infty} \mathrm{e}^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right) = \mathbb{E}\left[f\left(\frac{P_1 + \ldots + P_n}{n}\right)\right],$$

où $P_1, \ldots, P_n$ sont des variables aléatoires de Poisson de paramètre λ indépendantes. D'après la loi forte des grands nombres, $n^{-1}(P_1 + \ldots + P_n)$ converge p.s. et donc en loi vers λ. Ainsi,

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right) = f(\lambda).$$

Remarque. On a utilisé les résultats suivants que vous pourrez vérifier à titre d'exercice :

- ▷ La moyenne d'une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$ est λ.
- ▷ Soient $n \geq 1$, $\lambda_1, \ldots, \lambda_n > 0$ et $X_1, \ldots, X_n$ n variables aléatoires *indépendantes*, chaque X_i suivant une loi de Poisson de paramètre λ_i. Alors $X_1 + \ldots + X_n$ suit une loi de Poisson de paramètre $\lambda_1 + \ldots + \lambda_n$.



Recollement aléatoire de cercles
(par Delphin Sénizergues)

Fin