Convergences

Dans ce chapitre, les mesures considérées sont des mesures sur $\mathbb{R}^d$ mais tout ceci reste valable si on remplace $\mathbb{R}^d$ par un espace métrique localement compact dénombrable à l'infini (i.e. réunion dénombrable de compacts).

1 Convergence des mesures bornées sur $\mathbb{R}^d$

On commence par donner quelques notations utilisées tout au long du chapitre.

Ensembles de mesures (sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$) :
$\mathcal{M}$: ensemble des mesures positives de masse finie.
$\mathcal{M}_b$: ensemble des mesures de masse positive inférieure ou égale à b.
$\mathcal{M}^1$: ensemble des probabilités.

Ensembles de fonctions (sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$) :
C_K : ensemble des fonctions continues à support compact.
C_0 : ensemble des fonctions continues tendant vers 0 à l'infini.
C_b : ensemble des fonctions continues bornées.

Remarque 1.1. *C_0, C_K munis de la norme uniforme sont séparables mais C_b ne l'est pas (on peut construire des boules disjointes de rayon constant indexées par les partitions d'entiers en utilisant des fonctions en dents de scie 'indicatrices de ces ensembles').*
$(C_b, \|\cdot\|_\infty)$ est un espace de Banach dont C_0 est un sev fermé (mais pas C_K).
C_K est dense dans C_0 (mais pas dans C_b).

1.1 Topologies vague, faible et étroite

Définition 1.1. *Les topologies vague, faible et étroite sont respectivement les topologies les moins fines sur $\mathcal{M}$ (celles qui contiennent le moins d'ouverts) rendant continues les applications $\mu \to \int f \mu$ (avec $f \in C_K, C_0$ et C_b respectivement).*

C'est à dire que ce sont les topologies engendrées par les images réciproques des ouverts par les applications correspondantes. On notera que la topologie faible (vague ou étroite) ne coincide pas forcément avec la topologie faible de l'analyse fonctionnelle, c'est à dire la topologie des convergences des formes linéaires continues sur l'espace $\mathcal{M}$. Plus précisément, il faut déterminer le dual de $\mathcal{M}$. On sait par le theoreme de Riesz que le dual de C_0 (ou C_K) muni de la norme infinie est donné par $\mathcal{M}$. Mais cela ne suffit pas pour conclure. La topologie faible $\sigma(\mathcal{M}_b, C_K)$ (ou vague

$\sigma(\mathcal{M}_b, C_0) = \sigma(C_0(\mathbb{R}^d)', C_0(\mathbb{R}^d))$, ou étroite $\sigma(\mathcal{M}_b, C_b))$ ne coincide pas forcément avec $\sigma(\mathcal{M}_b, \mathcal{M}_b^*)$ (ou $\sigma(\mathcal{M}_b^*, \mathcal{M}_b)$).

On rappelle qu'une base de voisnage $\mathcal{V}_x$ de x pour la topologie $\mathcal{T}$ est une collection d'ensemble tel que tout élément de $\mathcal{T}$ contenant x contient un élement de $\mathcal{V}_x$. Dans ce cas, une base de voisinages de μ pour ces topologies (vague, faible, étroite) est composée des ensembles

$$V^{\mu}_{\varepsilon, f_1 \cdots f_n} = \{\nu \in \mathcal{M}, \forall i = 1, \cdots, n \ |\mu(f_i) - \nu(f_i)| < \varepsilon\} = \cap_{i=1}^n V^{\mu}_{\varepsilon, f_i}$$

où $\varepsilon > 0$ et $f_1, \ldots, f_n$ appartiennent à l'espace considéré (C_K, C_0, ou C_b). On notera en particulier que tout ouvert comme union de tels voisinages.

En particulier, $(\mu_n)_n$ converge vers μ :

- vaguement ssi $\forall f \in C_K, \lim_{n\to\infty} \mu_n(f) = \mu(f)$
- faiblement ssi $\forall f \in C_0, \lim_{n\to\infty} \mu_n(f) = \mu(f)$
- étroitement ssi $\forall f \in C_b, \lim_{n\to\infty} \mu_n(f) = \mu(f)$ (Notation $\mu_n \Rightarrow \mu$.)

La topologie la plus fine est l'étroite, puis vient la faible et enfin la vague. Ainsi, une suite convergeant étroitement converge faiblement et une suite convergeant faiblement converge vaguement.

Proposition 1.1. *On a les propriétés suivantes :*
1) Sur $\mathcal{M}_b$, les topologies vagues et faibles coïncident.
2) $\mathcal{M}_b$ est métrisable compact pour la topologie faible.
3) Si μ_n converge faiblement vers μ dans $\mathcal{M}_b$, alors $\mu(\mathbb{R}^d) \leq \liminf_{n\to\infty} \mu_n(\mathbb{R}^d)$. Si il y a égalité, alors la convergence est étroite.
4) Sur $\mathcal{M}^1$, les trois topologies coïncident.

Remarque 1.2. *Le point 1) vient du fait que C_K est dense dans C_0 pour la norme infinie, et le fait que C_0 n'est pas dense dans C_b empêche de faire pour la topologie étroite.*
Le point 2) vient de la séparabilité de C_K (ou C_0), qui permet de définir une distance et faire une extraction diagonale pour montrer la compacité. A nouveau, la non séparabilité de C_b empêche de généraliser à la topologie étroite.
Les points 3) et 4) pointent alors la différence fondamentale entre une convergence faible (ou vague) et une convergence étroite dans M_b : la convergence faible n'assure pas la conservation de la masse de la mesure (i.e. $\mu_n(\mathbb{R}^d) = 1$ n'implique pas $\mu(\mathbb{R}^d) = 1$). Ce problème va être gérer par la tension, juste après.

Démonstration. 1) Si μ_n converge vaguement dans $\mathcal{M}_b$ vers μ alors elle converge faiblemement (utiliser $|\mu_n(f) - \mu(f)| \leq |\mu_n(f_p) - \mu(f_p)| + 2b\|f_p - f\|_\infty$ où $f \in C_0$ et $f_p \in C_K$ et obtenir 1)).
Mais C_0 n'est pas dense dans dans C_b et par exemple δ_n converge vaguement mais pas étroitement vers la mesure nulle. Il faut travailler dans $\mathcal{M}^1$ pour que ca marche en utilisant une fonction $h_p \in C_K$ qui tend en croissant vers 1.
C'est ce qui fait que la convergence des transformées de Laplace pour les v.a. positives entraine la convergence étroite, voir aussi le théorème de Lévy avec la continuité de la fonction caractéristique

en 0.

On va pouvoir appliquer ce critère sequentiel pour montrer que les topologies coincident en utilisant les voisinages donnés précédemment.
Plus précisément, pour montrer que la topologie $\mathcal{T}$ contient la topologie $\mathcal{T}'$, il suffit de montrer que tout élément d'une base de voisinage de μ pour $\mathcal{T}'$ contient un élément de la base de voisinage de μ pour $\mathcal{T}$. C.a.d il faut montrer que pour tout ε, f, μ, il existe ε', f' tels que $V^{\mu}_{\varepsilon',f'} \subset V^{\mu}_{\varepsilon,f}$.
On rappelle que l'on dit qu'un sous-ensemble H d'un espace vectoriel E est *total* ssi $Vect(H)$ est dense dans E. Le résultat vient du lemme :

Lemme 1.1. *La topologie vague coincide avec la topologie la moins fine rendant continues les applications $\mu \to \int f\mu$ lorsque f parcourt un ensemble total H de C_K.*

Le cas de la convergence faible se traite de manière semblable.
Démonstration. Soit H un ensemble total de C_K. On peut toujours supposer que H est un espace vectoriel quitte à considérer $Vect(H)$. Soit alors $f \in C_K$ et $g \in H$ telle que $\|f-g\|_\infty \leq \varepsilon$. Si $\nu, \mu \in \mathcal{M}_b$, on a

$$\begin{aligned} |\mu(f)-\nu(f)| &\leq |\mu(f)-\mu(g)|+|\mu(g)-\nu(g)|+|\nu(g)-\nu(f)| \\ &\leq 2b\varepsilon + |\nu(g)-\mu(g)| \end{aligned}$$

et par suite $V^{\mu}_{\varepsilon,g} \subset V^{\mu}_{(2b+1)\varepsilon,f}$ donc puisque la topologie vague est celle qui contient le moins d'ouverts, les deux topologies coïncident. □

2) On prend $(f_n)_n$ une suite d'éléments dense dans C_K. On définit une distance d sur $\mathcal{M}_b$ par

$$d(\mu,\nu) = \sum_{n\geq 1} \frac{1}{2^n\|f\|_\infty}|\mu(f_n)-\nu(f_n)|$$

Il est aisé de montrer que c'est bien une distance. En suivant la preuve de 1), la topologie induite par d coïncide avec la topologie faible.
Montrons maintenant que $\mathcal{M}_b$ est faiblement compact. On peut se restreindre à montrer que $\mathcal{M}_b$ est sequentiellement compact (car métrisable). Soit donc $(\mu_p)_p \in \mathcal{M}_b^{\mathbb{N}}$. On veut en extraire une sous-suite faiblement convergente. Or

$$\forall n \in \mathbb{N},\ |\mu_p(f_n)| \leq b\|f_n\|_\infty$$

Par un procédé d'extraction diagonale, on peut extraire une sous-suite $(\mu_{p_k})_k$ telle que

$$\lim_{k\to\infty} \mu_{p_k}(f_n) = \Psi(f_n),\ \ n \in \mathbb{N}$$

Par densité (prolongement uniformément continue d'une application uniformément continue), il existe pour tout f de C_K, un réel $\Psi(f)$ tel que

$$\mu_{p_k}(f) \to_{k\to\infty} \Psi(f)$$

Le théorème de Riesz assure l'existence d'une unique mesure μ telle que $\mu(f) = \Psi(f)$ pour tout $f \in C_K$.
(Attention, Ouvrard ne détaille pas ce passage : il passe à C_0 directement). Il reste à montrer que

$\mu \in \mathcal{M}_b$. On prend alors une suite $h_p \in C_K$, $0 \le h_p \le 1$ tendant simplement en croissant vers 1. On a pour tout k, $\mu_{p_k}(h_p) \le b$ donc $\mu(h_p) \le b$ donc $\mu(\mathbb{R}^d) \le b$. Ceci conclut la preuve.

3) Pour $\mu(\mathbb{R}^d) \le \liminf_{n\to\infty} \mu_n(\mathbb{R}^d)$, il suffit de faire appel pour un compact K à une fonction continue à support compact f_K, qui vaut 1 sur K et qui bornée par 1. En effet, $\mu(K) \le \mu(f_K) \le \liminf_{n\to\infty} \mu_n(f_K) \le \liminf_{n\to\infty} \mu_n(\mathbb{R}^d)$. Pour la deuxième partie, on touche du doigt la tension : il faut montrer que si l'égalité est vraie, des inégalités réciproques peuvent être données,i.e. pour tout ϵ, il existe K tel que $\mu_n(K^c) \le \mu_n(\mathbb{R}_n) - \mu_n(f_{K'}) \le \epsilon$ pour tout $n \ge 0$.

4) Il nous suffit donc de montrer maintenant que sur $\mathcal{M}^1$, les topologies vagues et étroites coïncident. Il suffit pour cela de montrer que $\forall\, \mathbb{P} \in \mathcal{M}^1, \forall \varepsilon > 0, \forall f \in C_b$, $V^{\mathbb{P}}_{\varepsilon,f}$ contient un voisinage pour la topologie vague.
Soit $(h_p)_p$ une suite de fonctions positives de C_K qui convergent vers 1 simplement et en croissant. Pour toute probabilité $\mathbb{Q}$, on a

$$\begin{aligned} |\mathbb{P}(f) - \mathbb{Q}(f)| &\le |\mathbb{P}(f - fh_p)| + |\mathbb{P}(fh_p) - \mathbb{Q}(fh_p)| + |\mathbb{Q}(fh_p - f)| \\ &\le \|f\|_\infty \{\mathbb{Q}(1-h_p) + \mathbb{P}(1-h_p)\} + |\mathbb{P}(fh_p) - \mathbb{Q}(fh_p)| \\ &= \|f\|_\infty \{2 - \mathbb{Q}(h_p) - \mathbb{P}(h_p)\} + |\mathbb{P}(fh_p) - \mathbb{Q}(fh_p)| \\ &\le \|f\|_\infty \{2\mathbb{P}(1-h_p) + \mathbb{P}(h_p) - \mathbb{Q}(h_p)\} + |\mathbb{P}(fh_p) - \mathbb{Q}(fh_p)| \end{aligned}$$

Remarquons maintenant que par le théorème de la convergence monotone, $\lim_{p\to\infty} \mathbb{P}(1-h_p) = 0$. On fixe alors p tel que $2\mathbb{P}(1-h_p) \le \varepsilon$. Il est aisé de voir que l'inégalité précédente ne dit rien d'autre que

$$\mathcal{M}^1 \cap V^{\mathbb{P}}_{\varepsilon,h_p} \cap V^{\mathbb{P}}_{\varepsilon,fh_p} \subset \mathcal{M}^1 \cap V^{\mathbb{P}}_{(2\epsilon\|f\|_\infty+\epsilon),f}$$

autrement dit, que tout voisinage de la topologie étroite sur $\mathcal{M}^1$ contient un voisinage ouvert de la topologie vague sur $\mathcal{M}^1$, c.q.f.d. □

Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $\delta_{1/n} \Rightarrow \delta_0$ mais $\delta_{1/n}(\{0\}) = 0$ ne converge pas vers $\delta_0(\{0\}) = 1$. Aussi, $\mu_n \Rightarrow \mu$ n'implique pas que $(\mu_n(A))$ converge vers $\mu(A)$ pour tout borélien A. D'où la défintion et le résultat suivant :

Définition 1.2. *Soit μ une mesure sur $\mathbb{R}^d$. On dira qu'un borélien A est de μ-continuité ssi $\mu(\partial A) = 0$ (∂A désigne la frontière de A).*

Proposition 1.2. *Soit $(\mu_n)_n \in \mathcal{M}_b^{\mathbb{N}}$ et $\mu \in \mathcal{M}_b$. Il y a équivalence entre :*
a) $\mu_n \Rightarrow \mu$
b) Pour tout fermé F, $\limsup \mu_n(F) \le \mu(F)$ et $\lim_{n\to\infty} \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$.
c) Pour tout ouvert O, $\liminf \mu_n(O) \ge \mu(O)$ et $\lim_{n\to\infty} \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$.
d) Pour tout borélien A de μ-continuité, $\mu_n(A) \to_{n\to\infty} \mu(A)$.

Démonstration. Nous ne donnons ici que la démonstration de quelques implications. Pour plus de détails, se reporter à Ouvrard.
a) $\Rightarrow$ *b)* : La deuxième partie de b) revient à faire $f = 1$. Passons à la première. Soit $\phi_j : \mathbb{R} \to [0,1]$ continue définie par $\phi_j(x) = 1$ si $x \le 0$, $\phi_j(x) = 0$ si $x \ge 1/j$ et ϕ_j affine sur $[0, 1/j]$. On définit alors $f_j \in C_b$ par $f_j(x) = \phi_j(d(x,F))$ qui tend en décroissant vers la fonction caractéristique 1_F de F (car F fermé). D'après le théorème de convergence monotone décroissante,

$$\mu(F) = \lim_{j\to\infty} \mu(f_j)$$

donc pour tout $\varepsilon > 0$, il existe un entier j_0 tel que $\mu(F) \le \mu(f_{j_0}) \le \mu(F) + \varepsilon$.
Or pour tout $n \in \mathbb{N}$, $\mu_n(F) \le \mu_n(f_{j_0})$ donc

$$\limsup_n \mu_n(F) \le \limsup_n \mu_n(f_{j_0}) = \mu(f_{j_0}) \le \mu(F) + \varepsilon$$

$b) \iff c)$: Passer au complémentaire.
$c) \Rightarrow d)$: On a $A^\circ \subset A \subset \bar{A}$ et on sait b) et c) car $b \iff c$. Donc on a :

$$\mu(A^\circ) \le \liminf \mu_n(A^\circ) \le \liminf \mu_n(A) \le \limsup \mu_n(A) \le \limsup \mu_n(\bar{A}) \le \mu(\bar{A})$$

or $\mu(\partial A) = 0$ et par conséquent, toutes les inégalités sont en fait des égalités.
$c) \Rightarrow a)$ On rappelle qu'une application classique du théorème de Fubini permet d'écrire que pour toute fonction $f \in C_b$ positive et pour toute mesure $\mu \in \mathcal{M}_b$,

$$\int_{\mathbb{R}^d} f(x)\mu(x) = \int_0^{\|f\|_\infty} \mu(f \ge u)du$$

et

$$\int_{\mathbb{R}^d} f(x)\mu(x) = \int_0^{\|f\|_\infty} \mu(f > u)du \quad (*)$$

Soit alors $f \in C_b$ positive. Par le théorème de Fatou, on a

$$\begin{aligned}\limsup_n \int_{\mathbb{R}^d} f(x)\mu_n(x) &= \limsup_n \int_0^{\|f\|_\infty} \mu_n(f \ge u)du \\ &\le \int_0^{\|f\|_\infty} \limsup_n \mu_n(f \ge u)du\end{aligned}$$

Or $\{f \ge u\}$ est fermé et on a supposé $b)$ vraie (car $b) \Leftrightarrow c)$) donc on a

$$\limsup_n \int_{\mathbb{R}^d} f\mu_n \le \int_0^{\|f\|_\infty} \mu(f \ge u)du = \int_{\mathbb{R}^d} f\mu$$

On a de même, en utilisant $(*)$ et $c)$, que

$$\liminf_n \int_{\mathbb{R}^d} f\mu_n \ge \int_{\mathbb{R}^d} f\mu$$

On a donc montré que $\lim_{n\to\infty} \int_{\mathbb{R}^d} f\mu_n = \int_{\mathbb{R}^d} f\mu$ pour toute fonction continue bornée et positive. Par linéarité, on en déduit le résultat pour toute fonction f continue bornée de signe quelconque.
□

1.2 Tension

La propriété de tension est une propriété qui permet de limiter les phénomènes de perte de masse tel que $\delta_n \Rightarrow 0$ et donc d'extraire des sous suites convergentes dans $\mathcal{M}^1$.

Définition 1.3. *Une suite $(\mu_n)_n$ de mesures de $\mathcal{M}_b$ est tendue ssi pour tout $\varepsilon > 0$, il existe un compact K de $\mathbb{R}^d$ tel que*

$$\sup_n \mu_n(K^c) \le \varepsilon$$

Proposition 1.3. *1) Si $(\mu_n)_n \in \mathcal{M}_b$ converge étroitement vers $\mu \in \mathcal{M}_b$ alors $(\mu_n)_n$ est tendue.*
2) Si $(\mu_n)_n$ est tendue et si $(\mu_n)_n \in (\mathcal{M}^1)^{\mathbb{N}}$ alors il existe une sous-suite $(\mu_{n_k})_k$ de $(\mu_n)_n$ et une probabilité $\mu \in \mathcal{M}^1$ telle que $\mu_{n_k} \Rightarrow \mu$.

Démonstration. 1) Soit $\varepsilon > 0$ fixé. Il existe une boule ouverte O telle que $\mu(O) \geq \mu(\mathbb{R}^d) - \epsilon$. D'autre part, par la propriété précédente, $\liminf \mu_n(O) \geq \mu(O)$. Donc il existe un entier N_2 tel que pour $n \geq N_2$,

$$\mu_n(O) \geq \mu(\mathbb{R}^d) - \epsilon.$$

D'autre part, toujours par la propriété précédente, $\lim_{n\to\infty} \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$ donc il existe un rang $N_1 \geq N_2$ tel que pour $n \geq N_1$, $\mu_n(\mathbb{R}^d) \leq \mu(\mathbb{R}^d) + \epsilon$.
On en déduit que $\forall n \geq N_1, \mu_n(\bar{O}^c) \leq \mu_n(O^c) \leq 2\epsilon$.
N_1 est maintenant fixé. Pour chaque $n \leq N_1$, il existe un compact K_n tel que $\mu_n(K_n^c) \leq \epsilon$. Donc si on pose $K = \cup_{n \leq N_1} K_n \cup \bar{O}$, K est un compact et pour tout n,

$$\mu_n(K^c) \leq 2\epsilon$$

ce qui prouve la tension de la suite $(\mu_n)_n$.
2) Réciproquement, si $(\mu_n)_n$ est une suite de mesures tendues, on sait déjà qu'il existe une mesure μ de masse inférieure ou égale à 1 et une sous-suite $(\nu_n)_n$ de $(\mu_n)_n$ telle que $(\nu_n)_n$ converge faiblement vers μ. Pour montrer que $\nu_n \Rightarrow \mu$ et que $\mu \in \mathcal{M}_1$, il suffit de montrer que $\mu(\mathbb{R}^d) = 1$.
Or pour tout $\varepsilon > 0$, il existe un compact K_ϵ tel que $\mu_n(K_\epsilon) \geq 1 - \varepsilon$ pour tout n. Soit alors $f_\varepsilon \in C_0$ une fonction telle que :

$$1 \geq f_\varepsilon \geq 1_{K_\varepsilon} \geq 0$$

On a $\mu_n(f_\varepsilon) \geq \mu_n(K_\varepsilon) \geq 1 - \varepsilon$ et quand $n \to \infty$, on obtient donc $\mu(f_\varepsilon) \geq 1 - \varepsilon$. Puisque $\mu(\mathbb{R}^d) \geq \mu(f_\varepsilon)$, en faisant maintenant tendre ε vers 0, il en résulte que $\mu(\mathbb{R}^d) \geq 1$ et puisque l'on sait déjà que l'on a l'inégalité inverse, μ est une probabilité. $\square$

1.3 Le Théorème de convergence de Lévy

Ce théorème assure que la convergence en loi de v.a. via la convergence des fonctions caractéristiques. Il faut néanmoins demander à ce que la fonction limite soit continue, de façon à ce qu'elle corresponde bien la fonction caractéristique d'une v.a. On pourra méditer par exemple sur la limite de la fonction caractéristique de la loi uniforme sur $[-n, n]$ quand $n \to \infty$.

Application aux sommes de variables aléatoires : Théorèmes de événements rares (dont convergence de la loi binomiale vers la loi de Poisson), Théorème central limite.

Exercice 1.1. Soit $\varepsilon > 0$. Montrer qu'il existe une fonction $f_\varepsilon \in C_0(\mathbb{R}^d)$ telle que pour toute mesure bornée $\nu \in \mathbb{R}^d$ on ait

$$\int_{[0,\varepsilon]^d} \hat{\nu}(t)dt = \int_{\mathbb{R}^d} f_\varepsilon \nu.$$

On a $f_\varepsilon(x) = \Pi_{j=1}^d g_\varepsilon(x_j)$ où $g_\varepsilon(u) = \frac{e^{i\varepsilon u}-1}{iu}$ si $u \neq 0$, $= \varepsilon$ si $u = 0$ est bien continue. On utilise alors le lemme de Fubini.

Exercice 1.2. Notre but est de montrer le théorème de Lévy :

Théorème 1.1. *(a) Si $\mu_n \in \mathcal{M}_b$ converge étroitement vers μ, alors $\hat{\mu}_n \to \hat{\mu}$ simplement.*
(b) Inversement, si $\hat{\mu}_n \to \varphi$ simplement et si φ est continue en 0, alors il existe une unique mesure $\mu \in \mathcal{M}_b$ telle que $\varphi = \hat{\mu}$ et μ_n converge étroitement vers μ .

La preuve repose sur le compacité faible. La continuité en O de ϕ permet d'éviter la perte de masse, et donc assurer la convergence étroite. On montre ensuite qu'il y a une unique valeur d'adhérence en utilisant l'injectivité de la transformée de Fourier. Le cas $\mu_n(dx) = \frac{1}{2n} 1_{[-n,n]}(dx)$ illustre ce probleme de continuité : la limite de la transformée de Fourier est nulle partout sauf en O où elle vaut 1. Or la transformée de Fourier d'une mesure bornée est continue en 0, il ne peut donc y avoir convergence étroite vers une mesure bornée.

1) Montrer (a).

A partir de la question suivante on s'attache à démontrer (b) et on prend les hypothèses de (b).

2) Montrer que μ_n admet une sous-suite faiblement convergente, que l'on notera $\mu_{\psi(n)}$. On appellera μ sa limite.

3) Montrer en appliquant le théorème de Lebesgue que $\lim_n \int_{[0,\varepsilon]^d} \widehat{\mu_{\psi(n)}}(t)dt = \int_{[0,\varepsilon]^d} \varphi(t)dt.$

4) Soit f_ε définie dans l'exercice précédent. Montrer successivement en utilisant le 3) que

$$\lim_k \int f_\varepsilon \mu_{\psi(k)} = \int f_\varepsilon \mu,$$

$$\lim_k \int_{[0,\varepsilon]^d} \hat{\mu}_{\psi(k)}(t)dt = \int_{[0,\varepsilon]^d} \hat{\mu}(t)dt,$$

$$\frac{1}{\varepsilon^d} \int_{[0,\varepsilon]^d} \hat{\mu}(t)dt = \frac{1}{\varepsilon^d} \int_{[0,\varepsilon]^d} \varphi(t)dt.$$

5) a) Déduire de la question précédente que $\hat{\mu}(0) = \phi(0)$ et remarquer que $\lim_n \hat{\mu}_{\psi(n)}(0) = \phi(0) = \hat{\mu}(0)$.

b) Montrer que $(\mu_{\psi(n)})_n$ converge étroitement vers μ. En déduire que $\hat{\mu} = \phi$.

c) En déduire que $(\mu_n)_n$ converge faiblement vers μ puis que $(\mu_n)_n$ converge étroitement vers μ.

6) Sous l'hypothèse (a) ou (b), montrer que la suite des fonctions $\hat{\mu}_n$ est équicontinue : on utilisera la proposition **??** et l'inégalité (que l'on démontrera) $|e^{ix.t} - e^{ix.t'}| \leq |x.(t-t')|$, où $x.t$ désigne le produit scalaire dans $\mathbb{R}^d$ de x et de t. En déduire que la suite $\hat{\mu}_n(t)$ converge uniformément vers $\hat{\mu}$ sur tout compact.

Solution :
1) Immédiat par défintion de la convergence étroite utilisée pour les fonctions $sin(< x, t >)$ *L'application de l'exercice précédent est immédiate car* μ_n *est bornée puisque*

$$||\mu_n||_{C_0'} = \sup_{||\varphi|| \leq 1} \int \varphi \mu_n \leq \mu_n(\mathbb{R}^d) = b.$$

2) Utiliser la compacité de la convergence faible.

3) La première relation découle de la convergence faible de μ_k et de $f_\varepsilon \in C_0(\mathbb{R}^d)$. La seconde découle de l'exercice précédent car

$$\int_{[0,\varepsilon]^d} \hat{\mu}_k(t)dt = \int f_\varepsilon \mu_k \to \int f_\varepsilon \mu = \int_{[0,\varepsilon]^d} \hat{\mu}(t)dt.$$

La troisième relation vient immédiatement de la précédente et du 3).

5) On passe à la limite dans la dernière relation du 4), φ étant continue par hypothèse, et $\hat{\mu}$ comme transformée de Fourier de mesure bornée. Donc $\hat{\mu}(0) = \varphi(0)$. Enfin

$$\lim_n \mu_{\psi(n)}(\mathbb{R}^d) = \lim_n \hat{\mu}_{\psi(n)}(0) = \varphi(0) = \hat{\mu}(0) = \mu(\mathbb{R}^d.)$$

Donc $\mu_{\psi(n)}$ converge étroitement vers l'unique μ telle $\hat{\mu} = \varphi$. Mais remarque que si $\mu_{\psi(n)}$ tend étroitement vers μ, alors par le (a) $\hat{\mu}_{\psi(n)}$ tend simplement vers $\hat{\mu}$. Mais comme par hypothèse $\hat{\mu}_{\psi(n)}$ tend simplement vers φ, on déduit que $\hat{\mu} = \varphi$. Ce raisonnement s'appliquant à toute sous-suite extraite, toutes les sous-suites faiblement convergentes de μ_n ont la même limite par le théorème d'injectivité de la Transformée de Fourier. Donc la suite μ_n de l'espace métrique compact $\mathcal{M}(b)$ a une unique limite et donc toute la suite converge.

2 Convergence de variables aléatoires

2.1 Rappel sur les différents types de convergence

On rappelle les principales notions de convergence. Ici $(X_n)_n$ est une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$ et X une variable aléatoire à valeur dans $\mathbb{R}^d$ toutes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

- Convergence presque sûre : $X_n \to X$ $\mathbb{P}$p.s. ssi $\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1$.
- Convergence $\mathcal{L}^p$: $X_n \to X$ dans $\mathcal{L}^p$ ssi $\mathbb{E}(|X_n - X|^p) \to_{n\to\infty} 0$.
- Convergence en probabilité : $X_n \to X$ en probabilité ssi $\forall \varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \to_{n\to\infty} 0$.
- Convergence en loi : $X_n \to X$ en loi ssi $\mathbb{E}(f(X_n)) \to_{n\to\infty} \mathbb{E}(f(X))$ pour tout $f \in C_b$.

Puisque les topologies faibles, vagues et étroites coïncident sur $\mathcal{M}^1$, dans la définition ci-dessus, on peut en fait remplacer C_b par C_K ou C_0.
La convergence en loi est juste la transcription de la convergence faible en termes de variables aléatoires. Les liens qui existent entre ces différentes notions de convergence seront démontrées en TD. Il faut comprendre que la convergence en loi est très différente des autres, elle n'est pas trajectorielle. Les convergences p.s. et L^p impliquent la convergence en $\mathbb{P}$ qui implique la convergence en loi. La convergence en loi implique la convergence en $\mathbb{P}$ quand la limite est constante. La convergence en proba ou la convergence L^p implique la convergence p.s. d'une sous suite. Pour un exemple de fonction (bornée) convergeant en proba vers 0 mais qui presque surement ne converge pas vers 0, prendre un créneau qui se déplace sur $[0, 1]$ en réduisant sa taille.

3 Convergence en loi

Notation : On utilise indifférement la notation $X_n \overset{n\to\infty,\ \mathcal{L}}{\longrightarrow} X$ ou bien $X_n \Rightarrow X$.

Ce qui suit sera démontré en TD.

Proposition 3.1. *Si $(X_n)_n$ converge en loi vers X et si $f : \mathbb{R}^d \to \mathbb{R}^k$ est continue alors $(f(X_n))_n$ converge en loi vers $f(X)$.*

Exemple : Si (X_n, Y_n) converge en loi vers (X, Y) alors $(X_n + Y_n)_n$ converge en loi vers $X + Y$. Mais il n'est pas vrai que si $(X_n)_n$ (resp. $(Y_n)_n$) converge en loi vers X (resp. Y) alors $(X_n + Y_n)_n$ converge en loi vers $X + Y$. C'est cependant vrai si X_n est indépendante de Y_n pour tout n et X indépendante de Y.

Proposition 3.2. *La convergence en probabilités implique la convergence en loi. La réciproque est fausse mais vraie si la limite est une constante.*

3.1 V.a. discrètes et à densité

Pour <u>les variables aléatoires discrètes</u>, on a un critère très simple de convergence en loi.

Proposition 3.3. *Soit $(X_n)_n, X$ des variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $\mathbb{Z}$. Alors*

$$X_n \Rightarrow X \text{ ssi } \forall r \in \mathbb{Z},\ \lim_{n\to+\infty} \mathbb{P}(X_n = r) = \mathbb{P}(X = r)$$

Exemple : Si X_n est une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, \lambda/n)$ où $\lambda > 0$ est un paramètre alors $(X_n)_n$ converge en loi vers une loi de Poisson de paramètre λ (Cf. Ouvrard).

$$\mathbb{P}(X_n = k) = C_n^k (\lambda/n)^k (1 - \lambda/n)^{n-k} \sim \frac{n...(n-k+1)}{n^k k!} \lambda^k e^{-\lambda} \sim \frac{\lambda^k}{k!} e^{-\lambda}.$$

Pour <u>les valeurs aléatoires à densité</u>, le lemme de Scheffé donne une condition suffisante (Ouvarard p 316) : la convergence p.s. des densités implique la convergence en loi.
En effet on montre la convergence de $\int |f_n - f|$ vers 0 en utilisant $|a - b| = a + b - 2\min(a, b)$ puis le théorème de convergence dominée pour $0 \leq \min(f_n, f) \leq f$.
Ce n'est pas une condition nécessaire (Ouvrard Exemple 14.3 p319).

3.2 Convergence en loi en termes de fonctions de répartition (dimension $d = 1$)

Théorème 3.1. *$(X_n)_n$ converge en loi vers X ssi $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ en tout point de continuité x de F_X.*

Démonstration.
$\Rightarrow$: Si x est un point de continuité de F_X, alors $(-\infty; x]$ est un ensemble de continuité de $\mathbb{P}_X$ donc

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty} \mathbb{P}_{X_n}((-\infty; x]) = \mathbb{P}_X((-\infty; x]) = F_X(x)$$

$\Leftarrow$: On rappelle qu'il n'y a au plus qu'un nombre dénombrable de points de discontinuité de F_X (car il ne peut y avoir pour chaque $\varepsilon > 0$ qu'un nombre fini de sauts de taille plus grande que ε puisque F_X est bornée et croissante). Soit donc $f \in C_0(\mathbb{R})$. Pour chaque $\varepsilon > 0$, on peut trouver par uniforme continuité une fonction g de la forme

$$g = \sum_{j=1}^{k} \alpha_j 1_{(a_j;b_j]}, \quad a_j < b_j \leq a_{j+1} < b_{j+1}$$

avec a_j et b_j qui ne sont pas des points de discontinuité de F_X et

$$\|f - g\|_\infty \leq \varepsilon$$

On a alors

$$\int g\mathbb{P}_{X_n} \to_{n\to\infty} \sum_{j=1}^{k} \alpha_j (F_X(b_j) - F_X(a_j)) = \int g\mathbb{P}_X$$

Il fuffit maintenant d'appliquer l'inégalité triangulaire pour obtenir

$$\left| \int f\mathbb{P}_{X_n} - \int f\mathbb{P}_X \right| \leq 2\varepsilon + \left| \int g\mathbb{P}_{X_n} - \int g\mathbb{P}_X \right| \leq 3\varepsilon$$

dès que n est assez grand. $\square$

Nous donnons maintenant le théorème d'Helly-Bray dont on pourra trouver une preuve élémentaire dans Williams (voir TD : compacité, densité, monotonie). C'est en fait une conséquence de la compacité faible de $\mathcal{M}_b$.

Théorème 3.2. ***(Helly-Bray)***
Soit $(F_n)_n$ une suite de fonctions de répartition. Il existe alors une fonction coissante F continue à droite et à valeur dans $[0,1]$ et une sous-suite $(F_{n_k})_k$ de $(F_n)_n$ telle que

$$\lim_{k\to\infty} F_{n_k}(x) = F(x)$$

en tout point de continuité x de F.

3.3 Théorème de Lévy

4 TD Probabilités : Modes de convergence

Exercice 4.1. Lien entre la convergence en probabilité et la convergence en norme. $\mathbb{L}^p$
Montrer à l'aide de l'inégalité de Markov que la convergence $\mathbb{L}^p$ implique la convergence en probabilité.

Exercice 4.2. Lien entre la convergence en probabilité et la convergence presque sûre, Brémaud P., *Introduction aux probabilités : Modélisation des phénomènes aléatoires*, pp. 175-176.

1) Montrer que la convergence presque sûre implique la convergence en probabilité.
Indication : si $X_n \to X$ p.s., poser $B_k = \limsup_n \{|X_n - X| \geq \frac{1}{k}\}$ ou appliquer le théorème de convergence dominée.

2) Montrer que la réciproque est fausse. On pourra considérer une suite $(X_n)_n$ de Bernoulli indépendantes de paramètres respectifs $p_n = 1 - q_n$ avec $\lim_{n\to\infty} p_n = 0$ mais tels que la série de terme général p_n soit divergente.

3) Montrer que de toute suite de variables aléatoires convergeant en probabilité, on peut extraire une sous-suite convergeant presque sûrement.
Solutions : 1) $B_k = \cap_{n\in\mathbb{N}} \cup_{m\geq n} \{|X_m - X| \geq \frac{1}{k}\}$. *Si* $\omega \notin \bigcup_k B_k$, *on a* $\lim X_n(\omega) = X(\omega)$, *donc*

$$\mathbb{P}(B_k) = 0 = \lim_{n\to\infty} \mathbb{P}(\cup_{m\geq n}|X_m - X| \geq \frac{1}{k}) \geq \lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \frac{1}{k}).$$

2) Comme p_n *tend vers zéro,* X_n *converge en probabilité vers* 0 *puisque* $\mathbb{P}(|X_n| > \epsilon) \leq p_n$. *On pose* $N = \{X_n = 1 \ i.o.\}$ *et* $\mathbb{P}(N) = 1$ *par le deuxième lemme de Borel Cantelli et la divergence de la série associée à* p_n.
3)Si $(X_n)_n$ *converge en probabilité vers* X, *montrer que l'on peut construire une sous-suite* $(X_{n_k})_k$ *de* $(X_n)_n$ *telle que* $\mathbb{P}(\| X_{n_k} - X \| \geq 1/k) \leq 1/2^k$ *et conclure par le lemme de Borel-Cantelli*

Exercice 4.3. *Lien entre la convergence en probabilité et la convergence en loi, Ouvrard, p. 315.*
Rappel : $X_n \to X$ en loi signifie $\mathbb{P}_{X_n} \to \mathbb{P}_X$ étroitement, i.e.

$$\forall f \in C_b(\mathbb{R}), \ \int_{\mathbb{R}} f\mathbb{P}_{X_n} \to \int_{\mathbb{R}} f\mathbb{P}_X, \ \text{(soit } \mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)])$$

Soit $(X_n)_n$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$.

1) (Ouvrard tome II, proposition 10.5) On suppose que les $(X_n)_n$ convergent en probabilité vers X. Soit f une fonction continue de $\mathbb{R}^d$ dans $\mathbb{R}$.
a) Montrer que pour tout $\epsilon, a > 0$, il existe $0 < \eta < 1$ tel que

$$\{|f(X_n) - f(X)| > \epsilon\} \subset \{|X| > a\} \cup \{|X_n - X| > \eta\}$$

b) En choisissant convenablement a, en déduire que $(f(X_n))_n$ converge en probabilité vers $f(X)$.

2) On suppose que les $(X_n)_n$ convergent en probabilité vers X. Montrer que les $(X_n)_n$ convergent en loi vers X.

3) Montrer que la réciproque est fausse en considérant le contre-exemple suivant : pour tout n, $X_n = X$ et X est une Bernoulli de paramètre $1/2$. Montrer que $(X_n)_n$ converge en loi vers $(1 - X)$ mais que $(X_n)_n$ ne peut pas converger en probabilité vers $(1 - X)$.

4) Montrer néanmoins que si les X_n convergent en loi vers une constante a alors X_n convergent en probabilité vers a.

Solutions 1.a) Par uniforme continuité de f sur $[-a-1, a+1]$, il existe $0 < \eta \leq 1$ tel que $-a-1 \leq x, y \leq a+1$ et $| x - y | \leq \eta$ impliquent $| f(x) - f(y) | \leq \epsilon$. Alors $|f(x) - f(y)| > \epsilon$ entraine $y > a$ ou $| x - y | > 1$ ($y \in [-a, a]$ mais $x \notin [-a-1, a+1]$) ou $| x - y | > \eta$.
1.b) On peut majorer la probabilité cherchée grace à l'inclusion précédente et passer à la limite

$$\lim_{n\to\infty} \mathbb{P}(|f(X_n) - f(X)| > \epsilon) \leq \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) + \lim_{a\to\infty} \mathbb{P}(|X| > a) = 0.$$

2) On décompose suivant l'évenement $\{|f(X_n) - f(X)| > \varepsilon\}$ et son complémentaire :

$$\forall \varepsilon > 0,\ \forall f \in C_b(\mathbb{R}),\ \ |\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \leq \varepsilon + 2||f||_\infty \mathbb{P}(|f(X_n) - f(X)| > \varepsilon)$$

3) C'est juste que l'égalité en loi n'implique pas l'égalité en probabilité, c'est à dire l'égalité presque sur : X_n converge déjà en probabilité vers X donc on aurait $X = 1 - X$ p.s.
4) Remarquer que pour $\epsilon > 0$, si δ_a désigne la masse de Dirac en a et $B(a, \epsilon)$ la boule fermée de centre a et de rayon ϵ, $\delta_a(\partial B(a, \epsilon)) = 0$. En déduire $\lim_{n\to+\infty} \mathbb{P}(\| X_n - a \| \leq \epsilon) = \delta_a(B(a, \epsilon)) = 1$. Une méthode plus rapide consiste à appliquer $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$ à la fonction $f(x) = |x - a|$.

Conclusion : On vient de démontrer que la convergence en norme $\mathbb{L}^p$ et la convergence presque sûre impliquent la convergence en probabilités qui implique la convergence en loi. Les réciproques sont fausses et sont partiellement vraies si on rajoute certaines hypothèses. Regarder en détail le tableau de la hiérarchie des convergences dans Brémaud P., *Introduction aux probabilités : Modélisation des phénomènes aléatoires*, p.238.

Exercice 4.4. Ouvrard tome II, Exercice 14.7
Soient sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ deus suites $(X_n)_n$ et $(Y_n)_n$ de variables aléatoires qui convergent en loi respectivement vers X et Y.

1) Montrer que si pour tout n, X_n et Y_n sont indépendantes et si X et Y sont indépendantes alors $(X_n, Y_n)_n$ converge en loi vers (X, Y).

2) On étudie un contre-exemple dans le cas où l'on supprime l'hypothèse "pour tout n, X_n et Y_n sont indépendantes". Soit donc X et Y deux variables aléatoires réelles indépendantes de même loi de Bernoulli de paramètre 1/2. On pose pour tout n.

$$X_n = X + 1/n,\ \ Y_n = (1 - X) - 1/n$$

Etudier la convergence en loi des trois suites $(X_n)_n$, $(Y_n)_n$ et $(X_n + Y_n)_n$ et conclure.

Solutions :
1) On utilise les fonctions caractéristiques et le théorème de Lévy (ou bien les fonctions de répartition) :

$$\mathbb{E}(e^{it_1 X_n + it_2 Y_n}) = \mathbb{E}(e^{it_1 X_n})\mathbb{E}(e^{it_2 Y_n}) \overset{n\to\infty}{\longrightarrow} \mathbb{E}(e^{it_1 X})\mathbb{E}(e^{it_2 Y}) = \mathbb{E}(e^{it_1 + it_2 Y})$$

avec X et Y indépendants.

Exercice 4.5. Convergence en loi pour les variables aléatoires discrètes, Ouvrard tome II, page 317. Soit $(X_n)_n$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace probabilisé et toutes à valeurs dans $\mathbb{Z}$. Montrer que $(X_n)_n$ converge en loi vers X si et seulement si pour tout $r \in \mathbb{Z}$, $\mathbb{P}(X_n = r) \longrightarrow_{n\to\infty} \mathbb{P}(X = r)$.

Indications : pour la condition nécessaire, approcher la fonction indicatrice de r par un élément de $C_K(\mathbb{R})$ à support dans $[r-\frac{1}{2}, r+\frac{1}{2}]$, et pour la condition suffisante, remarquer que $\mathbb{E}(f(X_n)) = \sum_{r\in K} f(r)\mathbb{P}(X_n = r)$ pour toute fonction continue à support dans le compact K.
Solutions : Soit $f \in C_K(\mathbb{R})$, supp$(f) \subset [-r+\frac{1}{2}, r+\frac{1}{2}]$, $f(r) \neq 0$. On a $\int_{\mathbb{R}} f\mathbb{P}_{X_n} \to \int_{\mathbb{R}} f\mathbb{P}_X$ et donc $f(r)\mathbb{P}(X_n = r) \to f(r)\mathbb{P}(X = r)$. Réciproquement, par la relation donnée, on a $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$, $\forall f \in C_k(\mathbb{R})$. Donc $\mathbb{E}f(X_n) \to \mathbb{E}f(X)$ $\forall f \in C_K(\mathbb{R})$. Comme sur $\mathcal{M}'$ les topologies faibles, vagues et étroites coïncident, on conclut.

Exercice 4.6. *Convergence vers la loi de Poisson : Théorème des évenements rares, Ouvrard tome II p321.* À tout entier n on associe les variables aléatoires $(X_{n,m})_{1\leq m\leq n}$. On suppose que ces variables sont indépendantes et que

$$\begin{aligned}
&\mathbb{P}(X_{n,m} = 1) = p_{n,m}, \qquad \mathbb{P}(X_{n,m} = 0) = 1 - p_{n,m},\\
&\lim_{n\to\infty} \sum_{m=1}^{n} p_{n,m} = \lambda \in]0, \infty[,\\
&\lim_{n\to\infty} \max_{1\leq m\leq n} p_{n,m} = 0.
\end{aligned}$$

Montrer que $S_n = \sum_{m=1}^{n} X_{n,m}$ converge en loi vers une variable aléatoire de loi de Poisson de paramètre λ (on pourra utiliser les fonctions caractéristiques et l'inégalité $\left|\Pi_{m=1}^n z_m - \Pi_{m=1}^n w_m\right| \leq \theta^{n-1}\sum_{m=1}^{n} |z_m - w_m|$ pour z_i, w_i des complexes de module $\leq \theta$.)

Solution : On pose

$$\phi_{n,m}(t) = \mathbb{E}(\exp(itX_{n,m})) = (1 - p_{n,m}) + p_{n,m}\exp(it)$$

$$\textit{Donc} \qquad \mathbb{E}(\exp(itS_n)) = \Pi_m(1 + p_{n,m}(\exp(it) - 1)).$$

On utilise la propriété suivante : Soient (z_i), (w_i) des complexes de module $\leq \theta$ alors

$$\left|\Pi_{m=1}^n z_m - \Pi_{m=1}^n w_m\right| \leq \theta^{n-1}\sum_{m=1}^{n} |z_m - w_m|$$

qui se prouve par recurrence : $\left|\Pi_{m=1}^n z_m - \Pi_{m=1}^n w_m\right| \leq \left|z_1\Pi_{m=2}^n z_m - z_1\Pi_{m=2}^n w_m\right| + \left|z_1 - w_1\right| \left|\Pi_{m=2}^n w_m\right| \leq \theta\, \theta^{n-2}\sum_{m=2}^{n} |z_m - w_m| + \theta^{n-1}\left|z_1 - w_1\right|$. On utilisera aussi que pour tout complexe b avec $|b| \leq 1$, $\left|\exp(b) - (1+b)\right| \leq |b|^2(e^1 - 1 - 1) \leq |b|^2$. On calcule alors

$$\begin{aligned}
&\left|E(\exp(itS_n)) - \exp\Big(\sum_{m=1}^{n} p_{n,m}(e^{it} - 1)\Big)\right| = \left|\Pi_m(1 + p_{n,m}(\exp(it) - 1)) - \Pi_m \exp\left(p_{n,m}(e^{it} - 1)\right)\right|\\
&\leq \sum_{m=1}^{n} |1 + p_{n,m}(\exp(it) - 1) - \exp\left(p_{n,m}(e^{it} - 1)\right)| \leq \sum_{m=1}^{n} p_{n,m}^2|\exp(it) - 1|^2\\
&\leq 4 \max_{1\leq m\leq n} p_{n,m} \sum_{m=1}^{n} p_{n,m} \ \to 0.
\end{aligned}$$

On trouve donc $\lim_n \mathbb{E}(\exp(itS_n)) = \exp\left(\lambda(e^{it}-1)\right)$ *et et pour une loi de Poisson Z* $\mathbb{E}(\exp(itZ)) = \exp(-\lambda)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\exp(itk) = \exp(-\lambda)\sum_{k=0}^{\infty}\frac{[\exp(it)\lambda]^k}{k!}$.

Exercice 4.7. *Théorème de Weierstrass : Polynômes de Bernstein, Williams, probability with martingales* Soit $x \in [0,1]$, et (X_n^x) une suite de v.a. i.i.d. de loi de Bernoulli $\mathcal{B}(x)$, et on note $S_n^x = \sum_{k=1}^n X_k^x$.

1. Quelle est la loi de S_n^x ? En déduire que $\forall\delta > 0$
$$\mathbb{P}\left(\left|\frac{1}{n}S_n^x - x\right| \geq \delta\right) \leq \frac{1}{4n\delta^2}.$$
2. Soit $f : [0,1] \to \mathbb{R}$ une fonction continue, on définit $P_n^f(x) := \mathbb{E}\left(f(n^{-1}S_n^x)\right)$. Montrer que P_n^f est un polynôme de degré n dont on donnera l'expression.
3. Soit $M := \sup_{x\in[0,1]}|f(x)|$ et $\epsilon > 0$. Montrer qu'alors
$$|P_n^f(x) - f(x)| \leq \epsilon + 2M\ \mathbb{P}(|S_n^x/n - x| > \delta).$$
4. En déduire que P_n^f converge *uniformément* vers f (polynômes dits de Bernstein). Montrer également que si f est lipschitzienne, alors
$$\|P_n^f - f\|_\infty \leq \frac{c(f)}{n^{1/3}}.$$

Solutions :

1. *La loi de* S_n^x *est une loi binomiale :* $\mathbb{P}(S_n^x = j) = \binom{n}{j}x^j(1-x)^{n-j}$. *On applique l'inégalité de Chebychev :*
$$\mathbb{P}\left(\left|\frac{1}{n}S_n^x - x\right| \geq \delta\right) = \mathbb{P}\left(|S_n^x - nx| \geq n\delta\right) \leq \frac{\mathbb{E}\left(S_n^x - nx\right)^2}{n^2\delta^2} = \frac{\mathrm{Var}\left(S_n^x\right)}{n^2\delta^2} \leq \frac{1}{4n\delta^2}.$$
On a utilisé le fait que la variance de S_n^x *est égale à* $nx(1-x)$ *et que* $x(1-x) \leq \frac{1}{4}$ *si* $x \in [0,1]$.
2. *On vérifie aisément que*
$$P_n^f(x) = \sum_{k=0}^n \mathbb{P}(S_n^x = k)\ f(k/n) = \sum_{k=0}^n \binom{n}{k}\ x^k(1-x)^{n-k} f(k/n)\cdot$$
3. *Par continuité uniforme,* $\forall\epsilon > 0,\quad \exists\delta > 0\quad \text{t.q. } \forall x,y,\ |x-y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$. *On calcule* $f(x) - P_n^f(x)$ *:*
$$\begin{aligned}&\sum_{k=0}^n \binom{n}{k}\ x^k(1-x)^{n-k}\left(f(x) - f(k/n)\right)\\ =\ &\sum_{k:\left|k/n-x\right|<\delta}\binom{n}{k}x^k(1-x)^{n-k}(f(x)-f(k/n)) + \sum_{k:\left|k/n-x\right|\geq\delta}\binom{n}{k}x^k(1-x)^{n-k}(f(x)-f(k/n))\\ :=\ &\Sigma_1 + \Sigma_2.\end{aligned}$$

$$\begin{aligned}
\left|\Sigma_1\right| &\leq \sum_{k:\left|k/n-x\right|<\delta} \binom{n}{k} x^k(1-x)^{n-k} \left|f(x)-f(k/n)\right| \leq \epsilon \sum_{k:\left|k/n-x\right|<\delta} \binom{n}{k} x^k(1-x)^{n-k} \leq \epsilon. \\
\left|\Sigma_2\right| &\leq \sum_{k:\left|k/n-x\right|\geq\delta} \binom{n}{k} x^k(1-x)^{n-k} \left|f(x)-f(k/n)\right| \\
&\leq 2M \sum_{k:\left|k/n-x\right|\geq\delta} \binom{n}{k} x^k(1-x)^{n-k} = 2M\mathbb{P}\left(\left|\frac{S_n}{n}-x\right|\geq\delta\right).
\end{aligned}$$

4. *On obtient donc*

$$\sup_{x\in[0,1]} \left|f(x)-P_n^f(x)\right| \leq \epsilon + \frac{M}{2n\delta^2}$$

qu'on peut rendre aussi petit que l'on veut.
Si $f \in \mathrm{Lip}([0,1])$, *on a* $|f(x)-f(k/n)| \leq \mathrm{lip}(f)\ |x-k/n|$. *On alors* $\epsilon = \mathrm{lip}(f)\delta$. *L'erreur de l'approximation est maintenant* $\mathrm{lip}(f)\delta + \frac{M}{2n\delta^2}$. *La valeur optimale de* δ *s'obtient en égalant les deux termes en compétition, ce qui donne*

$$\delta = \left(\frac{M}{2\mathrm{lip}(f)n}\right)^{1/3}$$

et on trouve $c(f) = (M/2)^{1/3}(\mathrm{lip}(f))^{2/3}$.

Exercice 4.8. Exercice 14.8, Ouvrard, lemme de Slutsky
Soient sur le même espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ deus suites $(X_n)_n$ et $(Y_n)_n$ de variables aléatoires qui convergent en loi respectivement vers X et une constante y_0.

1) a) En utilisant le théorème de Stone-Weierstrass, montrer que l'ensemble

$$\mathcal{H} = \{(x,y) \to f(x)g(y) | f,g \in C_0(\mathbb{R})\}$$

est total dans $C_0(\mathbb{R}^2)$ ($C_0(\mathbb{R}^k)$ est l'espace des fonctions continues sur $\mathbb{R}^k$ tendant vers 0 en l'infini).

b) On rappelle que $(Y_n)_n$ converge en loi vers une constante donc converge en probabilité vers cette même constante (exercice 7.4.3). Soit f et g deux éléments de $C_0(\mathbb{R})$. Montrer que

$$\begin{aligned}
|\mathbb{E}(f(X_n)g(Y_n)) - \mathbb{E}(f(X)g(y_0))| &\leq \|f\|_\infty[\epsilon + 2\|g\|_\infty \mathbb{P}(|g(Y_n)-g(y_0)|>\epsilon)] \\
&+ \|g\|_\infty|\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))|
\end{aligned}$$

c) Conclure que (X_n,Y_n) converge en loi vers (X,y_0).

Exercice 4.9. Stabilité du caractère gaussien
Soit $(X_n)_n$ une suite de variables aléatoires gaussiennes qui convergent en moyenne quadratique vers une variable aléatoire X. Montrer que X est une variable gaussienne.
(*ind.* : on remarquera que la convergence en moyenne quadratique implique la convergence en loi, on passera aux fonctions caractéristiques et utilisera le théorème de convergence de Lévy.)

Exercice 4.10. (Théorème de Helly-Bray), Williams *Probability with martingales*, pp. 183-184. Soit $(F_n)_n$ une suite de fonctions de répartitions sur $\mathbb{R}$. Le but de l'exercice est de montrer qu'il existe une fonction croissante continue à droite F à valeurs dans $[0,1]$ et une sous-suite $(F_{n_k})_k$ telle que

$$\lim_k F_{n_k}(x) = F(x) \text{ en tout point de continuité } x \text{ de F.}$$

1) En utilisant le principe d'extraction diagonale, montrer qu'il existe une sous-suite $(F_{n_k})_k$ telle que pour tout $q \in \mathbb{Q}$, il existe un réel $H(q) \in [0,1]$ telle que

$$\lim_{k\to+\infty} F_{n_k}(q) = H(q).$$

2) Pour $x \in \mathbb{R}$, on définit $F(x)$ par

$$F(x) = \lim_{\substack{q\in\mathbb{Q}, q>x \\ q\to x}} H(q).$$

Montrer que F est continue à droite.

3) Conclure.

Exercice 4.11. Convergence en loi et fonction de répartition, Ouvrard, pp. 318-319
Soit $(X_n)_n$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On note F_n la fonction de répartition de X_n et F celle de X. On a le critère suivant :

$$X_n \Longrightarrow X \quad \text{ssi} \quad \lim_{n\to+\infty} F_n(x) = F(x)$$

en tout point de continuité x de F.

1) Pour la condition nécessaire, remarquer que si x est un point de continuité de F alors $(-\infty; x]$ est un ensemble de $\mathbb{P}_X$-continuité. (C'est-à-dire que $\mathbb{P}_X(\{x\}) = 0$ et donc $\mathbb{P}_X(]-\infty,\ y]) \to \mathbb{P}_X(]-\infty, x])$ quand $y \to x$.) Rappel : si μ_n converge étroitement vers μ, alors $\lim_n \mu(A_n) = \mu(A)$ pour tout borélien de μ-continuité.

2) Pour la condition suffisante, on se donne un $f \in C_0(\mathbb{R})$ et $\epsilon > 0$.

a) Se rappeler pourquoi l'ensemble des points de discontinuité de F est au plus dénombrable.

b) Justifier l'existence d'une fonction $g = \sum_{j=1}^k \alpha_j 1_{]a_j;b_j]}$ où $a_j < b_j \le a_{j+1} < b_{j+1}$ et a_j et b_j sont des points de continuité de F_X et telle que $\| f - g \|_\infty \le \epsilon$. (*ind.* : on pourra se servir du fait que f est uniformément continue).

c) Montrer que $\lim_{n\to+\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$ et conclure. *Indication :* $\mathbb{E}g(X_n) = \int_{\mathbb{R}} g d\mathbb{P}_{X_n} = \sum_j \alpha_j (F_n(b_j) - F_n(a_j)) \to \sum_j \alpha_j (F(b_j) - F(a_j)) = \mathbb{E}g(X)$.

Exercice 4.12. Représentation de Skorokhod d'une fonction de répartition, Williams *Probability with martingales*, p. 34.

On va montrer que $F: \mathbb{R} \to [0,1]$ est la fonction de répartition d'une variable aléatoire réelle si et seulement si F est croissante, continue à droite et si $\lim_{z\to-\infty} F(z) = 0$, $\lim_{z\to+\infty} F(z) = 1$.

1) Dessiner X^+ et X^- (cf. qusetion 3 pour la définition) pour les fonctions F suivantes : $F = \mathbb{1}_{\mathbb{R}^+}$, et $F(z) = 0$ pour $z \le 0$, $F(z) = z/2$ pour $0 \le z \le 1$, $F(z) = 1/2$ pour $1 \le z < 2$, $F(z) = 1$ pour $z \ge 2$.

2) Montrer que ces conditions sont nécessaires.

3) Supposons que F ait les quatre propriétés précédentes. On pose pour $\omega \in [0,1]$ muni de la tribu des boréliens et de la mesure de Lebesgue,

$$X^+(\omega) =: \inf\{z: F(z) > \omega\} = \sup\{y: F(y) \le \omega\},$$

$$X^-(\omega) =: \inf\{z: F(z) \ge \omega\} = \sup\{y: F(y) < \omega\}.$$

Remarquer que $(z > X^-(\omega)) \Rightarrow (F(z) \ge \omega)$. En utilisant la continuité à droite de F montrer que

$$F(X^-(\omega)) \ge \omega \text{ et que } (X^-(\omega) \le c) \Rightarrow (\omega \le F(X^-(\omega)) \le F(c)).$$

Déduire que

$$\mathbb{P}(X^- \le c) = F(c).$$

4) On va montrer que X^+ a aussi pour fonction de répartition F et que $\mathbb{P}(X^+ = X^-) = 1$. Expliquer successivement pourquoi $(\omega < F(c)) \Rightarrow (X^+(\omega) \le c)$, $F(c) \le \mathbb{P}(X^+ \le c)$, $X^- \le X^+$. En remarquant que

$$\mathbb{P}(X^- \neq X^+) = \bigcup_{c\in\mathbb{Q}} \{X^- \le c < X^+\},$$

déduire en utilisant le 2) que $X^+ = X^-$ presque partout. Conclure.

Solutions : 3) $z > X(\omega) \Rightarrow F(z) \ge \omega$ par définition de $X^-(\omega)$. On a donc $\forall z > X^-(\omega)$, $F(z) \ge \omega$ et comme F est continue à droite, $F(X^-(\omega)) \ge \omega$. Si $X^-(\omega) \le c$, on a $F(X^-(\omega)) \le F(c)$ et donc $\omega \le F(c)$. En résumé, $X^-(\omega) \le c \Rightarrow \omega \le F(c)$ et réciproquement $\omega \le F(c) \Rightarrow c \ge X^-(\omega)$ par définition de $X^-(\omega)$. Donc $X^-(\omega) \le c \Leftrightarrow \omega \le F(c)$. En conséquence, $\mathbb{P}(X^- \le c) = \mathbb{P}(\omega \le F(c)) = F(c) - 0 = F(c)$.

4) $(\omega < F(c)) \Rightarrow (X^+(\omega) \le c)$ par définition de X^+. $F(c) \le \mathbb{P}(X^+ \le c)$ s'en déduit immédiatement puisque $\mathbb{P}(\omega < F(c)) = F(c)$. $X^- \le X^+$ est évident. On a $\mathbb{P}(X^- \le c \le X^+) = \mathbb{P}(\{X^- \le c\}\setminus\{X^+ \le c\}) = \mathbb{P}(X^- \le c) - \mathbb{P}(X^+ \le c) \le F(c) - F(c) = 0$. Donc $\mathbb{P}(X^- \neq X^+) = 0$. On obtient $X^+ = X^-$ p.s., ce qui implique que X^+ et X^- ont tous deux F comme fonction de répartition.

Remarque 4.1. *Etant donné une fonction de répartition F, ce théorème montre comment simuler une variable aléatoire X^- de fonction de répartition F. Bien entendu, cette méthode suppose que l'on sache calculer X^- à partir de F...*

Exercice 4.13. La représentation de Skorokhod, Williams *Probability with martingales*, p.182.

1) Déduire des exercices précédents que si X_n, de fonction de répartition F_n, tend presque sûrement vers X, de fonction de répartition F, alors $F_n(x) \to F(x)$ en tout point x de continuité de F. Notre but est de montrer une sorte de réciproque :

Théorème 4.1. *Si F_n et F sont des fonctions de répartition et si $F_n \to F$ en tout point de continuité de F, alors il existe des v.a. X_n et X sur* $([0,1],\ \mathcal{B},\ \mathbb{P})$ *(où $\mathbb{P}$ est la mesure de Lebesgue) telles que $F = F_X$, $F_n = F_{X_n}$ et $X_n \to X$ p.s.*

2) On pose $X^+ =: \inf\{z : \ F(z) > \omega\}$, $X^- =: \inf\{z : \ F(z) \geq \omega\}$ et on définit de même X_n^+ et X_n^-. On rappelle (exercice **??**) que X^+ et X^- ont pour fonction de répartition F et que $\mathbb{P}(X^+ = X^-) = 1$. On fixe $\omega \in [0,1]$. Soit z un point non atomique pour F vérifiant $z > X^+(\omega)$. Montrer que $\limsup_n X_n^+(\omega) \leq z$ et en déduire que $\limsup_n X_n^+(\omega) \leq X^+(\omega)$.

3) Adapter l'argument précédent pour montrer que $\liminf_n X_n^-(\omega) \geq X^-(\omega)$ et conclure.

2) Par définition de X^+, $F(z) > \omega$. Donc pour n grand, $F_n(z) > \omega$. Par la définition de X_n^+, on a alors $X_n^+ \leq z$ et donc $\limsup X_n^+(\omega) \leq z$*. Comme ceci est vrai $\forall z > X^+(\omega)$, on obtient* $\limsup_n X_n^+(\omega) \leq X^+(\omega)$.