Prérequis. Caractérisation de la convergence en loi par les fonctions de répartition, la fonction caractéristique caractérise la loi

Définition 0.0.1. Une suite $(\mu_n)_{n\in\mathbf{N}}$ de probabilités sur $\mathbf{R}^n$ *converge étroitement* vers μ si pour tout $\phi : \mathbf{R}^n \to \mathbf{R}$ continue bornée on a :

$$\int_{\mathbf{R}^n} \phi \, \mathrm{d}\mu_n \to \int_{\mathbf{R}^n} \phi \, \mathrm{d}\mu.$$

Autrement dit, la convergence étroite des lois de variables aléatoires correspond à la convergence en loi de ces variables.

Définition 0.0.2. Une suite $(\mu_n)_{n\in\mathbf{N}}$ de probabilités sur $\mathbf{R}^n$ est *tendue* si pour tout $\varepsilon > 0$ il existe un compact K tel que pour tout $n \in \mathbf{N}$:

$$\mu_n(K) \geqslant 1 - \varepsilon.$$

Dans la suite on se place dans $\mathbf{R}$.

Théorème 0.0.3 (de Helly)**.** *Soit (F_n) une suite de fonctions de répartition. Alors il existe une extraction $(\alpha_n)_{n\in\mathbf{N}}$ et une fonction $F : \mathbf{R} \to [0,1]$ croissante et continue à droite telle qu'en tout point x de continuité de F on ait :*

$$F_{\alpha_n}(x) \to F(x).$$

Démonstration. Pour $q \in \mathbf{Q}$, la suite $(F_n(q))_{n\in\mathbf{N}}$ est dans $[0,1]$, on peut donc en extraire une sous-suite qui converge vers un $F_{\mathbf{Q}}(q) \in [0,1]$. Par extraction diagonale, il existe une extraction $(\alpha_n)_{n\in\mathbf{N}}$ telle que pour tout $q \in \mathbf{Q}$, $F_{\alpha_n}(q)$ converge vers $F_{\mathbf{Q}}(q)$. On pose alors pour $x \in \mathbf{R}$:

$$F(x) = \inf_{q\in Q, x<q} F_{\mathbf{Q}}(q).$$

On définit ainsi une fonction $F : \mathbf{R} \to [0,1]$ croissante. Pour la continuité à droite, soient $x \in \mathbf{R}$ et $\varepsilon > 0$. On trouve un $q > x$ tel que $F_{\mathbf{Q}}(q) \leqslant F(x) + \varepsilon$, de sorte que pour tout $y \in [x,q]$ on ait :

$$F(x) \leqslant F(y) \leqslant F_{\mathbf{Q}}(q) \leqslant F(x) + \varepsilon.$$

Il reste à démontrer le dernier point, on fixe alors x un point de continuité de F, et $\varepsilon > 0$. Par continuité on trouve $y < x$ tel que :

$$F(x) - \varepsilon \leqslant F(y).$$

On n'a plus qu'à intercaler $y < r < x < s$ avec $r, s \in \mathbf{Q}$ et $F_{\mathbf{Q}}(s) \leqslant F(x) + \varepsilon$ pour obtenir :

$$\begin{aligned} F(x) - \varepsilon \leqslant F(y) \leqslant F_{\mathbf{Q}}(r) = \lim F_{\alpha_n}(r) \leqslant \liminf F_{\alpha_n}(x) \\ \leqslant \limsup F_{\alpha_n}(x) \leqslant \limsup F_{\alpha_n}(s) = F_{\mathbf{Q}}(s) \leqslant F(x) + \varepsilon. \end{aligned}$$

Le résultat est démontré en faisant tendre ε vers 0. $\square$

Théorème 0.0.4 (de Prokhorov). *De toute suite tendue $(\mu_n)_{n\in\mathbf{N}}$ sur $\mathbf{R}$ on peut extraire une sous-suite qui converge étroitement.*

Démonstration. Soit F_n la fonction de répartition de μ_n. On applique le théorème de Helly pour obtenir une fonction F, et il suffit alors de montrer que c'est une fonction de répartition (la convergence étroite des lois est la convergence en loi des variables aléatoires), puisque l'on sait déjà qu'après extraction il y a convergence en tout point de continuité. Comme F est croissante et continue à droite, il manque seulement de montrer que F tend vers 0 (resp. 1) en $-\infty$ (resp. en $+\infty$).

On fixe $\varepsilon > 0$ et l'on obtient par tension un $M > 0$ tel que pour tout $n \in \mathbf{N}$:

$$\mu_n([-M, M]) \geqslant 1 - \varepsilon.$$

Comme F a un nombre au plus dénombrable de discontinuités, on suppose que M est un point de continuité de F, de sorte à pouvoir écrire en passant à la limite :

$$F(M) - F(-M) \geqslant 1 - \varepsilon.$$

Il ne reste plus qu'à faire tendre ε vers 0 pour obtenir le résultat. □

Théorème 0.0.5 (de Lévy). *Soit $(\mu_n)_{n\in\mathbf{N}}$ une suite de probabilités sur $\mathbf{R}$. On note ϕ_n la fonction caractéristique de μ_n. Si $(\phi_n)_{n\in\mathbf{N}}$ converge simplement vers une fonction ϕ continue en 0 alors celle-ci est la fonction caractéristique d'une probabilité μ vers laquelle $(\mu_n)_{n\in\mathbf{N}}$ converge étroitement.*

Démonstration. On veut appliquer le théorème de Prokhorov, donc on montre que $(\mu_n)_{n\in\mathbf{N}}$ est tendue. Pour $u > 0$, on a par Fubini :

$$\frac{1}{2u}\int_{-u}^{u}(1-\phi_n) = \int_{\mathbf{R}}\left(1 - \frac{\sin(ux)}{ux}\right)\mathrm{d}\mu_n(x) \geqslant \left(1 - \frac{2}{\pi}\right)\mu_n\left(\mathbf{R}\setminus\left[-\frac{\pi}{2u}, \frac{\pi}{2u}\right]\right).$$

Or, $\frac{1}{2u}\int_{-u}^{u}(1-\phi_n)$ converge vers $\frac{1}{2u}\int_{-u}^{u}(1-\phi)$ qui elle-même converge, lorsque $u \to 0$, vers 0 (par continuité de ϕ en 0). Ainsi pour tout $\varepsilon > 0$ il existe $u > 0$ tel qu'à partir d'un certain rang :

$$\mu_n\left(\mathbf{R}\setminus\left[\frac{-\pi}{2u}, \frac{\pi}{2u}\right]\right) \leqslant \varepsilon.$$

Donc la suite $(\mu_n)_{n\in\mathbf{N}}$ est tendue, on applique le théorème de Prokhorov pour obtenir une valeur d'adhérence μ. Mais comme toute valeur d'adhérence de $(\mu_n)_{n\in\mathbf{N}}$ a pour fonction caractéristique ϕ et que la fonction caractéristique caractérise la loi, μ est l'unique valeur d'adhérence de $(\mu_n)_{n\in\mathbf{N}}$, donc il y a convergence étroite. □

Remarques.

— C'est trop long pour quinze minutes, mais on peut se contenter de donner les idées du théorème de Prokhorov qui ne contient pas vraiment de raisonnement intéressant. La preuve intéressante se trouve dans le théorème de Helly et le théorème de Lévy, et on peut adapter en fonction de la leçon.

- Dans le théorème de Helly on utilise la séparabilité de $\mathbf{R}$ et une extraction diagonale pour obtenir une convergence à extraction près sur tout l'espace. Le théorème de Prokhorov est vrai de manière générale sur les espaces métriques séparables, mais la démonstration est bien sûr différente.
- Le théorème de Lévy sert à démontrer le théorème central limite : en posant $Y_i = \frac{X_i - \mu}{\sigma}$, on a le développement limité :

$$\phi_Y(t) = 1 + \frac{t^2}{2} + o(t^2)$$

  et la moyenne $Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ vaut $\sum_{i=1}^n Y_i/\sqrt{n}$, d'où :

$$\phi_{Z_n}(t) = \left(\phi_Y\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 + \frac{t^2}{2n} + o(t^2)\right)^n \to \mathrm{e}^{-t^2/2}$$

  et le théorème de Lévy conclut.
- Les fonctions $F : \mathbf{R} \to [0,1]$ croissantes continues à droite et de limite nulle en $-\infty$ et 1 en $+\infty$ sont des fonctions de répartition. Plus précisément, c'est la fonction de répartition de $F^{-1}(U)$ où F^{-1} est l'inverse généralisé de F et où U est une variable aléatoire de loi uniforme sur $]0,1[$.
- Le théorème de Prokhorov sert aussi en théorie du transport optimal, pour démontrer qu'il existe une solution au problème de Kantorovich.

Recasages.

- 203 : Le théorème de Prokhorov est très clairement un théorème de compacité à la Bolzano-Weierstrass. Pour l'utilisation de la compacité, on est en plein dans le thème : on utilise le théorème de Prokhorov pour démontrer le théorème de Lévy qui est très utile en probabilités.
- 228 : Il faut mettre l'accent sur le théorème de Helly, mais c'est vraiment dans le thème (avec plusieurs fois l'utilisation de la définition de la continuité et des limites inférieure et supérieure), surtout si l'on fait une petite partie à propos des fonctions de répartition par exemple.
- 229 : Les fonctions de répartition sont monotones, et on l'utilise intensivement dans la démonstration du théorème de Helly : c'est comme ça que l'on passe de $F_{\mathbf{Q}}$ à F. On utilise aussi la monotonie dans la démonstration du théorème de Prokhorov pour dire que M peut être choisi comme un point de continuité : les fonctions monotones n'ont qu'un nombre au plus dénombrable de discontinuités.
- 261 : Rien de mieux que de démontrer le théorème de Lévy dans cette leçon, puisqu'il énonce justement une caractérisation de la convergence en loi. On pourra passer plus rapidement sur le théorème de Helly et sur celui de Prokhorov.
- 262 : De même qu'en 261, c'est un théorème de convergence en loi. On peut même énoncer et démontrer le théorème central limite dans la foulée s'il reste du temps, quitte à éclipser encore plus les théorèmes de Helly et Prokhorov.