Théorème Central Limite, stabilité et principe de Lindeberg

Martin Donati, encadré par Julien Poisat et Emeric Bouin (CEREMADE, Paris-Dauphine)

30 Mai - 1^{er} Juillet 2016

Table des matières

Introduction au principe de Lindeberg

Le principe de Lindeberg, méthode issue de l'article [3], reprise par [6], est à l'origine une méthode pour démontrer le théorème Central Limite, sans avoir recours aux outils dérivant de l'analyse de Fourier. Le théorème Central Limite est fondamentalement une preuve d'une convergence en loi. Ainsi, ce qui va nous intéresser, c'est d'étudier des convergences du type :

$$\mathbb{E}(f(S_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(Y)), \tag{1}$$

où

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \tag{2}$$

est la somme qui nous intéresse dans ce théorème, dont on spécifiera tout au long de ce rapport les hypothèses plus précisément, et

$$Y = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \tag{3}$$

suit une loi normale centrée réduite que l'on peut écrire comme somme de lois normales centrées réduites indépendantes par stabilité de la loi normale.

Posons $Z_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$; $Z_n = X$.

Ainsi, l'on a :

$$|\mathbb{E}(f(X)) - \mathbb{E}(f(Y))| = \left| \sum_{i=1}^{n} \mathbb{E}\left(f\left(\frac{Z_i}{\sqrt{n}}\right) - f\left(\frac{Z_{i-1}}{\sqrt{n}}\right)\right)\right|. \tag{4}$$

L'idée principale sera donc, par un développement de Taylor, de contrôler l'erreur $\mathbb{E}\left(f\left(\frac{Z_i}{\sqrt{n}}\right) - f\left(\frac{Z_{i-1}}{\sqrt{n}}\right)\right)$.

Cette méthode s'étend remarquablement bien à des théorèmes de convergence beaucoup plus forts que le théorème Central Limite, et avec des hypothèses moins restrictives que, notamment, l'identique distribution des variables. En particulier, [4] donne une majoration de la quantité $|\mathbb{E}(g(f((X_1, \ldots, X_n))) - \mathbb{E}(g(f((Y_1, \ldots, Y_n)))|$, qui permet d'étendre les résultats de convergence à bien plus que seulement des sommes de variables aléatoires.

Enfin, nous étudierons également dans ce rapport les différentes lois possibles comme limites de suites de variables aléatoires indépendantes et identiquement distribuées.

Chapitre 1

Convergence en loi de séries de variables aléatoires de L^2

Dans ce chapitre, toutes les variables aléatoires considérées seront supposées réelles, centrées, et de variance finie. $\mathbb{R}$ est muni de la tribu Borélienne et de la mesure de Lebesgue. Nous noterons $\|f\| = \sup_{\mathbb{R}} |f(x)|$.

1.1 Théorème Central Limite et principe de Lindeberg

Nous allons ici nous intéresser à la démonstration du théorème Central Limite donnée par Lindeberg [1] en 1922 par une méthode qui porte désormais le nom de principe de Lindeberg. Plus précisément, nous suivrons l'article [6].

1.1.1 Mode de convergence

Avant toute chose, il est essentiel de définir les différents modes de convergence pour les objets mathématiques que l'on va étudier. Plus précisément nous allons parler de la convergence en loi, qui est une convergence des mesures associées aux variables aléatoires.

Définition 1.1 (Convergence étroite)**.** *Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilités sur $\mathbb{R}$. Elle converge étroitement vers μ mesure de probabilité sur $\mathbb{R}$ si et seulement si :*

$$\forall f \in C_b(\mathbb{R}), \quad \int_{\mathbb{R}} f(x)\mu_n(\mathrm{d}x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} f(x)\mu(\mathrm{d}x), \tag{1.1}$$

où $C_b(\mathbb{R})$ désigne l'ensemble des fonctions continues bornées de $\mathbb{R}$ dans $\mathbb{R}$.

Définition 1.2 (Convergence en loi)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires à valeurs réelles ; elle converge en loi vers une variable aléatoire X si $(\mathbb{P}_{X_n})_{n\in\mathbb{N}}$ converge étroitement vers $\mathbb{P}_X$, c'est à dire si :*

$$\forall f \in C_b(\mathbb{R}), \quad \mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X)). \tag{1.2}$$

L'article [6] s'appuie sur une autre définition de la convergence en loi, grâce à la fonction de répartition d'une variable aléatoire :

Définition 1.3 (Fonction de répartition)**.** *Soit X une variable aléatoire à valeurs réelles, on définit :*

$$F_X(x) = \mathbb{P}(X \leq x), \quad \forall x \in \mathbb{R}. \tag{1.3}$$

La fonction F_X, appelée fonction de répartition, est croissante et continue à droite, et est discontinue en x si et seulement si $\mathbb{P}(X = x) > 0$: on dit alors que x est un atome de la loi de X. L'ensemble des atomes étant dénombrable, la fonction F est entièrement déterminée par la donnée des $F(x)$ où x est un point de continuité.

La convergence simple des fonctions de répartition est équivalente à la convergence en loi. Nous retiendrons l'implication suivante :

Propriété 1.1. *Soit (X_n) une suite de variables aléatoires réelles, et X une variable aléatoire réelle de fonction de répartition F_X et A l'ensemble de ses points de discontinuité. Alors*

$$\Big(\forall x \notin A, F_{X_n}(x) \underset{n\to\infty}{\longrightarrow} F_X(x)\Big) \Longrightarrow X_n \xrightarrow[n\to\infty]{\text{loi}} X. \tag{1.4}$$

1. Voir [3].

On pourra trouver une démonstration de l'équivalence dans [1], chapitre 5, théorème (25.8).

Nous allons utiliser cette propriété presque exclusivement pour montrer la suivante, dont le but va être de réduire l'ensemble des fonctions à tester pour la convergence en loi.

Définition 1.4. *Notons C l'ensemble des fonctions uniformément continues et bornées sur $\mathbb{R}$, et notons C^2 l'ensemble des fonctions deux fois dérivables telles que $f, f', f'' \in C$.*

Propriété 1.2. *Soit (X_n) une suite de variables aléatoires à valeurs dans $\mathbb{R}$, telles que :*

$$\forall f \in C^2, \quad \mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X)). \tag{1.5}$$

Alors $X_n \xrightarrow[n\to\infty]{loi} X$.

Démonstration. Soit y un point en lequel F_X est continue. Fixons $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$F_X(y+\delta) - F_X(y-\delta) < \varepsilon. \tag{1.6}$$

Construisons alors f et g deux fonctions de l'espace C^2 telles que :
- $0 \leq f(x) \leq g(x) \leq 1$,
- $f(x) = 1$ si $x \leq y - \delta$,
- $f(x) = 0$ si $x \geq y$,
- $g(x) = 1$ si $x \leq y$,
- $g(x) = 0$ si $x \geq y + \delta$.

On a alors par construction :

$$F_X(y-\delta) \leq \mathbb{E}(f(X)) \leq F_X(y) \leq \mathbb{E}(g(X)) \leq F_X(y+\delta), \tag{1.7}$$

et également $\forall n \in \mathbb{N}$:

$$\mathbb{E}(f(X_n)) \leq F_{X_n}(y) \leq \mathbb{E}(g(X_n)). \tag{1.8}$$

Or par hypothèse, $\mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X))$ et $\mathbb{E}(g(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(g(X))$, donc :

$$\mathbb{E}(f(X)) \leq \liminf F_{X_n}(y) \leq \limsup F_{X_n}(y) \leq \mathbb{E}(g(X)), \tag{1.9}$$

et donc

$$F_X(y) - \varepsilon \leq \liminf F_{X_n}(y) \leq \limsup F_{X_n}(y) \leq F_X(y) + \varepsilon. \tag{1.10}$$

Ceci étant vrai pour tout $\varepsilon > 0$, on a finalement :

$$F_{X_n}(y) \underset{n\to\infty}{\longrightarrow} F_X(y), \tag{1.11}$$

et ce pour tout $y \notin A$. Ainsi, par la propriété 1.1, on a bien $X_n \xrightarrow[n\to\infty]{\text{loi}} X$. □

Ce résultat s'apparente à un résultat de densité, puisque finalement on a montré :

$$\left[\forall f \in C^2, \mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X))\right] \Longleftrightarrow \left[\forall f \in C_b, \mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X))\right]. \tag{1.12}$$

1.1.2 Opérateur de convolution

Définition 1.5. *Posons quelle que soit la variable aléatoire Z réelle : $T_Z \begin{cases} C & \to C \\ f & \mapsto (x \mapsto \mathbb{E}(f(Z+x))) \end{cases}$*

T est un opérateur linéaire bien défini de C dans C puisque d'une part :

$$\forall x \in \mathbb{R}, |\mathbb{E}(f(Z+x))| \leq \|f\|, \tag{1.13}$$

donc $\|T_Z(f)\| \leq \|f\|$ et d'autre part :

$$\forall x, y \in \mathbb{R}, |T_Z(f)(x) - T_Z(f)(y)| \leq \sup_z |f(x+z) - f(y+z)|, \tag{1.14}$$

donc $T_Z(f)$ est bornée et uniformément continue. De plus, $T_{Z_1}T_{Z_2} = T_{Z_1+Z_2}$, d'après le théorème de Fubini. En particulier T_{Z_1} et T_{Z_2} commutent, ce qui donne la propriété suivante.

Propriété 1.3. *L'opérateur T défini précédemment vérifie :*

$$\|(T_{Z_1}^n - T_{Z_2}^n)(f)\| \leq n\|(T_{Z_1} - T_{Z_2})(f)\| \tag{1.15}$$

Propriété 1.4 (Stabilité de la loi normale)**.** *Soit* (Y_n) *une suite de variables aléatoires indépendantes et suivant toutes une loi normale* $\mathcal{N}(0,1)$ *et soit* $Z_n = \frac{1}{\sqrt{n}} Y_n$*. Alors :*

$$T_Y = T_{Z_n}^n. \tag{1.16}$$

Ce résultat découle directement du calcul du produit de convolution de deux lois normales : si l'on prend deux variables aléatoires indépendantes X et Y avec $X \sim \mathcal{N}(0,\alpha)$ et $Y \sim \mathcal{N}(0,\beta)$, alors

$$X + Y \sim \mathcal{N}(0, \sqrt{\alpha^2 + \beta^2}). \tag{1.17}$$

1.1.3 Preuve du TCL

Théorème 1.1 (Central Limite)**.** *Soit* (X_n) *une suite de variables aléatoires réelles indépendantes et identiquement distribuées, centrées réduites. Alors*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1). \tag{1.18}$$

Démonstration. Posons $X = X_1$ et $\sigma = \frac{1}{\sqrt{n}}$. Notons $S_n = \displaystyle\sum_{k=1}^{n} \sigma X_k$.

Première étape : Réduction du problème

A l'aide de la propriété 1.2, il suffit de montrer qu'en notant Y une variable aléatoire suivant une loi normale $\mathcal{N}(0,1)$:

$$\forall f \in C^2, \quad \mathbb{E}(f(S_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(Y)). \tag{1.19}$$

Soit donc $f \in C^2$. Par récurrence, on a donc que $T_{S_n} = T_{\sigma X}^n$, et donc par les propriétés 1.3 et 1.4 :

$$\begin{aligned} |\mathbb{E}(f(S_n)) - \mathbb{E}(f(Y))| &= |T_{\sigma X}^n(f)(0) - T_{\sigma Y}^n(f)(0)| && (1.20)\\ &\leq \|T_{\sigma X}^n(f) - T_{\sigma Y}^n(f)\| && (1.21)\\ &\leq n\|(T_{\sigma X} - T_{\sigma Y})(f)\|. && (1.22) \end{aligned}$$

Deuxième étape : Développement de Taylor

Puisque f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$, on peut écrire la formule de Taylor avec reste intégral entre a et $a+x$ réels :

$$f(a+x) = f(a) + xf'(a) + \int_a^{a+x} (a+x-t) f''(t) \mathrm{d}t \tag{1.23}$$

Que l'on préfèrera écrire [2] :

$$f(a+x) = f(a) + xf'(a) + \frac{1}{2}x^2 f''(a) + \int_a^{a+x} (a+x-t)(f''(t) - f''(a)) \mathrm{d}t \tag{1.24}$$

Fixons $\varepsilon > 0$. Puisque f'' est uniformément continue, il existe $\delta > 0$ tel que :

$$\forall a \in \mathbb{R}, \quad |x| < \delta \Longrightarrow |f''(a+x) - f''(a)| \leq \varepsilon. \tag{1.25}$$

Mais pour toute variable aléatoire Z réelle de moyenne nulle et de variance σ^2, on a :

$$\begin{aligned} T_Z(f)(a) &= \int_{\mathbb{R}} \left(f(a) + zf'(a) + \frac{1}{2} z^2 f''(\eta_{a,z}) \right) \mathbb{P}_Z(\mathrm{d}z) && (1.26)\\ &= f(a) + \frac{\sigma^2}{2} f''(a) + \int_{\mathbb{R}} \left(\int_a^{a+z} (a+z-t)(f''(t) - f''(a)) \mathrm{d}t \right) \mathbb{P}_Z(\mathrm{d}z). && (1.27) \end{aligned}$$

Notons temporairement : $R(a,z) = \displaystyle\int_a^{a+z} (a+z-t)(f''(t) - f''(a)) \mathrm{d}t$. Le dernier terme de (1.27) se majore ainsi :

2. En remarquant que $\displaystyle\int_a^{a+x} (a+x-t) \mathrm{d}t = \frac{x^2}{2}$.

$$\left|\int_{\mathbb{R}} R(a,z)\mathbb{P}_Z(\mathrm{d}z)\right| \leq \left|\int_{|z|<\delta} R(a,z)\mathbb{P}_Z(\mathrm{d}z)\right| + \left|\int_{|z|\geq\delta} R(a,z)\mathbb{P}_Z(\mathrm{d}z)\right| \tag{1.28}$$

$$\leq \varepsilon \int_{|z|<\delta} \frac{1}{2} z^2 \mathbb{P}_Z(\mathrm{d}z) + \|f''\| \int_{|z|\geq\delta} z^2 \mathbb{P}_Z(\mathrm{d}z), \tag{1.29}$$

donc finalement :

$$\left|\int_{\mathbb{R}} \left(\int_a^{a+z} (a+z-t)(f''(t)-f''(a))\mathrm{d}t\right) \mathbb{P}_Z(\mathrm{d}z)\right| \leq \frac{\sigma^2}{2}\varepsilon + \|f''\| \int_{|z|\geq\delta} z^2 \mathbb{P}_Z(\mathrm{d}z). \tag{1.30}$$

Troisième étape : Majoration de (1.22)**.**

Lorsque l'on applique cette formule pour σX et σY, cette dernière intégrale s'écrit (si Z désigne maintenant X ou Y) :

$$\int_{|z|\geq\delta} z^2 \mathbb{P}_{\sigma Z}(\mathrm{d}z) = \sigma^2 \underbrace{\int_{|z|\geq\delta\sqrt{n}} z^2 \mathbb{P}_Z(\mathrm{d}z)}_{\underset{n\to\infty}{\longrightarrow} 0}, \tag{1.31}$$

puisque X et Y ont des variances 1, et donc en particulier cette intégrale est majorée par $\sigma^2\varepsilon$ pour n suffisamment grand. Finalement, on a pour n assez grand :

$$\begin{aligned} T_{\sigma X}(f)(a) - T_{\sigma Y}(f)(a) =& f(a) + \frac{\sigma^2}{2} f''(a) + \int_{\mathbb{R}} \left(\int_a^{a+z} (a+z-t)(f''(t)-f''(a))\mathrm{d}t\right) \mathbb{P}_X(\mathrm{d}x) \\ & - \left[f(a) + \frac{\sigma^2}{2} f''(a) + \int_{\mathbb{R}} \left(\int_a^{a+z} (a+z-t)(f''(t)-f''(a))\mathrm{d}t\right) \mathbb{P}_Y(\mathrm{d}y)\right], \end{aligned} \tag{1.32}$$

d'où, en injectant (1.30) :

$$\|(T_{\sigma X} - T_{\sigma Y})(f)\| \leq 2\left[\frac{\sigma^2}{2}\varepsilon + \|f''\| \int_{|z|\geq\delta} z^2 \mathbb{P}_Z(\mathrm{d}z)\right] \tag{1.33}$$

$$\leq \sigma^2 \varepsilon (1 + 2\|f''\|). \tag{1.34}$$

Finalement, on a, pour n assez grand

$$|\mathbb{E}(f(S_n)) - \mathbb{E}(f(Y))| \leq n|\|(T_{\sigma X} - T_{\sigma Y})(f)\| \tag{1.35}$$

$$\leq n\sigma^2\varepsilon(1 + 2\|f''\|) \tag{1.36}$$

$$\leq \varepsilon(1 + 2\|f''\|) \tag{1.37}$$

Et donc ceci étant valable quel que soit $\varepsilon > 0$, on a bien $\mathbb{E}(f(S_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(Y))$. Pour conclure, ceci est vrai quelle que soit $f \in C^2$, et par la propriété 1.2, $S_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$.

□

1.2 Cas non identiquement distribué

Dans toute la suite, on considère désormais une suite de variables aléatoires (X_n) indépendantes, de moyenne nulle, et notons $\mathbb{V}(X_n) = a_n^2 > 0$. On note également :

$$s_n = \left(\sum_{i=1}^n a_i^2\right)^{\frac{1}{2}}. \tag{1.38}$$

Définition 1.6 (Critère de Lindeberg)**.** *La suite (X_n) vérifie le critère de Lindeberg si et seulement si :*

$$\forall\delta > 0, \lim_{n\to\infty} \frac{1}{{s_n}^2} \sum_{i=1}^n \int_{|x|\geq\delta s_n} x^2 \mathbb{P}_{X_i}(\mathrm{d}x) = 0. \tag{1.39}$$

L'objectif va être d'étendre le théorème Central Limite au cas où les variables ne sont plus identiquement distribuées, mais à la place vérifient le critère de Lindeberg. Encore une fois, la méthode employée est reprise de [6].

1.2.1 Lemmes intermédiaires

Commençons par étendre la propriété 1.3 de l'opérateur T.

Propriété 1.5. *Si $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ sont deux suites de variables aléatoires, alors*

$$\|T_{X_1}T_{X_2}\dots T_{X_n}(f) - T_{Y_1}T_{Y_2}\dots T_{Y_n}(f)\| \leq \sum_{i=1}^{n} \|T_{X_i}(f) - T_{Y_i}(f)\| \tag{1.40}$$

Nous allons avoir également besoin d'un lemme relatif au critère de Lindeberg :

Propriété 1.6. *Soit (X_n) une suite de variables aléatoires indépendantes de moyenne nulle et de variance a_n^2, et (Y_n) une suite de variables aléatoires indépendantes et de même ditribution qu'une variable Y de moyenne nulle et de variance 1. Alors si (X_n) vérifie le critère de Lindeberg, (a_nY_n) le vérifie également.*

1.2.2 Théorème de Lindeberg

Théorème 1.2. *Soit (X_n) une suite de variables aléatoires indépendantes. Si (X_n) vérifie le critère de Lindeberg, alors :*

$$\left(\sum_{i=1}^{n} a_i^2\right)^{-\frac{1}{2}} \sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1). \tag{1.41}$$

Démonstration. Soit $S_n = \left(\sum_{i=1}^{n} a_i^2\right)^{-\frac{1}{2}} \sum_{i=1}^{n} X_i$, et notons $\sigma = \dfrac{1}{s_n} = \left(\sum_{i=1}^{n} a_i^2\right)^{-\frac{1}{2}}$. On note $T_{S_n} = T_{\sigma X_1}\dots T_{\sigma X_n}$, et $Y \sim \mathcal{N}(0,1)$.

La preuve se déroule comme celle du théorème Central Limite, en écrivant au départ, pour $f \in C^2$:

$$\begin{aligned}
|\mathbb{E}(f(S_n)) - \mathbb{E}(f(Y))| &= |T_{\sigma X_1}\dots T_{\sigma X_n}(f)(0) - T_{\sigma a_1 Y}\dots T_{\sigma a_n Y}(f)(0)| && (1.42)\\
&\leq \|T_{\sigma X_1}\dots T_{\sigma X_n}(f) - T_{\sigma a_1 Y}\dots T_{\sigma a_n Y}(f)\| && (1.43)\\
&\leq \sum_{i=1}^{n} \|T_{\sigma X_i}(f) - T_{\sigma a_i Y}(f)\|, && (1.44)
\end{aligned}$$

d'après la propriété 1.5.

On a pour n assez grand :

$$\sum_{i=1}^{n} \|T_{\sigma X_i}(f) - T_{\sigma a_i Y}(f)\| \leq \sum_{i=1}^{n} \left[\sigma^2 {a_i}^2 \varepsilon + \|f''\| \left(\int_{|\sigma x_i|\geq\delta} \sigma^2 {x_i}^2 \mathbb{P}_{X_i}(\mathrm{d}x_i) + \int_{a_i\sigma|y|\geq\delta} \sigma^2 {a_i}^2 y^2 \mathbb{P}_Y(\mathrm{d}y)\right)\right]. \tag{1.45}$$

Par hypothèse, (X_n) vérifie le critère de Lindeberg, et donc par la propriété 1.6, (a_nY_n) aussi (où $Y_n \sim \mathcal{N}(0,1)$), ce qui signifie que

$$\frac{1}{{s_n}^2}\sum_{i=1}^{n}\int_{|x_i|\geq s_n\delta} {x_i}^2 \mathbb{P}_{X_i}(\mathrm{d}x_i) \underset{n\to\infty}{\longrightarrow} 0 \quad \text{et} \quad \frac{1}{{s_n}^2}\sum_{i=1}^{n}\int_{a_i|y|\geq s_n\delta} {a_i}^2 y^2 \mathbb{P}_Y(\mathrm{d}y) \underset{n\to\infty}{\longrightarrow} 0. \tag{1.46}$$

Les conclusions sont les mêmes et l'on obtient :

$$S_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1). \tag{1.47}$$

□

1.2.3 Tableaux triangulaires

Ce théorème s'étend de façon encore plus générale aux tableaux triangulaires, dont nous allons tout de suite donner une définition.

Définition 1.7. $(X_{n,i})_{1\leq i\leq r_n, n\in\mathbb{N}}$ *est appelée tableau triangulaire dès que les variables de chaque étage $(X_{n,i})_{1\leq i\leq r_n}$ sont indépendantes, où pour tout entier n, $r_n \in \mathbb{N}^*$.*

Théorème 1.3. *Soit* $S_n = \left(\sum_{i=1}^{r_n} {a_{n,i}}^2\right)^{-\frac{1}{2}} \sum_{i=1}^{r_n} X_{n,i}$*, où* $(X_{n,i})$ *est un tableau triangulaire, dont les variables sont de variance* $a_{n,i}^2$*, et vérifient le critère de Lindeberg étendu :*

$$\forall \delta > 0, \quad \lim_{n\to\infty} \left(\sum_{i=1}^{r_n} a_{n,i}^2\right)^{-1} \sum_{i=1}^{r_n} \int_{|x|\geq \delta s_n} x^2 \mathbb{P}_{X_{n,i}}(\mathrm{d}x) = 0. \tag{1.48}$$

Alors $S_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$.

Démonstration. Une preuve de ce théorème est exposée dans [1], mais utilise les fonctions caractéristiques. Montrons brièvement que la preuve précédente s'adapte très bien dans ce cas.

Reprenons les étapes clés de la démonstration, avec $f \in C^2$, $\varepsilon > 0$ et $\delta > 0$ choisi par uniforme continuité de f'', et en notant $\sigma = \frac{1}{s_n}$:

$$\begin{aligned} |\mathbb{E}(f(S_n)) - \mathbb{E}(f(Y))| &\leq \sum_{i=1}^{r_n} \|T_{\sigma X_{n,i}}(f) - T_{\sigma a_{n,i} Y}(f)\| && (1.49)\\ &\leq \sum_{i=1}^{r_n} \left[\sigma^2 a_{n,i}^2 \varepsilon + \|f''\| \left(\int_{|\sigma x|\geq\delta} \sigma^2 x^2 \mathbb{P}_{X_{n,i}}(\mathrm{d}x) + \int_{a_{n,i}\sigma|y|\geq\delta} \sigma^2 a_{n,i}^2 y^2 \mathbb{P}_Y(\mathrm{d}y)\right)\right] && (1.50)\\ &\leq \varepsilon(1 + 2\|f''\|), && (1.51) \end{aligned}$$

pour n assez grand, grâce au critère de Lindeberg étendu. Les conclusions sont les mêmes et finalement, on obtient : $S_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$. □

1.3 Discussion sur le critère de Lindeberg

La première chose à vérifier, c'est que ce théorème qui veut généraliser le théorème Central Limite a bien comme corollaire ce dit théorème.

Pour prouver ce résultat, on est ramené à la simple propriété suivante, dont la preuve est directe :

Propriété 1.7. *Toute suite* (X_n) *de variables indépendantes identiquement ditribuées de moyenne nulle et de variance* σ^2 *vérifie le critère de Lindeberg.*

Considérons une suite de variables indépendantes. Notons $a_n^2 = \mathbb{V}(X_n)$. On trouvera dans [8], Theorem 2 p.520, que si $s_n \xrightarrow[n\to\infty]{} \infty$ et $\frac{a_n}{s_n} \xrightarrow[n\to\infty]{} 0$, alors le critère de Lindeberg est nécessaire et suffisant pour la convergence en loi vers la gaussienne.

Remarquons en particulier que :

$$0 < m \leq a_n \leq M \Longrightarrow s_n \xrightarrow[n\to\infty]{} \infty \text{ et } \frac{a_n}{s_n} \xrightarrow[n\to\infty]{} 0, \tag{1.52}$$

et que ceci est toujours vrai pour un tableau triangulaire de variables identiquement distribuées.

1.4 Vers l'intégration stochastique

Ici nous allons avoir recours au résultat sur les tableaux triangulaires. Soit $(X_{n,i})_{1\leq i\leq n}$ un tableau triangulaire de variables aléatoires, telles que $a_{n,i} = f\left(\frac{i}{n}\right)$, où $f \in C^0([0,1],]0,+\infty[)$, et supposons-le vérifier le critère de Lindeberg étendu. La fonction f est continue sur $[0,1]$ qui est compact, donc est bornée et atteint ses bornes. En particulier, on a donc $0 < m \leq a_{i,n} \leq M$ et tout tableau obtenu en pondérant des variables identiquement distribuées vérifie le critère de Lindeberg[3] de sorte que, avec $s_n^2 = \sum_{i=1}^{n} f\left(\frac{i}{n}\right)^2$,

$$\frac{1}{s_n} \sum_{i=1}^{n} X_{n,i} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1). \tag{1.53}$$

Or, étant une somme de Riemmann,

$$\frac{1}{n} {s_n}^2 = \frac{1}{n} \sum_{i=1}^{n} f\left(\frac{i}{n}\right)^2 \xrightarrow[n\to\infty]{} \int_0^1 f^2(x)\mathrm{d}x = \|f\|_2^2. \tag{1.54}$$

3. Il est donc raisonnable de penser que $(X_{n,i})$ le vérifie aussi.

Par conséquent, il est naturel de conjecturer le résultat suivant :

$$\frac{1}{\|f\|_2\sqrt{n}}\sum_{i=1}^{n} X_{n,i} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1). \tag{1.55}$$

Tentons donc de montrer ce résultat.

Soit $g \in C^2$ et $\varepsilon > 0$. Notons $K_g = \|g''\|$. Reprenons la preuve du théorème 1.2, en appliquant sur les tableaux $\frac{1}{\|f\|_2\sqrt{n}}X_{n,i}$ et $Y_{n,i} \sim \mathcal{N}(0, \frac{a_{n,i}}{s_n})$, ce qui donne, avec $\sigma = \frac{1}{s_n}$ et $\sigma' = \frac{1}{\|f\|_2\sqrt{n}}$:

$$\left|\mathbb{E}\left(g\left(\frac{1}{\|f\|_2\sqrt{n}}\sum_{i=1}^{n} X_{n,i}\right)\right) - \mathbb{E}(g(Y))\right| \leq \|T_{\sigma' X_{n,1}}\dots T_{\sigma' X_{n,n}}(f) - T_{\sigma a_{n,1} Y}\dots T_{\sigma a_{n,n} Y}(f)\| \tag{1.56}$$

$$\leq \sum_{i=1}^{n} \|T_{\sigma' X_{n,i}}(g) - T_{\sigma a_{n,i} Y}(g)\| \quad \text{(d'après la propriété 1.5)} \tag{1.57}$$

$$\leq \sum_{i=1}^{n}\left[K_g\frac{a_{n,i}^2}{2}|\sigma'^2-\sigma^2| + \frac{1}{2}(\sigma'^2+\sigma^2)a_{n,i}^2\varepsilon + K_g\left(\int_{|\sigma' x|\geq\delta}\sigma'^2x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x) + \int_{a_{n,i}\sigma|y|\geq\delta}\sigma^2a_{n,i}^2y^2\mathbb{P}_Y(\mathrm{d}y)\right)\right] \tag{1.58}$$

$$\leq \frac{1}{2}K_g\underbrace{\left|\frac{s_n^2}{n\|f\|_2^2}-1\right|}_{A} + \frac{\varepsilon}{2} + \varepsilon\underbrace{\frac{s_n^2}{\|f\|_2^2 n}}_{B} + K_g\left(\underbrace{\sum_{i=1}^{n}\frac{1}{\|f\|_2^2 n}\int_{|x|\geq\delta\|f\|_2\sqrt{n}}x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x)}_{C} + \underbrace{\sum_{i=1}^{n}\int_{a_{n,i}|y|\geq\delta s_n}a_{n,i}^2y^2\mathbb{P}_Y(\mathrm{d}y)}_{D}\right). \tag{1.59}$$

Or $A \leq \varepsilon$ et $B \leq 2$ pour n assez grand d'après (1.54) , $D \leq \varepsilon$ à partir d'un certain rang puisque le tableau $(Y_{n,i})$ vérifie le critère de Lindeberg étendu, et enfin :

$$C = \sum_{i=1}^{n}\frac{1}{\|f\|_2^2 n}\int_{|x|\geq\delta\|f\|_2\sqrt{n}}x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x) \tag{1.60}$$

$$= \underbrace{\frac{s_n^2}{\|f\|_2^2 n}}_{B\leq 2}\frac{1}{s_n^2}\sum_{i=1}^{n}\int_{|x|\geq\delta\|f\|_2\sqrt{n}}x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x). \tag{1.61}$$

De plus le tableau $(\frac{1}{s_n}X_{n,i})$ vérifie le critère de Lindeberg, c'est à dire :

$$\forall\delta' > 0, \lim_{n\to\infty}\frac{1}{s_n^2}\sum_{i=1}^{n}\int_{|x|\geq\delta' s_n}x^2\mathbb{P}_{X_{n,i}} = 0, \tag{1.62}$$

or puisque $s_n \geq \frac{1}{2}\|f\|_2^2\sqrt{n}$ à partir d'un certain rang, en prenant $\delta' = 2\delta$, on a bien :

$$\frac{1}{s_n^2}\sum_{i=1}^{n}\int_{|x|\geq\delta\|f\|_2\sqrt{n}}x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x) \leq \frac{1}{s_n^2}\sum_{i=1}^{n}\int_{|x|\geq\delta' s_n}x^2\mathbb{P}_{X_{n,i}}(\mathrm{d}x) \xrightarrow[n\to\infty]{} 0. \tag{1.63}$$

Finalement, pour n assez grand :

$$\left|\mathbb{E}\left(g\left(\frac{1}{\|f\|_2\sqrt{n}}\sum_{i=1}^{n} X_{n,i}\right)\right) - \mathbb{E}(g(Y))\right| \leq \tilde{K}_g\varepsilon, \tag{1.64}$$

où $\tilde{K}_g$ est une constante qui ne dépend que de g. Ceci étant vrai $\forall\varepsilon > 0$, on a donc

$$\left|\mathbb{E}\left(g\left(\frac{1}{\|f\|_2\sqrt{n}}\sum_{i=1}^{n} X_{n,i}\right)\right) - \mathbb{E}(g(Y))\right| \xrightarrow[n\to\infty]{} 0, \tag{1.65}$$

et ce pour toute fonction $g \in C^2$, ce qui par la propriété 1.2 donne $\frac{1}{\|f\|_2\sqrt{n}}\sum_{i=1}^{n} X_{n,i} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$. Récapitulons cela sous la forme :

Théorème 1.4. *Soit $(X_{n,i})_{1\leq i\leq n}$ un tableau triangulaire de variables centrées réduites. $f \in C^0([0,1],]0,+\infty[)$, $a_{n,i} = f\left(\frac{i}{n}\right)$ tels que $(a_{n,i}X_{n,i})$ vérifie le critère de Lindeberg étendu. Alors*

$$\sum_{i=1}^{n} f\left(\frac{i}{n}\right)\frac{X_{n,i}}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \|f\|_2^2\right). \tag{1.66}$$

Chapitre 2

Généralisation

Cette section est fortement inspirée de la lecture de [4].

2.1 Un théorème de Chatterjee

Définition 2.1. *On note $C_b^k(\mathbb{R}^n, \mathbb{R})$ l'ensemble des fonctions f k-fois différentiables ($k \geq 1$) de $\mathbb{R}^n$ dans $\mathbb{R}$ telles que toutes les dérivées partielles $\left(\frac{\partial^p f}{\partial x_i^p}\right)_{1 \leq p \leq k}$ soient bornées.*

Sur cet espace, on pose, pour $r \leq k$:

$$\lambda_r = \sup \left\{ \left\| \frac{\partial^p f}{\partial x_i^p} \right\|_\infty^{r/p}, 1 \leq i \leq n, 1 \leq p \leq r \right\}. \tag{2.1}$$

Pour toute fonction g appartenant à $C_b^3(\mathbb{R})$, on note $C_1(g) = \|g'\| + \|g''\|$ et $C_2(g) = \frac{1}{6}\|g'\| + \frac{1}{2}\|g''\| + \frac{1}{6}\|g'''\|$.

Théorème 2.1. *Soit $f_n \in C_b^3(\mathbb{R}^n, \mathbb{R})$, quel que soit n entier. Soit (X_n) et (Y_n) deux suites de variables aléatoires réelles toutes indépendantes et de variances finies, telles que :*

$$\mathbb{E}(X_n) = \mathbb{E}(Y_n) \text{ et } \mathbb{V}(X_n) = \mathbb{V}(Y_n). \tag{2.2}$$

Alors : $\forall g \in C_b^3(\mathbb{R}), \forall \delta > 0$,

$$\begin{aligned}|\mathbb{E}(g(f_n(X_1, \ldots, X_n)) - \mathbb{E}(g(f_n(Y_1, \ldots, Y_n))| \leq C_1(g)\lambda_2(f_n) \sum_{i=1}^n \left(\int_{|x|>\delta} x^2 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|>\delta} y^2 \mathbb{P}_{Y_i}(\mathrm{d}y) \right) \\ + C_2(g)\lambda_3(f_n) \sum_{i=1}^n \left(\int_{|x|\leq\delta} |x|^3 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|\leq\delta} |y|^3 \mathbb{P}_{Y_i}(\mathrm{d}y) \right).\end{aligned} \tag{2.3}$$

Démonstration. Fixons pour la suite un entier n, et notons $X = (X_1, \ldots, X_n)$ et $Y = (Y_1, \ldots, Y_n)$.

Soit $f \in C_b^3(\mathbb{R}^n, \mathbb{R})$ et $g \in C_b^3(\mathbb{R})$. Posons $h = g \circ f$

Première étape : Majoration des dérivées partielles

Le foncion h est bien entendu 3 fois différentiable de $\mathbb{R}^n$ dans $\mathbb{R}$. Commençons par établir des majorations uniformes sur les dérivées partielles de h à l'aide de λ_2 et λ_3.

Remarquons que : $\forall x \in \mathbb{R}^n, \forall 1 \leq i \leq n$,

$$\frac{\partial^2 h}{\partial x_i^2}(x) = g'(f(x))\frac{\partial^2 f}{\partial x_i^2}(x) + g''(f(x))\left(\frac{\partial f}{\partial x_i}\right)^2 \tag{2.4}$$

$$\frac{\partial^3 h}{\partial x_i^3}(x) = g'(f(x))\frac{\partial^3 f}{\partial x_i^3}(x) + 3g''(f(x))\frac{\partial^2 f}{\partial x_i^2}(x)\frac{\partial f}{\partial x_i}(x) + g'''(f(x))\left(\frac{\partial f}{\partial x_i}\right)^3. \tag{2.5}$$

Puisque

$$\left\| \frac{\partial^p f}{\partial x_i^p} \right\|_{\infty, \mathbb{R}^n} \leq \lambda_r^{p/r}, \quad \forall 1 \leq p \leq r \leq 3, \tag{2.6}$$

(2.4) donne

$$\left\| \frac{\partial^2 h}{\partial x_i^2} \right\|_{\infty, \mathbb{R}^n} \leq C_1(g)\lambda_2(f), \tag{2.7}$$

et (2.5) donne

$$\left\| \frac{\partial^3 h}{\partial x_i^3} \right\|_{\infty,\mathbb{R}^n} \leq 6C_2(g)\lambda_3(f). \tag{2.8}$$

Deuxième étape : Majoration des incréments

Posons $Z_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$ et $W_i = (X_1, \ldots, X_{i-1}, 0, Y_{i+1}, \ldots, Y_n)$; $Z_n = X$.
Ainsi, l'on a :

$$|\mathbb{E}(g(f(X)) - \mathbb{E}(g(f(Y)))| = \left| \sum_{i=1}^{n} \mathbb{E}(h(Z_i) - h(Z_{i-1})) \right|. \tag{2.9}$$

Etudions les quantités :

$$R_i = h(Z_i) - X_i \frac{\partial h}{\partial x_i}(W_i) - \frac{1}{2} X_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i), \tag{2.10}$$

$$T_i = h(Z_{i-1}) - Y_i \frac{\partial h}{\partial x_i}(W_i) - \frac{1}{2} Y_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i). \tag{2.11}$$

Appliquons maintenant l'inégalité de Taylor-Lagrange sur chacune des fonctions

$$h_{i,a}(x) = h(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n), \tag{2.12}$$

à l'ordre 3 :

$$|R_i| \leq |X_i|^3 \left\| \frac{\partial^3 h}{\partial x_i^3} \right\|_{\infty,\mathbb{R}^n} \frac{1}{3!}, \tag{2.13}$$

et en injectant (2.8) :

$$|R_i| \leq |X_i|^3 C_2(g)\lambda_3(f). \tag{2.14}$$

De plus, en majorant à l'aide de l'inégalité à l'ordre 2, on a :

$$\left| R_i + \frac{1}{2} X_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i) \right| \leq |X_i|^2 \left\| \frac{\partial^2 h}{\partial x_i^2} \right\|_{\infty,\mathbb{R}^n} \frac{1}{2!}, \tag{2.15}$$

et donc :

$$|R_i| \leq |X_i|^2 \left\| \frac{\partial^2 h}{\partial x_i^2} \right\|_{\infty,\mathbb{R}^n}, \tag{2.16}$$

ce qui en injectant (2.7) donne :

$$|R_i| \leq |X_i|^2 C_1(g)\lambda_2(f). \tag{2.17}$$

Le même raisonnement donne également :

$$|T_i| \leq |Y_i|^3 C_2(g)\lambda_3(f), \tag{2.18}$$

$$|T_i| \leq |Y_i|^2 C_1(g)\lambda_2(f). \tag{2.19}$$

Troisième étape : Majoration de l'erreur totale

Faisons donc apparaitre de force ces développements :

$$|\mathbb{E}(g(f(X))) - \mathbb{E}(g(f(Y)))| =$$
$$\left| \sum_{i=1}^{n} \mathbb{E}\left[R_i + X_i \frac{\partial h}{\partial x_i}(W_i) + \frac{1}{2} X_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i) - T_i - \left(Y_i \frac{\partial h}{\partial x_i}(W_i) + \frac{1}{2} Y_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i) \right) \right] \right|. \tag{2.20}$$

Remarquons maintenant, puisque X_i et Y_i sont indépendantes de W_i, que d'une part :

$$\mathbb{E}\left(X_i \frac{\partial h}{\partial x_i}(W_i) - Y_i \frac{\partial h}{\partial x_i}(W_i) \right) = \mathbb{E}(X_i - Y_i)\mathbb{E}\left(\frac{\partial h}{\partial x_i}(W_i) \right) = 0, \tag{2.21}$$

et d'autre part :

$$\mathbb{E}\left(X_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i) - Y_i^2 \frac{\partial^2 h}{\partial x_i^2}(W_i) \right) = \mathbb{E}(X_i^2 - Y_i^2)\mathbb{E}\left(\frac{\partial^2 h}{\partial x_i^2}(W_i) \right) = 0. \tag{2.22}$$

Finalement, on a quel que soit $\delta > 0$:

$$|\mathbb{E}(g(f(X))) - \mathbb{E}(g(f(Y)))| = \left|\sum_{i=1}^{n} \mathbb{E}\left(R_i - T_i\right)\right| \tag{2.23}$$

$$\leq \sum_{i=1}^{n} \mathbb{E}(|R_i|(\mathbb{1}_{|X_i|\leq\delta} + \mathbb{1}_{|X_i|>\delta})) + \mathbb{E}(|T_i|(\mathbb{1}_{|Y_i|\leq\delta} + \mathbb{1}_{|Y_i|>\delta})) \tag{2.24}$$

$$\leq C_1(g)\lambda_2(f_n) \sum_{i=1}^{n} \left(\int_{|x|>\delta} x^2 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|>\delta} y^2 \mathbb{P}_{Y_i}(\mathrm{d}y)\right) \tag{2.25}$$

$$+ C_2(g)\lambda_3(f_n) \sum_{i=1}^{n} \left(\int_{|x|\leq\delta} |x|^3 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|\leq\delta} |y|^3 \mathbb{P}_{Y_i}(\mathrm{d}y)\right). \tag{2.26}$$

□

2.2 Vitesse de convergence

Ce théorème donne en particulier une borne pour la vitesse de convergence des suites. Cela nécessite une condition supplémentaire : on suppose désormais dans ce paragraphe que les variables possèdent des moments d'ordre 3 finis, c'est à dire

$$\forall n \geq 1, \quad \gamma_n = \max_{k\in[\![1,n]\!]} \left\{\mathbb{E}(|X_k|^3), \mathbb{E}(|Y_k|^3)\right\} < \infty. \tag{2.27}$$

On a alors le théorème suivant :

Théorème 2.2. *Sous les mêmes conditions que le théorème 2.1 et avec la condition* (2.27) *en plus, l'on a :*

$$|\mathbb{E}(g(f_n((X_1, \ldots, X_n))) - \mathbb{E}(g(f_n((Y_1, \ldots, Y_n)))| \leq 2C_2(g)\lambda_3(f_n)n\gamma_n. \tag{2.28}$$

Démonstration. La structure de la preuve est la même que pour le théorème 2.1. Nous reprenons donc les mêmes notations. En revanche, la majoration va être sensiblement plus simple grâce à l'existence des moments d'ordre 3. Ecrivons l'équation (2.23) qui est toujours valable :

$$|\mathbb{E}(g(f(X))) - \mathbb{E}(g(f(Y)))| = \left|\sum_{i=1}^{n} \mathbb{E}\left(R_i - T_i\right)\right| \tag{2.29}$$

$$\leq \sum_{i=1}^{n} C_2(g)\lambda_3(f)\mathbb{E}(|X_i|^3 + |Y_i|^3) \tag{2.30}$$

$$\leq nC_2(g)\lambda_3(f)2\gamma_n. \tag{2.31}$$

□

2.3 Application au TCL

On pose :

$$f_n(x_1, \ldots, x_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i, \tag{2.32}$$

alors $f_n \in C_b^3(\mathbb{R}^n, \mathbb{R})$ et $\lambda_r(f_n) = \dfrac{1}{n^{r/2}}, \quad \forall 1 \leq r \leq 3$. Le théorème 2.1 donne, pour $Y_n \sim \mathcal{N}(0,1)$, et la famille (X_n) identiquement distribuée, centrée réduite : $\forall g \in C_b^3(\mathbb{R})$,

$$\left|\mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_n\right)\right) - \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_n\right)\right)\right| \leq C_1(g)\frac{1}{n}\sum_{i=1}^{n}\left(\int_{|x|>\delta} x^2 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|>\delta} y^2 \mathbb{P}_{Y_i}(\mathrm{d}y)\right) \tag{2.33}$$

$$+ C_2(g)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\left(\int_{|x|\leq\delta} |x|^3 \mathbb{P}_{X_i}(\mathrm{d}x) + \int_{|y|\leq\delta} |y|^3 \mathbb{P}_{Y_i}(\mathrm{d}y)\right) \tag{2.34}$$

$$\leq C_1(g)\left(\int_{|x|>\delta} x^2 \mathbb{P}_{X_1}(\mathrm{d}x) + \int_{|y|>\delta} y^2 \mathbb{P}_{Y_1}(\mathrm{d}y)\right) \tag{2.35}$$

$$+ C_2(g)\frac{1}{\sqrt{n}}\left(\delta\mathbb{V}(X_1) + \delta\mathbb{V}(Y_1)\right). \tag{2.36}$$

Soit $\varepsilon > 0$ et δ tels que

$$\left(\int_{|x|>\delta} x^2 \mathbb{P}_{X_1}(\mathrm{d}x) + \int_{|y|>\delta} y^2 \mathbb{P}_{Y_1}(\mathrm{d}y)\right) \leq \varepsilon(\mathbb{V}(X_1) + \mathbb{V}(Y_1)), \tag{2.37}$$

et n_0 tel que :

$$\forall n \geq n_0, \quad \frac{1}{\sqrt{n}}\delta \leq \varepsilon. \tag{2.38}$$

On a donc pour n assez grand :

$$\left|\mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_n\right)\right) - \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_n\right)\right)\right| \leq \varepsilon(C_1(g) + C_2(g))(\mathbb{V}(X_1) + \mathbb{V}(Y_1)), \tag{2.39}$$

et ce, quel que soit $\varepsilon > 0$, donc finalement :

$$\forall g \in C_b^3(\mathbb{R}), \quad \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_n\right)\right) - \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_n\right)\right) \underset{n\to\infty}{\longrightarrow} 0 \tag{2.40}$$

Pour aboutir au théorème Central Limite, il reste deux résultats essentiels : la stabilité de la loi normale, pour affirmer que [1] :

$$\mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_n\right)\right) = \mathbb{E}(g(Y)), \quad Y \sim \mathcal{N}(0,1), \tag{2.41}$$

et l'extension de la propriété 1.2, pour obtenir :

$$\left[\forall g \in C_b^3(\mathbb{R}), \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_n\right)\right) \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left(g(Y)\right)\right] \Longrightarrow \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_n \underset{n\to\infty}{\overset{\text{loi}}{\longrightarrow}} Y. \tag{2.42}$$

Pour ce dernier point, la seule chose à adapter dans la preuve de la propriété 1.2 est la possibilité de construire les fonctions f et g dans l'espace C_b^3 vérifiant les propriétés voulues, ce qui ne soulève pas de difficulté (on peut même demander d'être dans l'espace C_b^∞).

De plus, il est facile d'adapter ces preuves afin d'obtenir en corollaire le théorème de Lindeberg, et même pour le théorème relatif aux tableaux triangulaires.

En conclusion, le principe de Lindeberg permet efficacement de se passer de l'analyse de Fourier. En particulier, elle permet de se passer de l'hypothèse de distribution identique, et même dans d'autres cas non présentés ici, de l'hypothèse d'indépendance des variables, notamment grâce à la généralisation à toute fonction suffisamment régulière, permettant l'introduction de phénomènes non linéaires.

Intéressons-nous désormais à une généralisation supplémentaire. L'objectif est désormais d'étudier le comportement asymptotique lorsque l'on retire l'existence du moment d'ordre 2 et la condition de moyenne nulle, et les nouvelles lois limites.

1. Déjà démontré au chapitre 1.

Chapitre 3

Lois stables et théorèmes limites associés

Dans toute la suite, on notera : $X \overset{\text{loi}}{=} Y$ l'égalité en loi sur les variables aléatoires, c'est à dire l'égalité $\mathbb{P}_X = \mathbb{P}_Y$, ou encore $F_X = F_Y$. On ne parlera ici que de variables réelles.

3.1 Définitions équivalentes

3.1.1 Stabilité

Dans les premières parties, on a pu constater que la propriété de stabilité de la loi normale était essentielle dans le théorème Central Limite, plus précisément le fait que la somme $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_n$ ne soit autre, en loi, que Y, dès que toutes ces variables suivent des lois normales centrées réduites. Donnons une définition de ce que nous appelons la stabilité [1] :

Définition 3.1. *La loi d'une variable aléatoire X est dite stable s'il existe des réels a, b et c strictement positifs et d tels que, pour toutes copies indépendantes X_1 et X_2 de X :*

$$aX_1 + bX_2 \overset{\text{loi}}{=} cX + d. \tag{3.1}$$

On parle de stabilité stricte si $d = 0$.

Cette définition est évidemment équivalente à la définition suivante. [2] :

Définition 3.2. *La loi d'une variable aléatoire X est dite stable s'il existe des réels $(c_n)_{n\in\mathbb{N}}$ strictement positifs et $(d_n)_{n\in\mathbb{N}}$ tels que, pour toutes copies indépendantes (X_n) de X :*

$$\forall n \geq 1, \quad \sum_{i=1}^{n} X_i \overset{\text{loi}}{=} c_n X + d_n. \tag{3.2}$$

Toute loi stable vérifie un principe fondamental quant à la pondération des incréments qui composent sa somme :

Théorème 3.1. *Soit X une variable aléatoire dont la loi est stable, c'est à dire vérifiant* (3.1)*. Alors :*

$$\exists!\alpha, \quad 0 < \alpha \leq 2, \quad a^\alpha + b^\alpha = c^\alpha. \tag{3.3}$$

La preuve de ce théorème est exposée dans [8], VI.1.
En employant l'autre définition, on obtient cette écriture :

Théorème 3.2. *Soit X une variable aléatoire dont la loi est stable, c'est à dire vérifiant* (3.2)*. Alors :*

$$\exists!\alpha, \quad 0 < \alpha \leq 2, \quad c_n = n^{1/\alpha}. \tag{3.4}$$

Bien entendu, l'indice α défini par ces deux théorèmes est bien le même quelle que soit la définition employée.

Définition 3.3. *Soit X une variable aléatoire dont la loi est stable. L'indice α défini par les théorèmes 3.1 et 3.2 est appelé indice de stabilité de la loi. La loi est alors dite α-stable.*

Propriété 3.1. *La loi normale 2-stable, et la loi normale centrée est strictement stable.*

1. Cette définition provient de [5].
2. Cette définition provient de [8].

Propriété 3.2. *Les lois stables d'indice 2 sont les seules lois stables possédant une variance finie.*

Démonstration. Supposons X ayant une loi stable et une variance finie. Appliquons la variance à l'équation 3.1. On obtient :

$$a^2 + b^2 = c^2, \tag{3.5}$$

donc l'indice de stablilité de la loi de X est 2. □

3.1.2 Domaine d'attraction

Définition 3.4. *La loi d'une variable aléatoire X possède un domaine d'attraction s'il existe une suite (X_n) de variables aléatoires indépendantes identiquement distribuées, (s_n) une suite de réels strictement positifs et (b_n) une suite de réels tels que :*

$$\frac{1}{s_n}\left(\sum_{i=1}^{n} X_i - b_n\right) \xrightarrow[n\to\infty]{\text{loi}} X. \tag{3.6}$$

On dit alors que la loi commune aux variables X_n appartient au domaine d'attraction de la loi de X.

Propriété 3.3. *Toute loi possédant un domaine d'attraction est stable.*

La preuve de ce théorème est exposée dans [8], VI.1.

Propriété 3.4. *Toute loi stable a un domaine d'attraction et appartient à son propre domaine d'attraction.*

Les propriétés 3.3 et 3.4 donnent ensemble le théorème suivant :

Théorème 3.3. *Une loi est stable si et seulement si elle possède un domaine d'attraction.*

3.1.3 Forme explicite

Définition 3.5. *Soient $\alpha, \sigma, \beta, \mu$ des réels tels que $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$. Notons $\mathcal{S}_\alpha(\sigma, \beta, \mu)$ la loi dont la fonction caractéristique de toute variable X qui la suit est :*

$$\varphi_X(t) = \exp\left[-\sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu t\right] \quad si\ \alpha \neq 1, \tag{3.7}$$

ou

$$\varphi_X(t) = \exp\left[-\sigma|t| \left(1 - i\beta\frac{2}{\pi}\operatorname{sgn}(t)\ln|t|\right) + i\mu t\right] \quad si\ \alpha = 1. \tag{3.8}$$

Théorème 3.4. *Les lois $S_\alpha(\sigma, \beta, \mu)$, où $0 < \alpha \leq 2$, $\sigma \geq 0$ et $-1 \leq \beta \leq 1$ décrivent l'ensemble des lois stables.*

La preuve de ce théorème peut être trouvée sous une autre forme dans [8] ou sous cette forme dans [2].

Propriété 3.5. *Les lois normales sont les seules lois 2-stables.*

Démonstration. Appliquons $\alpha = 2$ dans (3.8) :

$$\varphi_X(t) = \exp\left[-\sigma^2|t|^2 + i\mu t\right], \tag{3.9}$$

ce qui est bien la fonction caractéristique d'une loi normale $\mathcal{N}(\mu, 2\sigma^2)$. □

3.2 Limite de somme de variables i.i.d.

3.2.1 Appartenance à un domaine

Intéressons-nous maintenant aux conditions pour qu'une loi appartienne à un domaine d'attraction.

Définition 3.6. *On dit d'une fonction réelle f qu'elle est à variation lente si elle vérifie :*

$$\lim_{x\to\pm\infty} \frac{f(cx)}{f(x)} = 1, \quad \forall c > 0. \tag{3.10}$$

Définition 3.7. *On définit le moment tronqué μ_X d'une variable aléatoire réelle X de la façon suivante :*

$$\forall x \in \mathbb{R}, \quad \mu_X(x) = \int_{-x}^{x} t^2 \mathbb{P}_X(\mathrm{d}t). \tag{3.11}$$

Théorème 3.5. *Soit X une variable aléatoire non constante. La loi de X appartient à un domaine d'attraction si et seulement si d'une part le moment tronqué μ_X vérifie en l'infini :*

$$\mu_X(x) \sim x^{2-\alpha} L(x), \tag{3.12}$$

où L est une fonction à variation lente, et d'autre part soit $\alpha = 2$, soit $0 < \alpha < 2$ et en notant F la fonction de répartition de X :

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} \underset{x \to +\infty}{\longrightarrow} p \quad \text{et} \quad \frac{F(-x)}{1 - F(x) + F(-x)} \underset{x \to +\infty}{\longrightarrow} q, \tag{3.13}$$

où $p+q = 1$.

Pour une preuve de ce théorème, voir [8] chapitre XVII. Les conditions proposées ici peuvent être écrites sous différentes formes. En particulier, la lecture de [7], chapitre IV, théorème 14, nous donne l'écriture suivante :

Théorème 3.6. *La loi d'une variable X appartient à un domaine d'attraction d'une loi stable de paramètre $0 < \alpha < 2$ si et seulement si :*

$$F_X(-x) = \frac{c_1 + o(1)}{x^\alpha} h(x) \quad \text{quand } x \to +\infty, \tag{3.14}$$

et

$$1 - F_X(x) = \frac{c_2 + o(1)}{x^\alpha} h(x) \quad \text{quand } x \to +\infty, \tag{3.15}$$

où c_1 et c_2 sont des constantes positives non toutes deux nulles, et h est une fonction à variation lente.

3.2.2 Théorème limite

En premier lieu, citons la propriété suivante, que l'on pourra trouver dans [5], propriété 1.2.5 :

Propriété 3.6. *La loi $S_\alpha(\sigma, \beta, \mu)$ est symétrique autour de μ si et seulement si $\beta = 0$.*

Ainsi, l'on va pouvoir écrire le théorème suivant :

Théorème 3.7. *Soit (X_n) une suite de variables aléatoires indépendantes et identiquement distribuées, dont la loi appartient à un domaine d'attraction, et soit α défini par le théorème 3.5. Alors il existe deux suites (s_n) et (b_n) de réels, $\sigma \geq 0$, $-1 \leq \beta \leq 1$, et μ tels que :*

$$\frac{1}{s_n} \sum_{i=1}^{n} (X_i - b_n) \underset{n \to \infty}{\overset{\text{loi}}{\longrightarrow}} S_\alpha(\sigma, \beta, \mu). \tag{3.16}$$

De plus on a :
- *si $\mathbb{E}(X_1) < \infty$, alors $b_n = \mathbb{E}(X_1)$ convient et alors $\mu = 0$.*
- *si la loi de X_1 est symétrique alors $\beta = 0$,*
- *si $\mathbb{V}(X_1) < \infty$, $\alpha = 2$ et $2\sigma^2 = \mathbb{V}(X_1)$.*

Malheureusement, en dehors des cas simples évoqués, la recherche des paramètres de la loi limite n'est pas triviale, et nécessite une étude approfondie, dont on pourra trouver un premier avancement dans [7] et [5].

Bibliographie

[1] Billingsley. *Probability and Measure.* Wiley, 1995.

[2] Gnedenko and Kolmogorov. *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, 1954.

[3] Lindeberg. Eine neue herleitung des exponentialgesetzes in der wahrscheinlichkeit-srechnung. *Math. Zeitschr*, pages 211–225, 1922.

[4] Chatterjee S. A simple invariance theorem. *arxiv*, 2008.

[5] Samorodnitsky and Taqqu. *Stable Non-Gaussian Random Processes.* Chapman & Hall, 1994.

[6] Trotter. Elementary proof of the central limit theorem. *Archiv der Mathem*, pages 226–234, 1959.

[7] Petrov V. V. *Sums of Independent Random Variables.* Springer-Verlag, 1975.

[8] Feller W. *An Introduction to Probability Theroy and Its Applications*, volume 2. Wiley, 1971.