Théorème de Lindeberg

Leçons : 223, 261, 262, 266

Réf :
- De l'intégration aux probabilités, O. Garet

Théorème (Lindeberg)

Soient $(X_{n,k})_{n,k\in\mathbb{N}^*}$ des variables aléatoires centrées admettant un moment d'ordre 2 et (N_n) une suite d'entiers naturels tendant vers $+\infty$.

Supposons que pour tout $n \in \mathbb{N}$ les variables $(X_{n,k})_{1\leq k\leq N_n}$ sont indépendantes.

Notons $S_n = \sum_{i=1}^{N_n} X_{n,i}$ et $\sigma_n^2 = \mathrm{Var}(S_n)$.

Si la condition de Lindeberg

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \lim_{n\to+\infty} \sum_{k=1}^{N_n} \frac{1}{\sigma_n^2} \mathbb{E}\left[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}|\geq \varepsilon\sigma_n\}}\right] = 0$$

est vérifiée, alors on a

$$\frac{S_n}{\sigma_n} \xrightarrow[n\to+\infty]{\text{Loi}} \mathcal{N}(0,1)$$

Preuve :

Soient n, k dans $\mathbb{N}^*$. Notons $\phi_{n,k}$ la fonction caractéristique de $\frac{X_{n,k}}{\sigma_n}$. On a alors par indépendance, $\varphi_{\frac{S_n}{\sigma_n}} = \prod_{k=1}^{N_n} \phi_{n,k}$.

On va montrer que pour tout $t \in \mathbb{R}$ $\varphi_{\frac{S_n}{\sigma_n}}(t)$ converge vers $e^{-\frac{t^2}{2}}$ et appliquer le théorème de Lévy.

Etape 1 : soit $\psi_{n,k}$ fonction caractéristique d'une loi normale centrée de même variance que $\frac{X_{n,k}}{\sigma_n}$. Soit $t \in \mathbb{R}$. On a

$$\prod_{k=1}^{N_n} \phi_{n,k}(t) - \exp\left(-\frac{t^2}{2}\right) = \prod_{k=1}^{N_n} \phi_{n,k}(t) - \prod_{k=1}^{N_n} \psi_{n,k}(t)$$

car par indépendance $\frac{1}{\sigma_n^2} \sum_{k=1}^{n} \mathrm{Var}(X_{n,k}) = 1$.

Lemme

Soient $m \in \mathbb{N}^*$, $a_1, \dots, a_m, b_1, \dots, b_m$ des nombres complexes de module inférieur à 1.

Alors on a
$$\left| \prod_{j=1}^{m} a_j - \prod_{j=1}^{m} b_j \right| \leqslant \sum_{j=1}^{m} |b_j - a_j|$$

Preuve: par récurrence sur m. ∎

Fort de ce lemme, on obtient

$$\left| \prod_{k=1}^{N_n} \phi_{n,k}(t) - \prod_{k=1}^{N_n} \psi_{n,k}(t) \right| \leqslant \sum_{k=1}^{N_n} |\phi_{n,k}(t) - \psi_{n,k}(t)|$$

Etape 2 :

Lemme

Soit $x \in \mathbb{R}$. On a
$$\left| e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) \right| \leqslant \min\left(x^2, \frac{|x|^3}{3!}\right)$$

Preuve: on a par formule de Taylor avec reste intégral à l'ordre 2

$$e^{ix} = 1 + ix - x^2 \int_0^1 (1-u) e^{iux} du.$$

Comme $\int_0^1 (1-u) du = \frac{1}{2}$, on obtient

$$e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) = -x^2 \int_0^1 (1-u)(e^{iux} - 1) du.$$

d'où $\left| e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) \right| \leqslant x^2.$

Puis avec une formule à l'ordre 3 on a

$$e^{ix} = 1 + ix - \frac{x^2}{2} - i\frac{x^3}{2} \int_0^1 (1-u)^2 e^{iux} du$$

Or on a $\int_0^1 (1-u)^2 du = \left[-\frac{1}{3}(1-u)^3 \right]_0^1 = \frac{1}{3}$ d'où le lemme ∎

Du lemme s'en suit

$$\left| \phi_{n,k} - \left(1 - \frac{1}{2} t^2 \frac{\mathrm{Var}(X_{n,k})}{\sigma_n^2}\right) \right| \leqslant \mathbb{E}[Y]$$

où $Y = \min\left(t^2 \left(\frac{X_{n,k}}{\sigma_n} \right)^2, t^3 \left(\frac{|X_{n,k}|}{\sigma_n} \right)^3 \right)$.

Soit $\varepsilon \in \mathbb{R}_+^*$. On a

$$\mathbb{E}[Y] = \mathbb{E}[Y \mathbb{1}_{\{|X_{n,k}| < \varepsilon \sigma_n\}}] + \mathbb{E}[Y \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon \sigma_n\}}]$$

d'où
$$\mathbb{E}[Y] \leqslant |t|^3 \mathbb{E}\left[\frac{|X_{n,k}|^3}{\sigma_n^3} \mathbb{1}_{\left\{ \frac{|X_{n,k}|}{\sigma_n} < \varepsilon \right\}} \right] + t^2 \mathbb{E}\left[X_{n,k}^2 \frac{1}{\sigma_n^2} \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon \sigma_n\}} \right]$$
$$\leqslant \varepsilon |t|^3 \frac{\mathrm{Var}(X_{n,k})}{\sigma_n^2} + \frac{t^2}{\sigma_n^2} \mathbb{E}[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon \sigma_n\}}]$$

et ainsi on obtient

$$\sum_{k=1}^{N_n} \left| \phi_{n,k}(t) - \left(1 - \frac{1}{2} t^2 \frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2}\right) \right| \leqslant \varepsilon |t|^3 + t^2 \sum_{k=1}^{N_n} \frac{1}{\sigma_n^2} \mathbb{E}[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \varepsilon \sigma_n\}}]$$

d'où par condition de Lindeberg

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \limsup_{n \to +\infty} \sum_{k=1}^{N_n} \left| \phi_{n,k}(t) - \left(1 - \frac{t^2}{2} \frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2}\right) \right| \leqslant \varepsilon |t|^3$$

et donc $\displaystyle \lim_{n \to +\infty} \sum_{k=1}^{N_n} \left| \phi_{n,k}(t) - \left(1 - \frac{t^2}{2} \frac{\mathrm{Var}(X_{n,k})}{\sigma_n^2}\right) \right| = 0$

<u>Etape 3</u> : il ne reste plus qu'à montrer qu'on a

$$\lim_{n \to +\infty} \sum_{k=1}^{N_n} \left| \varphi_{n,k}(t) - \left(1 - \frac{t^2}{2} \frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2}\right) \right| = 0$$

<u>Lemme</u>

On a $\forall z \in \mathbb{C} \quad |e^z - 1 - z| \leqslant \frac{e^{|z|}}{2} |z|^2$

<u>Preuve</u> : avec le développement en série entière. ■

Par ce lemme on obtient

$$\sum_{k=1}^{N_n} \left| \exp\left(-\frac{t^2}{2} \frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2}\right) - \left(1 - \frac{t^2}{2} \frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2}\right) \right| \leqslant \frac{t^4}{8} e^{\frac{t^2}{2}} \frac{1}{\sigma_n^4} \sum_{k=1}^{N_n} (\mathrm{Var}\, X_{n,k})^2$$

car $\frac{\mathrm{Var}\, X_{n,k}}{\sigma_n^2} \leqslant 1$. et $x \longmapsto e^x$ est croissante.

Notons M_n la quantité $\max\{\operatorname{Var} X_{n,k} \,;\, 1 \leq k \leq N_n\}$. On a alors

$$\frac{1}{\sigma_n^4}\sum_{k=1}^{N_n} \operatorname{Var}(X_{n,k})^2 \leq \frac{1}{\sigma_n^4}\sum_{k=1}^{N_n} M_n \operatorname{Var} X_{n,k} = \frac{M_n}{\sigma_n^2}$$

Montrons qu'on a $\lim\limits_{n\to+\infty} \frac{M_n}{\sigma_n^2} = 0$ et cela conclura.

Soient $\varepsilon \in]0, \frac{1}{2}[$ et $n \in \mathbb{N}$. Pour tout $k \in [\![1, N_n]\!]$ on a

$$\mathbb{E}\left[\frac{X_{n,k}^2}{\sigma_n^2}\right] = \frac{1}{\sigma_n^2}\mathbb{E}\left[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| < \varepsilon\sigma_n\}}\right] + \frac{1}{\sigma_n^2}\mathbb{E}\left[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon\sigma_n\}}\right]$$

$$\leq \varepsilon^2 + \frac{1}{\sigma_n^2}\mathbb{E}\left[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon\sigma_n\}}\right]$$

D'où $$\frac{M_n}{\sigma_n^2} \leq \varepsilon^2 + \sum_{k=1}^{N_n} \frac{1}{\sigma_n^2}\mathbb{E}\left[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| \geq \varepsilon\sigma_n\}}\right]$$

et Lindeberg conclut comme précédemment ■

<u>Application</u> : (Théorème central limite)

Soit (X_n) une suite de vaiid centrées admettant une variance σ^2. Notons S_n la somme $\sum_{k=1}^{n} X_i$.

Alors on a

$$\frac{S_n}{\sqrt{n}\,\sigma} \xrightarrow[n\to+\infty]{\text{Loi}} \mathcal{N}(0,1)$$

<u>Preuve</u> : avec les notations précédentes, on considère $N_n = n$, $(X_{n,k}) = (X_k)$, $\sigma_n^2 = n\sigma^2$, $S_n = \sum_{k=1}^{n} X_k$. La condition de Lindeberg s'écrit

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{n\sigma^2}\mathbb{E}\left[X_k^2 \mathbb{1}_{\{|X_k| \geq \varepsilon\sqrt{n}\sigma\}}\right] = 0$$

c-à-d

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \lim_{n\to+\infty} \frac{1}{\sigma^2}\mathbb{E}\left[X_1^2 \mathbb{1}_{\{|X_1| \geq \varepsilon\sqrt{n}\sigma\}}\right] = 0$$

ce qui est bien vérifié car $X_1 \in L^2(\Omega)$ ■

Proposition (Lyapounov)

Soient $(X_{n,k})$ des variables aléatoires centrées admettant un moment d'ordre 2, et (N_n) une suite d'entiers positifs de limite infinie.

Supposons que pour tout $n \in \mathbb{N}$ $(X_{n,k})_{1 \leq k \leq N_n}$ sont indépendantes. Soit $\delta \in \mathbb{R}_+^*$.

Si les variables $(X_{n,k})$ admettent un moment d'ordre $2+\delta$ et satisfont la condition de Lyapounov

$$\lim_{n \to +\infty} \sum_{k=1}^{N_n} \frac{1}{\sigma_n^{2+\delta}} \mathbb{E}\left[|X_{n,k}|^{2+\delta}\right] = 0$$

Alors on a

$$\frac{S_n}{\sigma_n} \xrightarrow[n \to +\infty]{\text{Loi}} \mathcal{N}(0,1)$$

Preuve: pour tout $(n,k) \in \mathbb{N}^2$ et tout $\varepsilon \in \mathbb{R}_+^*$ on a

$$\mathbb{1}_{\{|X_{n,k}| \geq \varepsilon \sigma_n\}} \leqslant \left(\frac{|X_{n,k}|}{\varepsilon \sigma_n}\right)^{\delta}$$

et ainsi on obtient la condition de Lindeberg. Cela conclut ∎

Corollaire

Soient $(X_{n,k})$ des vai centrées, bornées par M p.s. et (N_n) une suite d'éléments de $\mathbb{N}$ telle que $\lim_{n \to +\infty} N_n = +\infty$.

On note S_n la somme $\sum_{k=1}^{N_n} X_{n,k}$ et σ_n^2 la variance $\mathrm{Var}(S_n)$.

Si $\lim_{n \to +\infty} \sigma_n = +\infty$ alors on a

$$\frac{S_n}{\sigma_n} \xrightarrow[n \to +\infty]{\text{Loi}} \mathcal{N}(0,1)$$

Preuve: Soit $\delta \in \mathbb{R}_+^*$. On a

$$\sum_{k=1}^{N_n} \frac{1}{\sigma_n^{2+\delta}} \mathbb{E}\left[|X_{n,k}^{2+\delta}|\right] \leqslant \sum_{k=1}^{N_n} \frac{M^{\delta}}{\sigma_n^{2+\delta}} \mathbb{E}\left[X_{n,k}^2\right] = \frac{M^{\delta}}{\sigma_n^{\delta}}$$

et donc la condition de Lyapounov est vérifiée. Cela conclut ∎

Application directe de Lindeberg :

On a la convergence en loi suivante :

$$\sqrt{n}\left(\Gamma(n,n)-1\right)\xrightarrow[n\to+\infty]{\text{Loi}}\mathcal{N}(0,1)$$

(tiré d'un texte de modélisation, la loi $\Gamma(n,n)$ représente la loi d'un quotient $\frac{\beta}{\hat\beta}$ où β est un paramètre à estimer et $\hat\beta$ l'estimateur de vraisemblance du modèle en jeu)

Preuve :

La loi gamma $\Gamma(n,n)$ peut s'écrire comme une somme de n lois $\mathcal{E}(n)$ indépendantes.

En reprenant les notations du théorème de Lindeberg, on note $X_{n,k}$ une v.a de loi $\mathcal{E}(n)$, (N_n) la suite (n) et on considère $(X_{n,k})_{1\leq k\leq n}$ indépendantes.

On a alors

$$\forall n\in\mathbb{N}^* \qquad \sigma_n^2 := \operatorname{Var}\Big(\sum_{k=1}^n X_{n,k}\Big)=\frac{1}{n}.$$

On va appliquer le théorème de Lindeberg à la suite $(\tilde X_{n,k})$ définie par $\tilde X_{n,k}=X_{n,k}-\frac{1}{n}$ qui est centrée.

On s'assure que $(\tilde X_{n,k})$ vérifie la condition de Lindeberg

Soit $\varepsilon\in\mathbb{R}_+^*$. On a

$$\sum_{k=1}^n\frac{1}{\sigma_n^2}\mathbb{E}\left[\tilde X_{n,k}^2\mathbb{1}_{\{|\tilde X_{n,k}|\geq\varepsilon\sigma_n\}}\right] = n^2\,\mathbb{E}\left[\tilde X_{n,1}^2\mathbb{1}_{\{|\tilde X_{n,1}|\geq\varepsilon\frac{1}{\sqrt{n}}\}}\right]$$

$$\leq n^2\,\mathbb{E}\left[\tilde X_{n,1}^4\right]^{\frac12}\mathbb{P}\Big(|\tilde X_{n,1}|\geq\varepsilon\frac{1}{\sqrt{n}}\Big)^{\frac12}$$

par inégalité de Cauchy-Schwarz.

Par inégalité de Markov, on obtient

$$\mathbb{P}\Big(|\tilde X_{n,1}|\geq\frac{\varepsilon}{\sqrt{n}}\Big)=\mathbb{P}\Big(\big(X_{n,1}-\tfrac{1}{n}\big)^2\geq\frac{\varepsilon^2}{n}\Big)\leq\frac{\mathbb{E}\left[\left(X_{n,1}-\frac1n\right)^2\right]}{\varepsilon^2\frac{1}{n}}=\frac{1}{\varepsilon^2 n}$$

Par ailleurs on a

$$\mathbb{E}[\tilde{X}_{n,1}^4] = \mathbb{E}[(X_{n,1}-\tfrac{1}{n})^4] = \mathbb{E}\left[X_{n,1}^4 - 4X_{n,1}^3\tfrac{1}{n} + 6X_{n,1}^2\tfrac{1}{n^2} - \tfrac{4}{n^3}X_{n,1} + \tfrac{1}{n^4}\right]$$

$$= \frac{24}{n^4} - \frac{4\times 6}{n^4} + \frac{6\times 2}{n^4} - \frac{4}{n^4} + \frac{1}{n^4} = \frac{9}{n^4}$$

D'où

$$\sum_{k=1}^{n} \frac{1}{\sigma_n^2}\mathbb{E}[\tilde{X}_{n,k}^2 \mathbb{1}_{\{|X_{n,k}|\geq \varepsilon\sigma_n\}}] \leq n^2 \frac{3}{n^2}\frac{1}{\varepsilon\sqrt{n}} = \frac{3}{\varepsilon\sqrt{n}}$$

et donc la condition de Lindeberg est vérifiée.

On obtient alors

$$\frac{\sum_{k=1}^{n}(X_{n,k}-\frac{1}{n})}{\frac{1}{\sqrt{n}}} \xrightarrow[n\to+\infty]{\text{Loi}} \mathcal{N}(0,1)$$

c-à-d $\quad \sqrt{n}\,(\Gamma(n,n)-1) \xrightarrow[n\to+\infty]{\text{Loi}} \mathcal{N}(0,1)$ ∎

<u>Rmq</u>: on a besoin de calculer les moments d'une loi exponentielle. Pour cela, il suffit de dériver la fonction caractéristique de cette loi et d'évaluer en 0.

<u>Rmq</u>: a priori, le théorème de Lyapounov n'est peut-être pas suffisant pour obtenir cette convergence en loi.