Bernis (p207)

THÉORÈME CENTRAL LIMITE
ET UNE APPLICATION

261, 262, 266

<u>Thm</u> (central limite) : Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur un espace probabilisé $(X,\mathcal{A},\mathbb{P})$, indépendantes et de même loi admettant une espérance m et une variance $\sigma^2>0$.

$$\sqrt{n}\,\frac{\overline{X}_n - m}{\sigma} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1) \quad \text{où} \quad \overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

<u>Application</u> : Soit $(X_1,\dots,X_n)$ un n-échantillon de loi de Bernoulli $\mathcal{B}(p)$, $0<p<1$. Soit $\alpha\in\,]0,1[$. On déterminera un intervalle de confiance asymptotique au niveau de risque α.

Quitte à centrer et réduire les X_n, $n\in\mathbb{N}$ (i.e. les remplacer par $\frac{X_n-m}{\sigma}$) supposons que $m=0$ et $\sigma=1$. Soit $t\in\mathbb{R}$. Comme les $X_n, n\in\mathbb{N}$ sont mutuellement indépendantes, on a $\forall n\in\mathbb{N}^*$,

$$\varphi_{\sqrt{n}\overline{X}_n}(t) = \mathbb{E}\left[\exp(i\sqrt{n}\,t\,\overline{X}_n)\right] = \mathbb{E}\left[\prod_{k=1}^{n}\exp(i\,t/\sqrt{n}\,X_k)\right] = \prod_{k=1}^{n}\mathbb{E}\left[\exp(i\,t/\sqrt{n}\,X_k)\right] = \prod_{k=1}^{n}\varphi_{X_k}(t/\sqrt{n})$$

puis $\varphi_{\sqrt{n}\overline{X}_n}(t) = \varphi_{X_1}(t/\sqrt{n})^n$ car les X_k, $k\in\mathbb{N}$ sont identiquement distribuées. Comme X_1 est de carré intégrable, φ_{X_1} est deux fois dérivable sur $\mathbb{R}$, et :

$$\varphi'_{X_1}(0) = i\,\mathbb{E}[X_1] = 0 \qquad\qquad \varphi''_{X_1}(0) = i^2\,\mathbb{E}[X_1^2] = -1$$

Ainsi, $\varphi_{\sqrt{n}\overline{X}_n}(t) = \left(1-\frac{1}{2}\left(\frac{t}{\sqrt{n}}\right)^2 + o\left(\left(\frac{t}{\sqrt{n}}\right)^2\right)\right)^n = \left(1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)\right)^n = \left(1-\frac{t^2}{2n}\right)^n\left(1+o\left(\frac{1}{n}\right)\right)^n.$

Classiquement, $\left(1-\frac{t^2}{2n}\right)^n \xrightarrow[n\to+\infty]{} e^{-t^2/2}$. Un argument rapide pour montrer que $(1+o(\frac{1}{n}))^n \longrightarrow 1$ serait d'utiliser un logarithme complexe. Pour esquiver cette subtilité théorique, on procède comme suit : $\left|(1+o(\frac{1}{n}))^n - 1\right| = \left|\sum_{k=1}^{n}\binom{n}{k}\left[o\left(\frac{1}{n}\right)\right]^k\right| \leqslant \sum_{k=1}^{n}\binom{n}{k}\left|o\left(\frac{1}{n}\right)\right|^k = \left(1+\left|o\left(\frac{1}{n}\right)\right|\right)^n - 1$

$$= \exp\left(n\ln\left(1+\left|o\left(\tfrac{1}{n}\right)\right|\right)\right)-1 \xrightarrow[n\to+\infty]{} 0$$

Comme $t\longmapsto e^{-t^2/2}$ est la fonction caractéristique de la $\mathcal{N}(0,1)$, on déduit du théorème de Lévy que $\sqrt{n}\,\overline{X}_n \xrightarrow[n\to+\infty]{} \mathcal{N}(0,1)$. ∎

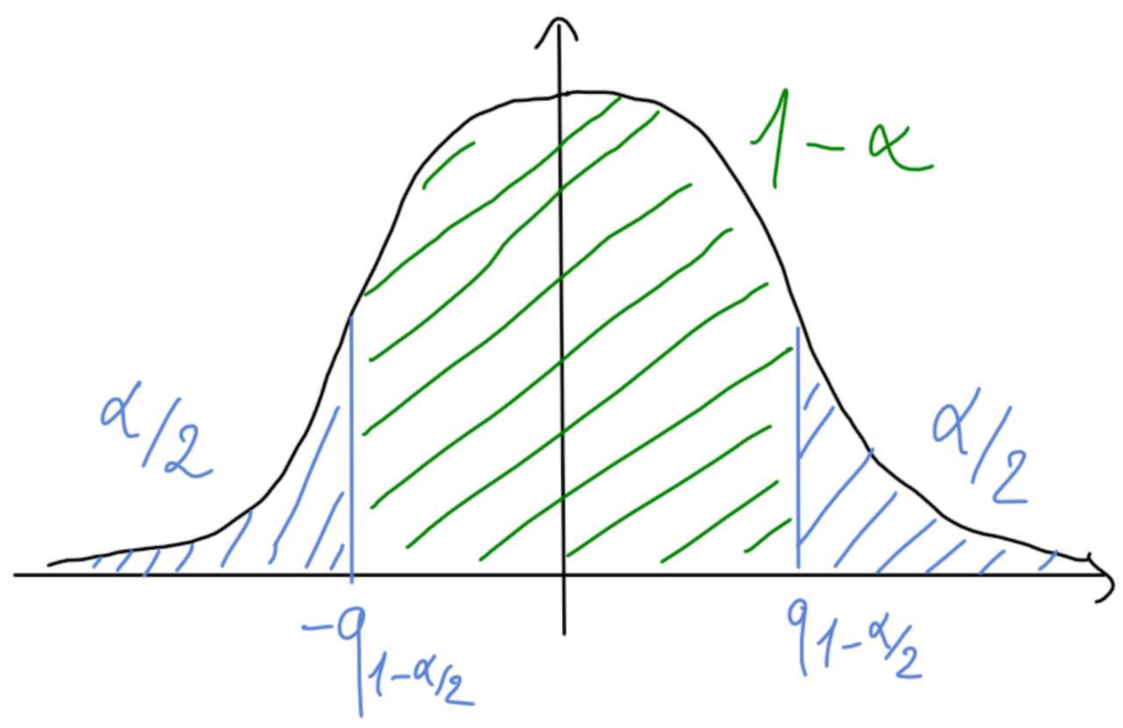
Application : Comme X_1 est intégrable, la loi faible des grands nombres assure que :

$$\overline{X}_n \xrightarrow{\mathbb{P}} \mathbb{E}[X_1] = p \quad \text{donc} \quad \sqrt{\overline{X}_n(1-\overline{X}_n)} \xrightarrow{\mathbb{P}} \sqrt{p(1-p)}$$

D'après le théorème central limite et le lemme de SLUTSKY,

$$\sqrt{n}\,\frac{\overline{X}_n - p}{\sqrt{\overline{X}_n(1-\overline{X}_n)}} = \underbrace{\frac{\sqrt{p(1-p)}}{\sqrt{\overline{X}_n(1-\overline{X}_n)}}}_{\xrightarrow{\mathbb{P}} 1} \underbrace{\frac{\overline{X}_n - p}{\sqrt{p(1-p)}}\sqrt{n}}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0,1)} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Notons $q_{1-\alpha/2}$ le $(1-\alpha/2)$-quantile de $\mathcal{N}(0,1)$, i.e. le réel vérifiant $\mathbb{P}(Z \leq q_{1-\alpha/2}) = 1-\alpha/2$ où $Z \sim \mathcal{N}(0,1)$.
Remarquons que $\mathbb{P}(|Z| \leq q_{1-\alpha/2}) = 1-\alpha$ par symétrie de la $\mathcal{N}(0,1)$ (voir dessin ci-contre).



On a alors, par convergence en loi :

$$\mathbb{P}\left(\left|\sqrt{n}\,\frac{\overline{X}_n - p}{\sqrt{\overline{X}_n(1-\overline{X}_n)}}\right| \leq q_{1-\alpha/2}\right) \xrightarrow[n\to+\infty]{} \mathbb{P}(|Z| \leq q_{1-\alpha/2}) = 1-\alpha$$

et $$\left|\sqrt{n}\,\frac{\overline{X}_n - p}{\sqrt{\overline{X}_n(1-\overline{X}_n)}}\right| \leq q_{1-\alpha/2} \iff \overline{X}_n - \frac{q_{1-\alpha/2}\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}} \leq p \leq \overline{X}_n + \frac{q_{1-\alpha/2}\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}}$$

Ainsi, $\left[\overline{X}_n - \frac{q_{1-\alpha/2}\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}} \,;\, \overline{X}_n + \frac{q_{1-\alpha/2}\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}}\right]$ est un intervalle de confiance asymptotique au niveau de confiance $1-\alpha$ pour p. ■