Une preuve élémentaire du théorème limite central [*]

Jérôme Depauw [†]

Octobre 2009

Résumé

Cet article propose une démonstration élémentaire du théorème limite central, sans transformation de Fourier, inspirée d'un article de H. F. Trotter. Un intérêt de cette preuve, outre sa simplicité, est qu'elle met en lumière l'importance de la propriété de stabilité de la loi de Gauss dans le théorème limite central.
MOTS-CLÉS : *théorème limite central, point fixe.*

1 Introduction

Le théorème limite central est, comme son nom l'indique, un résultat fondamental du calcul des probabilités. Il peut être résumé de la façon suivante : la somme d'un grand nombre de données numériques aléatoires (où « variables aléatoires ») indépendantes et d'ordre de grandeur comparable suit approximativement une loi de Gauss (la loi de Gauss est aussi appelée loi normale). Avant d'en donner un énoncé mathématique, rappelons la définition de la loi de Gauss.

Définition 1 (Loi de Gauss) *Une variable aléatoire Y suit la loi de Gauss de moyenne m et de variance σ^2 si pour tout intervalle $[a,b] \subset \mathbb{R}$, la probabilité de l'évènement $(Y \in [a,b])$ est donnée par*

$$P(Y \in [a,b]) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(y-m)^2}{\sigma^2}} \, \mathrm{d}y.$$

On notera $Y \sim \mathcal{N}(m,\sigma^2)$. La fonction sous l'intégrale est la densité de la loi de Gauss de moyenne m et de variance σ^2.

La courbe de la densité de la loi de Gauss de moyenne m et de variance σ^2 est la fameuse « courbe en cloche », centrée autour de m, et d'autant plus étalée que σ est grand.

[*]Publiée par la Revue de la Filière Mathématiques : RMS **120**, 1, (2009-2010) pages 33-38
[†]Université de Tours, Département de Mathématiques, Parc de Grandmont, 37200 Tours, France. E-mail `jerome.depauw@univ-tours.fr`.

Dans la suite, on utilisera la caractérisation de la loi d'une variable aléatoire X par les valeurs moyennes (où espérances mathématiques) des variables aléatoires de la forme $f(X)$, où f est une fonction continue bornée. Pour Y suivant la loi de Gauss $\mathcal{N}(m, \sigma^2)$, l'espérance mathématique $E(f(Y))$ s'écrit, toujours à l'aide de la densité

$$E(f(Y)) = \int_{-\infty}^{+\infty} f(y) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(y-m)^2}{\sigma^2}} \, \mathrm{d}y.$$

La famille des lois de Gauss est stable par changement d'échelle affine : Si Y suit la loi de Gauss $\mathcal{N}(m, \sigma^2)$, et $a \neq 0$ et b sont des nombres réels, alors $Y' = aY + b$ suit la loi de Gauss $\mathcal{N}(m', \sigma'^2)$ de moyenne $m' = am + b$ et de variance $\sigma'^2 = a^2\sigma^2$. Cela se vérifie aisément par changement de variable $y' = ay + b$ dans l'intégrale de la définition ci-dessus.

Nous pouvons maintenant donner un énoncé du théorème limite central.

Théorème 1 (Théorème limite central) *Soit une suite $(X_n)_n$ de variables aléatoires indépendantes de même loi. On les suppose de moyenne m, et de variance σ^2 finie. Pour toute fonction f continue bornée, on a*

$$E\Big[f\Big(\frac{X_1 + \cdots + X_n - nm}{\sqrt{n}}\Big)\Big] \to E[f(Y)]$$

pour n tendant vers l'infini, où Y suit la loi de Gauss de moyenne nulle et de variance σ^2.

La question qui vient à l'esprit est alors : pourquoi la loi limite est la loi de Gauss ? Une explication est que la loi de Gauss est la loi pour laquelle l'approximation que donne le théorème limite central devient vraie pour tout n (sans passage à la limite), comme c'est exprimé par la proposition suivante.

Proposition 1 (Stabilité de la loi de Gauss) *Soit $(Y_k)_{1 \leqslant k \leqslant n}$ un vecteur aléatoire dont les coordonnées sont des variables aléatoires indépendantes de loi de Gauss de moyenne m, et de variance σ^2. Alors la variable aléatoire*

$$\frac{Y_1 + \cdots + Y_n - nm}{\sqrt{n}}$$

suit la loi de Gauss de moyenne nulle et de variance σ^2.

Ainsi, le théorème limite central peut être interprété comme un théorème de point fixe.

Pour finir cette introduction, notons que le théorème limite central admet le corollaire intuitif suivant.

Corollaire 1 *Soit une suite $(X_n)_n$ de variables aléatoires indépendantes de même loi. On les suppose de moyenne m, et de variance σ^2 finie. Pour tout intervalle $[a, b]$*

$$P\Big(\frac{X_1 + \cdots + X_n - nm}{\sqrt{n}} \in [a, b]\Big) \to P(Y \in [a, b])$$

pour n tendant vers l'infini, où Y suit la loi de Gauss de moyenne nulle et de variance σ^2.

Ce corollaire se déduit du théorème limite central en encadrant, pour tout $\varepsilon > 0$, la fonction indicatrice de l'intervalle $[a, b]$ entre deux fonctions continues bornées $f < \mathbf{1}_{[a,b]} < g$ telles que $E(g(Y) - f(Y)) < \varepsilon$.

On trouve énoncé et démonstration du théorème limite central dans tout traité de calcul des probabilités. Les démonstrations proposées reposent en général sur l'analyse de Fourier. C'est aussi le cas dans la note de L. Gallado et E. Lesigne publiée par la RMS en 2005 (voir [1]). La démonstration présentée ci-dessous est différente. Celle-ci est une version simplifiée de celle de H. F. Trotter (voir [2]). Cette démonstration utilise la propriété de stabilité de la loi Gauss (proposition 1 ci-dessus) comme un ingrédient essentiel. La proposition 1 elle-même est démontré au paragraphe 3.

2 Démonstration du théorème

Quitte à remplacer X_n par $X_n - m$, on peut supposer les X_n centrées. On considère d'abord le cas d'une fonction f de classe $\mathcal{C}^2$ à support compact. Soit $(Y_n)_n$ une suite de variables aléatoires indépendantes, de loi de Gauss $\mathcal{N}(0, \sigma^2)$, indépendante de la première suite. On pose

$$I_n(f) = E\Big[f\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big)\Big] - E\Big[f\Big(\frac{Y_1 + \cdots + Y_n}{\sqrt{n}}\Big)\Big].$$

Comme, d'après la proposition ci-dessus, la loi de $\dfrac{Y_1 + \cdots + Y_n}{\sqrt{n}}$ est la loi de Gauss $\mathcal{N}(0, \sigma^2)$, le théorème est prouvé si l'on montre que $\lim\limits_{n\to\infty} I_n(f) = 0$.

Considérons $U_k = X_1 + \cdots + X_k + Y_{k+1} + \cdots + Y_n$. On a par simplification

$$\begin{aligned} f\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big) - f\Big(\frac{Y_1 + \cdots + Y_n}{\sqrt{n}}\Big) &= f\Big(\frac{U_n}{\sqrt{n}}\Big) - f\Big(\frac{U_0}{\sqrt{n}}\Big) \\ &= \sum_{k=1}^{n} f\Big(\frac{U_k}{\sqrt{n}}\Big) - f\Big(\frac{U_{k-1}}{\sqrt{n}}\Big) \end{aligned}$$

Pour tout k, réécrivons le terme correspondant de la somme ci-dessus :

$$f\Big(\frac{U_k}{\sqrt{n}}\Big) - f\Big(\frac{U_{k-1}}{\sqrt{n}}\Big) = f\Big(Z_k + \frac{X_k}{\sqrt{n}}\Big) - f\Big(Z_k + \frac{Y_k}{\sqrt{n}}\Big)$$

où $Z_k = \dfrac{1}{\sqrt{n}}(X_1 + \cdots + X_{k-1} + Y_{k+1} + \cdots + Y_n)$. D'après la formule de Taylor-Lagrange il existe des quantités $|s_{X_k,Z_k}| \leqslant |X_k/\sqrt{n}|$, et $|t_{Y_k,Z_k}| \leqslant |Y_k/\sqrt{n}|$ telles que cela soit encore égal à

$$= f'(Z_k)\frac{X_k}{\sqrt{n}} + \frac{f''(s_{X_k,Z_k})}{2}\frac{X_k^2}{n} - f'(Z_k)\frac{Y_k}{\sqrt{n}} - \frac{f''(t_{Y_k,Z_k})}{2}\frac{Y_k^2}{n}$$

On introduit $f''(Z_k)$ pour réécrire cela sous la forme

$$= f'(Z_k)\frac{X_k - Y_k}{\sqrt{n}} + \frac{f''(Z_k)}{2}\frac{X_k^2 - Y_k^2}{n} + \\ + \frac{f''(s_{X_k,Z_k}) - f''(Z_k)}{2}\frac{X_k^2}{n} - \frac{f''(t_{Y_k,Z_k}) - f''(Z_k)}{2}\frac{Y_k^2}{n}.$$

Comme f est de classe $\mathcal{C}^2$, et à support compact, f'' est uniformément continue. Soit $\varepsilon > 0$. Il existe donc un $\delta > 0$ tel que, si $|z-s| < \delta$, alors $|f''(z)-f''(s)| < \varepsilon$. Si $|X_k| < \delta\sqrt{n}$, alors $|s_{X_k,Z_k}| < \delta$, et $|f''(Z_k + s_{X_k,Z_k}) - f''(Z_k)| < \varepsilon$. Ainsi on obtient

$$f\Big(\frac{U_k}{\sqrt{n}}\Big) - f\Big(\frac{U_{k-1}}{\sqrt{n}}\Big) = f'(Z_k)\frac{X_k - Y_k}{\sqrt{n}} + \frac{f''(Z_k)}{2}\frac{X_k^2 - Y_k^2}{n} + R \qquad (1)$$

où la variable aléatoire R vérifie $R \leqslant \dfrac{\varepsilon}{2n}(X_k^2 + Y_k^2)$ si $|X_k|,\ |Y_k| < \delta\sqrt{n}$. En notant $\mathbf{1}_{|X_k|>\delta\sqrt{n}}$ la variable aléatoire valant 1 si l'événement $|X_k| > \delta\sqrt{n}$ a lieu, et 0 sinon, et en notant $\|f''\|_\infty = \sup_z |f''(z)|$, cela se réécrit

$$|R| \leqslant \frac{1}{n}\Big[\|f''\|_\infty\Big(X_k^2\mathbf{1}_{|X_k|>\delta\sqrt{n}} + Y_k^2\mathbf{1}_{|Y_k|>\delta\sqrt{n}}\Big) + \frac{\varepsilon}{2}(X_k^2 + Y_k^2)\Big].$$

Calculons l'espérance de ces variables aléatoires. Par indépendance des variables aléatoires X_k, Y_k et Z_k, et les égalités $E(X_k) = E(Y_k) = 0$ et $E(X_k^2) = E(Y_k^2) = \sigma^2$, l'espérance des deux premiers termes du membre de droite de l'égalité (1) ci-dessus est nulle. Il reste

$$\Big|E\Big[f\Big(\frac{U_k}{\sqrt{n}}\Big) - f\Big(\frac{U_{k-1}}{\sqrt{n}}\Big)\Big]\Big| \leqslant \\ \frac{1}{n}\Big[\|f''\|_\infty E\Big(X_k^2\mathbf{1}_{|X_k|>\delta\sqrt{n}} + Y_k^2\mathbf{1}_{|Y_k|>\delta\sqrt{n}}\Big) + \varepsilon\sigma^2\Big].$$

Cela ne dépend pas de k. Donc pour la somme sur $k = 1, \ldots, n$ on a

$$|I_n(f)| \leqslant \|f''\|_\infty\Big(E(X_1^2\mathbf{1}_{|X_1|>\delta\sqrt{n}}) + E(Y_1^2\mathbf{1}_{|Y_1|>\delta\sqrt{n}})\Big) + \varepsilon\sigma^2.$$

Pour n tendant vers l'infini, on voit que, d'après le théorème de convergence dominée, $\overline{\lim}_{n\to\infty} |I_n(f)| \leqslant \varepsilon\sigma^2$. Comme ε est aussi petit que l'on veut, la convergence souhaitée est démontrée pour les fonctions de classe $\mathcal{C}^2$ à support compact. Le passage aux fonctions continues à support compact se fait aisément par un argument d'approximation uniforme. Pour le passage aux fonctions continues bornées on remarque que, pour tout $M > 0$, l'inégalité de Bienaymé-Tchebycheff donne les majorations

$$P\Big(\Big|\frac{U_0}{\sqrt{n}}\Big| \geqslant M\Big) \leqslant \frac{\sigma^2}{M^2} \qquad \text{et} \qquad P\Big(\Big|\frac{U_n}{\sqrt{n}}\Big| \geqslant M\Big) \leqslant \frac{\sigma^2}{M^2}.$$

Si f est continue bornée, et $|g_M| \leqslant |f|$ est continue à support compact, et coincide avec f sur $[-M, M]$, alors $|I_n(f) - I_n(g_M)| \leqslant 4\|f\|_\infty\dfrac{\sigma^2}{M^2}$. Cette majoration est uniforme en n. La démonstration de théorème s'achève en laissant n puis M tendre vers l'infini.

3 Démonstration de la proposition 1

La proposition 1 est une conséquance immédiate des propriétés de changements d'échelles affines évoquées dans l'introduction, et de la proposition suivante.

Proposition 2 *Soient* Y *et* Y' *deux variables aléatoires suivant respectivement les lois de Gauss* $\mathcal{N}(m,\sigma^2)$ *et* $\mathcal{N}(m',\sigma'^2)$*. Si elles sont indépendantes, alors la variable aléatoire* $Y+Y'$ *suit la loi de Gauss* $\mathcal{N}(m+m',\sigma^2+\sigma'^2)$*.*

Démonstration.— L'expression de la densité de la loi de Gauss a été rappelée dans la définition 1. Or la densité ϕ de la somme $Y+Y'$ de deux variables aléatoires Y et Y' indépendantes de densité respectivement φ et ψ est donnée par le produit de convolution

$$\phi(u) = \int_v \varphi(u-v)\psi(v)\ dv. \qquad (2)$$

Quitte à remplacer Y et Y' par $Y-m$ et $Y'-m'$, on peut supposer que $m=m'=0$. D'après la formule (2) ci-dessus et l'expression de la densité de la loi de Gauss (voir définition 1), on a

$$\phi(u) = \frac{1}{2\pi\sigma\sigma'}\int e^{-\frac{1}{2}\left(\frac{(u-v)^2}{\sigma^2}+\frac{v^2}{\sigma'^2}\right)}\ dv.$$

Or l'exposant de l'exponentiel se réécrit, au facteur $-1/2$ près :

$$\begin{aligned} \frac{(u-v)^2}{\sigma^2}+\frac{v^2}{\sigma'^2} &= \frac{\sigma'^2u^2-2\sigma'^2uv+(\sigma^2+\sigma'^2)v^2}{\sigma^2\sigma'^2} \\ &= \frac{u^2}{\sigma^2+\sigma'^2}+\frac{\sigma^2+\sigma'^2}{\sigma'^2\sigma^2}\left(v-\frac{\sigma'^2}{\sigma^2+\sigma'^2}u\right)^2 \end{aligned}$$

donc $\phi(u) = \dfrac{e^{-\frac{1}{2}\frac{u^2}{\sigma^2+\sigma'^2}}}{2\pi\sigma\sigma'}\displaystyle\int e^{-\frac{1}{2}\frac{\sigma^2+\sigma'^2}{\sigma'^2\sigma^2}\left(v-\frac{\sigma'^2}{\sigma^2+\sigma'^2}u\right)^2}\ dv$. Sous l'intégrale en v, on reconnait, au facteur multiplicatif $\sqrt{\dfrac{\sigma^2+\sigma'^2}{2\pi\sigma'^2\sigma^2}}$ près, la densité d'une loi de Gauss (dont la moyenne dépend de u). L'intégrale d'une densité valant 1, on obtient finalement

$$\phi(u) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+\sigma'^2}}e^{-\frac{1}{2}\frac{u^2}{\sigma^2+\sigma'^2}}.$$

Donc U suit la loi $\mathcal{N}(0,\sigma^2+\sigma'^2)$, ce qui démontre la proposition 2, et la proposition 1.

Références

[1] L. Gallardo, E. Lesigne. *Approximation gaussienne de la loi binomiale.* RMS **116**, 1, (2005) pages 10-16

[2] H. F. Trotter. *An elementary proof of the central limit theorem.* Arch. Math. **10** (1959) pages 226–234.