

Le théorème de Moivre-Laplace.

Une démonstration complète dans le cas $p = 1/2$.

1 - Enoncé du théorème.

2 - Démonstration du théorème de Moivre-Laplace lorsque $p = 1/2$.

a - Les étapes de la démonstration. b - Convergence de $f_n(t)$ vers $\frac{e^{-t^2/2}}{\sqrt{2\pi}}$.

4 - Calcul de l'intégrale de Gauss.

5 - Démonstration du théorème de Moivre-Laplace lorsque $p \neq 1/2$.

6 - La loi des erreurs.

1) Enoncé du théorème.

Théorème 1. *On suppose que pour tout n , X_n suit la loi binomiale $\mathcal{B}(n, p)$ avec $p \in]0, 1[$.*

On pose $Z_n = \frac{X_n - np}{\sqrt{pqn}}$ avec $q = 1 - p$. Alors pour tous réels a et b tels que $a < b$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(a \leq Z_n \leq b) = \int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}}$$

Nous commençons par le cas $p = 1/2$. Le résultat suivant se démontre avec des outils accessibles au niveau terminale (rappelons que cette démonstration est hors programme).

Théorème 2. *On suppose que pour tout n , X_n suit la loi binomiale $\mathcal{B}(n, 1/2)$. On pose*

$Z_n = \frac{X_n - n/2}{\sqrt{n}/2}$. Alors pour tous réels a et b vérifiant $a < b$, il existe un réel $K > 0$

indépendant de n tel que

$$\left| \mathbb{P}(a \leq Z_n \leq b) - \int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}} \right| \leq \frac{K}{\sqrt{n}} \quad (1)$$

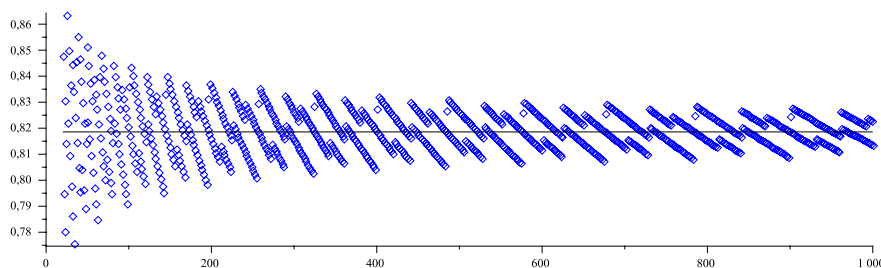


FIGURE 1 – La suite $(\mathbb{P}(a \leq Z_n \leq b))$ converge vers $\int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}}$ ($a = -1, b = 2$).

Remarquons que X_n prend des valeurs entières k comprises entre 0 et n , et donc Z_n prend les valeurs $t_k = \frac{k - n/2}{\sqrt{n}/2}$. La vitesse de convergence en $1/\sqrt{n}$ est optimale puisque si a et b sont dans le même intervalle $[t_k, t_{k+1}]$ avec $a > 0$ et $b - a \geq 1/\sqrt{n}$ (possible car $t_{k+1} - t_k = 2/\sqrt{n}$), alors $P(a \leq Z_n \leq b) = 0$ et donc

$$\left| \mathbb{P}(a \leq Z_n \leq b) - \int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}} \right| = \int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}} \geq \frac{e^{-b^2/2}}{\sqrt{2\pi n}}$$

2) Démonstration du théorème de Moivre-Laplace lorsque $p = 1/2$.

a) Les étapes de la démonstration.

Nous allons construire une fonction en escalier f_n telle que $\mathbb{P}(a \leq Z_n \leq b) = \int_{a_n}^{b_n} f_n(t) dt$, où a_n est proche de a et b_n est proche de b ($\mathbb{P}(a \leq Z_n \leq b)$ est l'aire du domaine représenté sur la figure 1 lorsque $a = -1$, $b = 2$ et $n = 192$), et telle que $f_n(t)$ est proche de $\frac{e^{-t^2/2}}{\sqrt{2\pi}}$.

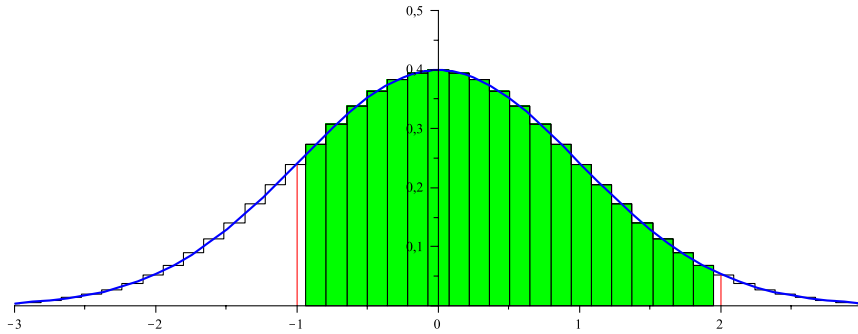


FIGURE 2 – Représentation graphique de f_n et de la fonction $t \rightarrow \frac{e^{-t^2/2}}{\sqrt{2\pi}}$

Définissons tout d'abord la fonction en escalier f_n : pour tout k tel que $0 \leq k \leq n$, f_n est constante égale à $\frac{\sqrt{n}}{2} \mathbb{P}(X_n = k)$ sur $J_k = [t_k - \frac{1}{\sqrt{n}}, t_k + \frac{1}{\sqrt{n}}[$, et f_n est nulle en dehors de l'union de ces intervalles.

Proposition 3. Si $t_k = \frac{k - n/2}{\sqrt{n}/2}$, $k' = \min\{k; t_k \in [a, b]\}$ et $k'' = \max\{k; t_k \in [a, b]\}$,

$$a_n = t_{k'} - \frac{1}{\sqrt{n}} \text{ et } b_n = t_{k''} + \frac{1}{\sqrt{n}}, \text{ alors } \mathbb{P}(a \leq Z_n \leq b) = \int_{a_n}^{b_n} f_n(t) dt.$$

Preuve. Par la formule de l'aire d'un rectangle, $\int_{t_k - 1/\sqrt{n}}^{t_k + 1/\sqrt{n}} f_n(t) dt = \mathbb{P}(X_n = k) = \mathbb{P}(Z_n = t_k)$.

On a donc :

$$\mathbb{P}(a \leq Z_n \leq b) = \mathbb{P}(t_{k'} \leq Z_n \leq t_{k''}) = \sum_{k=k'}^{k''} \mathbb{P}(Z_n = t_k) = \sum_{k=k'}^{k''} \int_{t_k - 1/\sqrt{n}}^{t_k + 1/\sqrt{n}} f_n(t) dt = \int_{a_n}^{b_n} f_n(t) dt.$$

Proposition 4. Si $A = \sup\{|a|, |b|\}$ et $n > A^2$, alors

$$I = \left| \int_{a_n}^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt - \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \leq \frac{1}{\sqrt{n}}$$

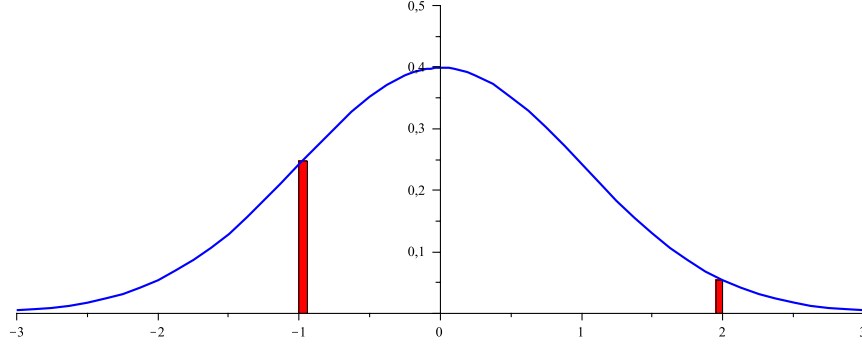


FIGURE 3 – L'aire de la réunion des deux domaines représentés est égal à I ($a = -1$, $b = 2$, $n = 192$).

Preuve. En effet,

$$\left| \int_{a_n}^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt - \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \leq \left| \int_{a_n}^a \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| + \left| \int_b^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| = I_1 + I_2$$

Comme $n > A^2$, alors $t_0 = -\sqrt{n} < a$ et $t_n = \sqrt{n} > b$, donc $1 \leq k' \leq k'' \leq n - 1$, d'où $t_{k'-1} = t_{k'} - \frac{2}{\sqrt{n}} < a \leq t_{k'}$, et comme $a_n = t_{k'} - \frac{1}{\sqrt{n}}$, on a $|a - a_n| \leq \frac{1}{\sqrt{n}}$. De plus, $\left| \frac{e^{-t^2/2}}{\sqrt{2\pi}} \right| \leq 1/2$, par conséquent $I_1 \leq \frac{1}{2\sqrt{n}}$. De même, $|b - b_n| \leq \frac{1}{\sqrt{n}}$ et $I_2 \leq \frac{1}{2\sqrt{n}}$.

Le résultat suivant montre que f_n est proche de $e^{-t^2/2}/\sqrt{2\pi}$ et est démontré en section 3.

Proposition 5. Il existe $C > 0$ tel que si $t \in [a_n, b_n]$ et $n > 4A^2$, alors $\left| f_n(t) - \frac{e^{-t^2/2}}{\sqrt{2\pi}} \right| \leq \frac{C}{\sqrt{n}}$.

Le théorème de Moivre-Laplace découle facilement de ces propositions. En effet, d'après la proposition 3 et l'inégalité triangulaire,

$$\begin{aligned} \left| \mathbb{P}(a \leq Z_n \leq b) - \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| &\leq \left| \int_{a_n}^{b_n} f_n(t) dt - \int_{a_n}^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \\ &\quad + \left| \int_{a_n}^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt - \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \end{aligned}$$

D'après la proposition 4, le deuxième terme de droite est majoré par $1/\sqrt{n}$, et d'autre part, d'après la proposition 5,

$$\left| \int_{a_n}^{b_n} f_n(t) dt - \int_{a_n}^{b_n} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \leq \int_{a_n}^{b_n} \left| f_n(t) - \frac{e^{-t^2/2}}{\sqrt{2\pi}} \right| dt \leq \int_{a_n}^{b_n} \frac{C}{\sqrt{n}} dt \leq \frac{C(b - a + 2)}{\sqrt{n}}$$

b) Convergence de $f_n(t)$ vers $\frac{e^{-t^2/2}}{\sqrt{2\pi}}$.

Nous allons montrer la proposition 5 : il existe $C > 0$ tel que si $n > 4A^2$, alors pour tout $t \in [a_n, b_n]$, $\left|f_n(t) - \frac{e^{-t^2/2}}{\sqrt{2\pi}}\right| \leq \frac{C}{\sqrt{n}}$. Ceci résulte des deux propositions suivantes lorsque n est pair. Le cas n impair est traité en fin de section.

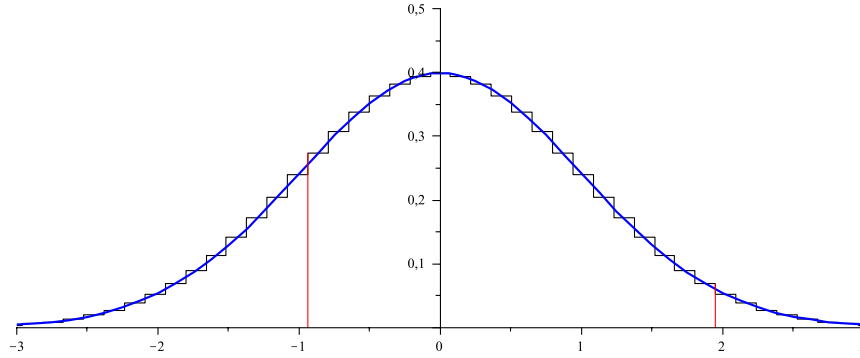


FIGURE 4 – Comparaison des graphes de f_n et de la fonction $x \rightarrow \exp(-x^2/2)/\sqrt{2\pi}$.

Proposition 6. Si $n \geq 2$ est pair, alors $1 - \frac{1}{n} \leq \sqrt{2\pi}f_n(0) \leq 1$.

Ceci implique, d'après le théorème des gendarmes, que $\lim_{n \rightarrow +\infty, n \text{ pair}} f_n(0) = \frac{1}{\sqrt{2\pi}}$.

Proposition 7. Si $n \geq 2$ est pair, alors $f_n(t_k) \leq f_n(0)$. Si de plus $n \geq 4A^2$ et $|t_k| \leq A$, il existe un réel $L > 0$ tel que

$$\left| \frac{f_n(t_k)}{f_n(0)} - e^{-t_k^2/2} \right| \leq \frac{L}{\sqrt{n}}$$

Preuve de la proposition 5 pour n pair.

Si $n \geq 4A^2$ et $t \in [a_n, b_n]$, il existe k tel que $k' \leq k \leq k''$ et $t \in J_k$ car l'union des J_k pour $k' \leq k \leq k''$ est égale à $[a_n, b_n]$. Par conséquent, $f_n(t) = f_n(t_k)$ et $|e^{-t_k^2/2} - e^{-t^2/2}| \leq |t - t_k| \leq \frac{1}{\sqrt{n}}$. Comme $k' \leq k \leq k''$, on a $|t_k| \leq A$, et donc, d'après la proposition 6 et la proposition 7,

$$\begin{aligned} \left| \sqrt{2\pi}f_n(t) - e^{-t^2/2} \right| &\leq \frac{f_n(t_k)}{f_n(0)} \left| \sqrt{2\pi}f_n(0) - 1 \right| + \left| \frac{f_n(t_k)}{f_n(0)} - e^{-t_k^2/2} \right| + \left| e^{-t_k^2/2} - e^{-t^2/2} \right| \\ &\leq \frac{1}{n} + \frac{L}{\sqrt{n}} + \frac{1}{\sqrt{n}} \leq \frac{L+2}{\sqrt{n}} \end{aligned}$$

Nous démontrons maintenant les propositions 6 et 7. Les valeurs de la fonction f_n font intervenir les coefficients $p_{n,k} = \mathbb{P}(X_n = k)$ où X_n suit la loi binomiale $\mathcal{B}(n, 1/2)$. Le lemme suivant donne une propriété des $p_{n,k}$ (propriété évidente si on sait que $\binom{n}{k} = \frac{n!}{k!(n-k)!}$).

1. voir exercice en partie III

Lemme 8. $p_{n,k+1} = \frac{n-k}{k+1} p_{n,k}$.

Démonstration par récurrence sur n . Comme $p_{n,k} = \binom{n}{k}/2^n$, il suffit d'établir :

$$(P_n) \quad \text{pour tout } 0 \leq k \leq n-1, \quad \binom{n}{k+1} = \frac{n-k}{k+1} \binom{n}{k}$$

La propriété (P_1) est vérifiée puisque $p_{1,1} = p_{1,0} = 1$. Si (P_n) est vérifié et si $0 \leq k \leq n$, on a alors, en utilisant la relation de Pascal $\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}$ et la relation (P_n) :

$$\begin{aligned} (k+1) \binom{n+1}{k+1} &= (k+1) \binom{n}{k+1} + (k+1) \binom{n}{k} = (n-k) \binom{n}{k} + (k+1) \binom{n}{k} \\ &= (n-k+1) \binom{n}{k} + k \binom{n}{k} = (n-k+1) \binom{n}{k} + (n-(k-1)) \binom{n}{k-1} \\ &= (n+1-k) \binom{n+1}{k} \end{aligned}$$

Lemme 9. Soit $p_m = p_{2m,m}$ la probabilité d'obtenir m piles après avoir lancé une pièce $2m$ fois. Alors $\frac{1}{1 + \frac{1}{2m}} \leq m\pi p_m^2 \leq 1$.

Notons $m!$ le produit des entiers de 1 à m . En appliquant le lemme 8, on obtient

$$p_m = p_{2m,m} = \frac{m+1}{m} p_{2m,m-1} = \dots = \frac{(m+1)(m+2) \dots 2m}{m(m-1) \dots 1} p_0 = \frac{(2m)!}{2^{2m}(m!)^2}$$

Posons $I_m = \int_0^1 (1-t^2)^{\frac{m}{2}} dt$. La suite (I_m) est décroissante.

En effet, si $t \in [0, 1]$, alors $(1-t^2)^{\frac{m+1}{2}} \leq (1-t^2)^{\frac{m}{2}}$. En intégrant, $I_{m+1} \leq I_m$.

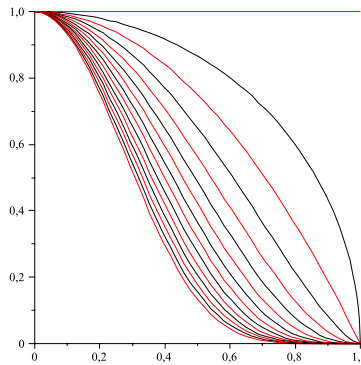


FIGURE 5 – graphe des fonctions $t \rightarrow (1-t^2)^{m/2}$.

On a $I_0 = 1$, $I_1 = \pi/4$ (aire du quart du disque unité), et la relation de récurrence $I_m = \frac{m}{m+1} I_{m-2}$.

En effet, si $f(t) = t(1 - t^2)^{\frac{m}{2}}$, il suffit de calculer $f'(t)$ et d'intégrer la relation obtenue entre 0 et 1.

En itérant la relation de récurrence,

$$I_{2m} = \frac{2m}{2m+1} I_{2m-1} = \frac{2m}{2m+1} \frac{2m-2}{2m-1} \cdots \frac{2}{3} I_0 = \frac{2^{2m}(m!)^2}{(2m+1)!} = \frac{1}{(2m+1)p_m} \quad (2)$$

de même,

$$I_{2m-1} = \frac{2m-1}{2m} \frac{2m-3}{2m-2} \cdots \frac{3}{4} I_1 = \frac{(2m)!}{[2m(2m-2)(\cdots)2]^2} \frac{\pi}{2} = \frac{(2m)!}{2^{2m}(m!)^2} \frac{\pi}{2} = \frac{\pi}{2} p_m$$

et donc $\frac{I_{2m-1}}{I_{2m}} = (1 + \frac{1}{2m}) m \pi p_m^2$. Comme (I_m) est décroissante, $I_{2m} \leq I_{2m-1} \leq I_{2m-2}$, et

en divisant par I_{2m} , on a $1 \leq (1 + \frac{1}{2m}) m \pi p_m^2 \leq \frac{I_{2m-2}}{I_{2m}} = 1 + \frac{1}{2m}$, d'où le résultat.

Preuve de la proposition 6. Si $n = 2m$, alors $(1 - \frac{1}{n})^2 \leq \frac{1}{1 + \frac{1}{n}} \leq \frac{n\pi}{2} p_m^2 \leq 1$. Ceci résulte du lemme 9. Or, $f_n(0) = \frac{\sqrt{n}}{2} p_m$. On conclut en prenant la racine carrée.

Nous utiliserons dans la démonstration de la proposition 7 les résultats suivants :

Lemme 10. $\frac{1-x}{1+x} = \exp\{-2x + \varepsilon(x)\}$, avec $|\varepsilon(x)| \leq x^3$ si $|x| \leq 1/2$.

Un calcul donne $\varepsilon(u) = \ln(1-u) - \ln(1+u) + 2u$ et $\varepsilon'(u) = \frac{-2u^2}{1-u^2}$. Si $|u| \leq 1/2$, on a $-3u^2 \leq \varepsilon'(u) \leq 3u^2$. En intégrant entre 0 et x , on a $-x^3 \leq \varepsilon(x) \leq x^3$.

Lemme 11. Si $\ell \geq 1$, alors $\sum_{i=1}^{\ell-1} i = \frac{\ell(\ell-1)}{2}$ et $\sum_{i=1}^{\ell-1} i^3 \leq \ell^4/4$ (par récurrence).

Il résulte de ces deux lemmes que si $\ell, m \in \mathbb{N}$ et si $1 \leq \ell \leq m/2$ et si $v_\ell = \frac{\prod_{i=1}^{\ell-1} (1 - \frac{i}{m})}{\prod_{i=1}^{\ell-1} (1 + \frac{i}{m})}$, alors

$$v_\ell = \exp\left(\frac{-\ell^2}{m} + \frac{\ell}{m} + \sum_{i=1}^{\ell-1} \varepsilon(i/m)\right), \text{ avec } \left| \sum_{i=1}^{\ell-1} \varepsilon(i/m) \right| \leq \sum_{i=1}^{\ell-1} |\varepsilon(i/m)| \leq \frac{1}{m^3} \sum_{i=1}^{\ell-1} i^3 \leq \frac{\ell^4}{4m^3}.$$

Preuve de la proposition 7. Si $n = 2m$, alors $\frac{f_n(t_k)}{f_n(0)} = \frac{p_{n,k}}{p_m}$. Comme $\binom{n}{k} = \binom{n}{n-k}$, on peut supposer que $k > n/2$. Le produit des relations $p_{2m,i+1} = \frac{2m-i}{i+1} p_{2m,i}$ pour i variant de m à $k-1$, donne :

$$\frac{f_n(t_k)}{f_n(0)} = \frac{p_{n,k}}{p_m} = \frac{p_{n,k}}{p_{2m,m}} = \frac{(2m-k+1)(2m-k+2)(\cdots)m}{(m+1)(m+2)(\cdots)k}$$

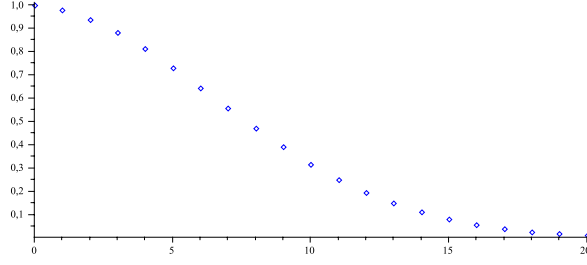


FIGURE 6 – La suite (v_ℓ) (ici $m = 96$).

En divisant chaque terme du numérateur et du dénominateur par m , $\frac{f_n(t)}{f_n(0)} = \frac{m}{k} \cdot v_{k-m}$. Ceci implique $f_n(t) \leq f_n(0)$, ce qui est également vérifié si t n'est dans aucun J_k car dans ce cas $f_n(t) = 0$. Supposons de plus que $|t_k| \leq A$ et $n \geq 4A^2$.

Comme $k - m = \sqrt{n}t_k/2$, $(k - m)/m = t_k/\sqrt{n} \leq 1/2$ et $(k - m)^2/m = t_k^2/2$, on a :

$$\left| \frac{f_n(t_k)}{f_n(0)} - v_{k-m} \right| = \frac{k - m}{k} v_{k-m} \leq \frac{k - m}{m} = \frac{t_k}{\sqrt{n}} \leq \frac{A}{\sqrt{n}}$$

et

$$v_{k-m} = \exp \left\{ -\frac{t_k^2}{2} \right\} \cdot \exp \left\{ \frac{t_k}{\sqrt{n}} + \delta_k \right\} \quad \text{avec} \quad |\delta_k| \leq \frac{(k - m)^4}{4m^3} = \frac{t_k^4}{8n} \leq \frac{t_k^4}{8\sqrt{n}}$$

Par conséquent, $|v_{k-m} - e^{-t_k^2/2}| \leq \exp \left\{ (t_k + t_k^4/8)/\sqrt{n} \right\} - 1 \leq \exp \left\{ C/\sqrt{n} \right\} - 1$, avec $C = A + A^4/8$. D'après l'inégalité de convexité $e^x - 1 \leq xe^x$ pour $x > 0$,

$$|v_{k-m} - e^{-t_k^2/2}| \leq K/\sqrt{n} \quad \text{avec} \quad K = Ce^C$$

D'où,

$$\left| \frac{f_n(t_k)}{f_n(0)} - e^{-t_k^2/2} \right| \leq \left| \frac{f_n(t_k)}{f_n(0)} - v_{k-m} \right| + |v_{k-m} - e^{-t_k^2/2}| \leq \frac{A + K}{\sqrt{n}} = \frac{L}{\sqrt{n}}$$

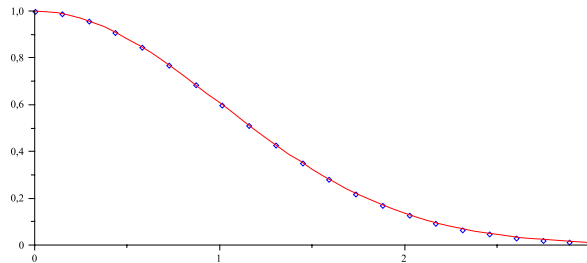


FIGURE 7 – La fonction $t \rightarrow e^{-t^2/2}$ et la suite de points $(t_k, f_n(t_k)/f_n(0))$ (ici $n = 192$).

Preuve de la proposition 5 pour n impair. On va se ramener au cas pair de la façon suivante. Rappelons que $t_k = \frac{k-n/2}{\sqrt{n}/2}$ et notons $\tau_k = \frac{k-(n-1)/2}{\sqrt{n-1}/2}$ la subdivision associée au

rang $n - 1$ ($n - 1$ est pair). Pour tout $1 \leq k \leq n - 1$, on a $\tau_{k-1} \leq t_k \leq \tau_k$. Si $t \in J_k$, alors $|t - \tau_k| \leq |t - t_k| + |t_k - \tau_k| \leq 3/\sqrt{n}$ et $|t - \tau_{k-1}| \leq 3/\sqrt{n}$. D'après la relation de Pascal,

$$f_n(t) = \frac{\sqrt{n}}{2} p_{n,k} = \frac{\sqrt{n}}{4} (p_{n-1,k} + p_{n-1,k-1}) = \frac{1}{2\sqrt{1-1/n}} (f_{n-1}(\tau_k) + f_{n-1}(\tau_{k-1}))$$

Or $\frac{1}{\sqrt{1-1/n}} - 1 \leq 1/n$, et d'après la proposition 6, $\sqrt{2\pi} f_{n-1}(\tau_k) \leq 1$ et $\sqrt{2\pi} f_{n-1}(\tau_{k-1}) \leq 1$, donc $0 \leq \sqrt{2\pi} (2f_n(t) - f_{n-1}(\tau_k) - f_{n-1}(\tau_{k-1})) \leq 2/n$. Par conséquent

$$2|\sqrt{2\pi} f_n(t) - e^{-t^2/2}| \leq \frac{2}{n} + |\sqrt{2\pi} f_{n-1}(\tau_k) - e^{-\tau_k^2/2}| + |\sqrt{2\pi} f_{n-1}(\tau_{k-1}) - e^{-\tau_{k-1}^2/2}| + |e^{-\tau_k^2/2} - e^{-t^2/2}| + |e^{-\tau_{k-1}^2/2} - e^{-t^2/2}|$$

D'après le cas pair, $|\sqrt{2\pi} f_{n-1}(\tau_k) - e^{-\tau_k^2/2}| \leq C/\sqrt{n}$ et $|\sqrt{2\pi} f_{n-1}(\tau_{k-1}) - e^{-\tau_{k-1}^2/2}| \leq C/\sqrt{n}$. De plus, $|e^{-\tau_k^2/2} - e^{-t^2/2}| \leq |\tau_k - t| \leq 3/\sqrt{n}$ et $|e^{-\tau_{k-1}^2/2} - e^{-t^2/2}| \leq |\tau_{k-1} - t| \leq 3/\sqrt{n}$. Il existe donc une constante C' telle que $\left| f_n(t) - \frac{e^{-t^2}}{\sqrt{2\pi}} \right| \leq \frac{C'}{\sqrt{n}}$.

4) Calcul de l'intégrale de Gauss.

Théorème 12. $\lim_{x \rightarrow +\infty} \int_{-x}^x e^{-t^2/2} dt = \sqrt{2\pi}$.

Si $G(x) = \int_{-x}^x e^{-t^2/2} dt$, il suffit de montrer que si $n = n(x)$ la partie entière de $x^2/2$, alors

$$\frac{\sqrt{2\pi}}{1+1/2n} \leq G(x) \leq \sqrt{2\pi}$$

En effet, comme $\lim_{x \rightarrow +\infty} n(x) = +\infty$, le théorème des gendarmes permet de conclure que $\lim_{n \rightarrow \infty} G(x)$ existe et est égal à $\sqrt{2\pi}$.

Majoration. Pour tout k , $\mathbb{P}(-x \leq Z_k \leq x) \leq 1$, donc, d'après le théorème de Moivre-Laplace :

$$G(x) = \sqrt{2\pi} \lim_{k \rightarrow \infty} \mathbb{P}(-x \leq Z_k \leq x) \leq \sqrt{2\pi}$$

Minoration. Notons que $\sqrt{2n} \leq x$. Par ailleurs, pour $0 < v < 1$, on a $(1-v)^n \leq e^{-nv}$ (pour le voir, passer au logarithme et utiliser la concavité de la fonction logarithme). Par conséquent, avec le changement de variable $u = t/\sqrt{2n}$,

$$\int_0^x e^{-t^2/2} dt \geq \int_0^{\sqrt{2n}} e^{-t^2/2} dt = \sqrt{2n} \int_0^1 e^{-nu^2} du \geq \sqrt{2n} \int_0^1 (1-u^2)^n du = \sqrt{2n} I_{2n}$$

Par conséquent, d'après (2) et le lemme 9,

$$G(x) \geq 2\sqrt{2n} I_{2n} = \frac{2\sqrt{2n}}{(2n+1)p_n} \geq \frac{\sqrt{2\pi}}{1+\frac{1}{2n}}$$

5) Démonstration du théorème de Moivre-Laplace lorsque $p \neq 1/2$.

Notre but est de démontrer le théorème 1. Supposons que X_n suit la loi binomiale $\mathcal{B}(n, p)$ et posons $Z_n = \frac{X_n - np}{\sqrt{pqn}}$ avec $q = 1 - p$. Si $X_n = k$, alors $Z_n = t_k = h_n(k - np)$ avec $h_n = \frac{1}{\sqrt{pqn}}$, et $\mathbb{P}(Z_n = t_k) = \mathbb{P}(X_n = k) = p_{n,k} = \binom{n}{k} p^k q^{n-k}$. Soit f_n la fonction définie par $f_n(t) = \frac{1}{h_n} \sum_{k=0}^n p_{n,k} \cdot \mathbb{1}_{[t_k - h_n/2, t_k + h_n/2]}(t)$. Le graphique suivant représente f_n pour $n = 192$ et $p = 0,2$. La convergence est plus lente que pour $p = 0,5$ (rectangles plus gros) et f_n n'est plus paire.

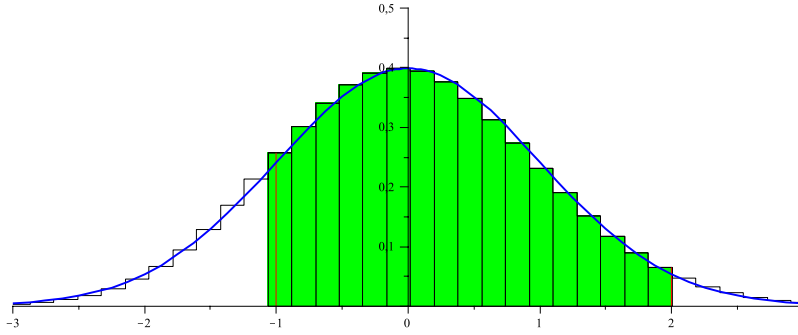


FIGURE 8 – Représentation graphique de f_n et de la fonction $t \rightarrow \frac{e^{-t^2/2}}{\sqrt{2\pi}}$

De manière analogue au cas $p = 1/2$, on montre que $\mathbb{P}(a \leq Z_n \leq b) = \int_{a_n}^{b_n} f_n(t) dt$ avec $|a_n - a| \leq h_n/2$ et $|b_n - b| \leq h_n/2$. Le résultat suivant généralise la proposition 5. Sa démonstration est d'un niveau plus élevé et nous ne la détaillerons pas complètement.

Proposition 13. *Il existe une constante $C > 0$ tel que pour tout $t \in [-A, A]$,*

$$\left| f_n(t) - \frac{e^{-t^2}}{\sqrt{2\pi}} \right| \leq \frac{C}{\sqrt{n}}$$

Preuve de la proposition 13. Rappelons la formule de Stirling : $n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} (1 + \varepsilon_n)$ avec $\lim_{n \rightarrow \infty} \varepsilon_n = 0$. Plus précisément, pour tout n , $0 \leq \varepsilon_n \leq 1/n$. Pour n et t fixés, il existe k tel que $t \in [t_k, t_{k+1}[$, d'où, en appliquant la formule de Stirling trois fois,

$$f_n(t) = \frac{p_{n,k}}{h_n} = \binom{n}{k} p^k q^{n-k} \sqrt{pqn} = \frac{n!}{k!(n-k)!} p^k q^{n-k} \sqrt{pqn} = M_n R_n (1 + \varepsilon_{n,k})$$

$$\text{avec } M_n = \frac{p^k q^{n-k}}{\binom{k}{n}^k \left(1 - \frac{k}{n}\right)^{n-k}}, \quad R_n = \frac{\sqrt{pq}}{\sqrt{2\pi \frac{k}{n} \frac{(n-k)}{n}}} \quad \text{et} \quad 1 + \varepsilon_{n,k} = \frac{(1 + \varepsilon_n)}{(1 + \varepsilon_k)(1 + \varepsilon_{n-k})}$$

Il existe une constante S telle que pour tout n , $|\varepsilon_{n,k}| \leq S/n$. En effet, $0 \leq \varepsilon_n \leq 1/n$, et, si n est assez grand, pour tout $t \in [-A, A]$, $n/4 \leq k \leq 3n/4$ et donc $0 \leq \varepsilon_k \leq 4/n$ et $0 \leq \varepsilon_{n-k} \leq 4/n$.

Etude de M_n . Notons que $k = np + \frac{t_k}{h_n} = np(1 + u_k)$ et $n - k = nq(1 + v_k)$, avec $u_k = \sqrt{\frac{q}{np}}t_k$ et $v_k = -\sqrt{\frac{p}{nq}}t_k$, et donc

$$M_n = \exp \left\{ -np\varphi(u_k) - nq\varphi(v_k) \right\}$$

où $\varphi(u) = (1+u) \ln(1+u)$. On peut montrer que $\varphi(u) = u + u^2/2 + \beta(u)$ avec $|\beta(u)| \leq |u|^3/3$ si $u \in [-1/2, 1/2]$. Par conséquent,

$$\begin{aligned} np\varphi(u_k) + nq\varphi(v_k) &= n(pu_k + pu_k^2/2 + p\beta(u_k) + qv_k + qv_k^2/2 + q\beta(v_k)) \\ &= t_k^2/2 + n(p\beta(u_k) + q\beta(v_k)) \end{aligned}$$

Comme les t_k sont dans $[-A, A]$ il existe $K > 0$ tel que $|u_k|, |v_k| \leq K/\sqrt{n} \leq 1/2$ si n est assez grand, et donc $|p\beta(u_k) + q\beta(v_k)| \leq K^3/n^{3/2}$. Par suite,

$$M_n = e^{-t_k^2/2}(1 + \varepsilon'_{n,k}) \quad \text{et il existe une constante } M > 0 \text{ tel que } \varepsilon'_{n,k} \leq M/\sqrt{n}$$

Etude de R_n . Comme $R_n = \frac{1}{\sqrt{2\pi(1+u_k)(1+v_k)}}$, on a

$$R_n = \frac{1}{\sqrt{2\pi}}(1 + \varepsilon''_{n,k}) \quad \text{et il existe une constante } R > 0 \text{ tel que } \varepsilon''_{n,k} \leq R/\sqrt{n}.$$

En conclusion, $f_n(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}(1 + \delta_{n,k})$ et il existe $C_0 > 0$ tel que $\delta_{n,k} \leq C_0/\sqrt{n}$. Comme

$|t - t_k| \leq h_n$, on en déduit qu'il existe une constante $C > 0$ tel que $\left| f_n(t) - \frac{e^{-t^2/2}}{\sqrt{2\pi}} \right| \leq \frac{C}{\sqrt{n}}$.

Remarque 14. Lorsque $p = 1/2$, il existe $L > 0$ tel que $\left| \frac{f_n(t_k)}{f_n(0)} - e^{-t_k^2/2} \right| \leq \frac{L}{n}$ pour tout n et pour tout $k \leq n$. Des simulations numériques indiquent qu'on peut prendre $L = 0,1$ (voir figure 7). L'ordre de grandeur de $f_n(t_k) - e^{-t_k^2/2}/\sqrt{2\pi}$ est donc en $1/n$, tandis que l'ordre de grandeur de $f_n(t) - e^{-t^2/2}/\sqrt{2\pi}$ est en général seulement en $1/\sqrt{n}$. Ceci confirme l'impression visuelle que la fonction f_n est très proche de la fonction $\frac{e^{-t^2/2}}{\sqrt{2\pi}}$ aux points t_k .

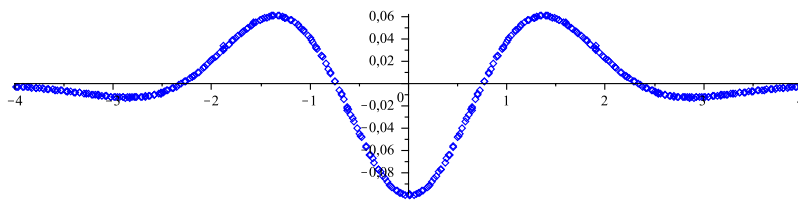


FIGURE 9 – Suites $(t_k, n(f_n(t_k) - \frac{e^{-t_k^2/2}}{\sqrt{2\pi}}))_{0 \leq k \leq n}$ quand $p = 1/2$ pour diverses valeurs de n

Par contre, lorsque $p \neq 1/2$, la vitesse de convergence est en $1/\sqrt{n}$. Plus précisément les points $(t_k, \frac{\sqrt{np(1-p)}}{p-1/2}(f_n(t_k) - \frac{e^{-t_k^2/2}}{\sqrt{2\pi}}))$ pour n assez grand se trouvent tous proches d'une même courbe comme indiquée sur la figure 8.

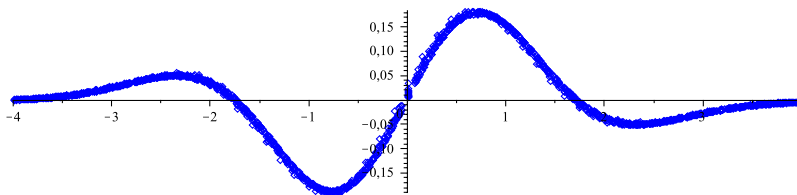


FIGURE 10 – Suites $(t_k, \frac{\sqrt{np(1-p)}}{p-1/2} (f_n(t_k) - \frac{e^{-t_k^2/2}}{\sqrt{2\pi}}))_{0 \leq k \leq n}$ quand $p \neq 1/2$
pour diverses valeurs de n et p .

6) La loi des erreurs.

Une variable X_n de loi binomiale représente le nombre de succès lors d'une suite de n épreuves indépendantes avec probabilité de succès p à chaque épreuve. Pour $1 \leq k \leq n$, on note Y_k la variable aléatoire égale à 1 si on a un succès à la k -ème épreuve et 0 sinon. On a donc $X_n = \sum_{k=1}^n Y_k$. D'autre part, l'espérance commune aux Y_k est $m = p$, la variance de Y_k est $\mathbb{E}(Y_k^2) - \mathbb{E}(Y_k)^2 = m - m^2 = p(1-p)$ et l'écart-type de Y_k est donc $\sigma = \sqrt{p(1-p)}$. Rappelons le théorème de Moivre-Laplace

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(a \leq \frac{\sum_{k=1}^n Y_k - nm}{\sigma\sqrt{n}} \leq b\right) = \int_a^b e^{-t^2/2} \frac{dt}{\sqrt{2\pi}} \quad (3)$$

Le théorème limite central énonce que (3) reste vrai si on suppose que les variables Y_k sont indépendantes, de même loi, d'espérance m et d'écart-type σ fini (on ne définira pas l'indépendance de variables aléatoires et on s'appuiera sur l'intuition). On peut donc supposer que les Y_k suivent une loi uniforme sur $[0, 1]$, ou qu'elles suivent une loi exponentielle. Dans ce cas bien entendu, $X_n = \sum_{k=1}^n Y_k$ ne suit plus une loi binomiale, mais d'après la relation (3), les intervalles de fluctuation asymptotiques restent les mêmes. La figure

suivante représente les densités f_n des variables $Z_n = \frac{\sum_{k=1}^n Y_k - nm}{\sigma\sqrt{n}}$ lorsque les variables Y_k suivent une loi exponentielle et que $n = 4, 16, 35, 64$ et 100.

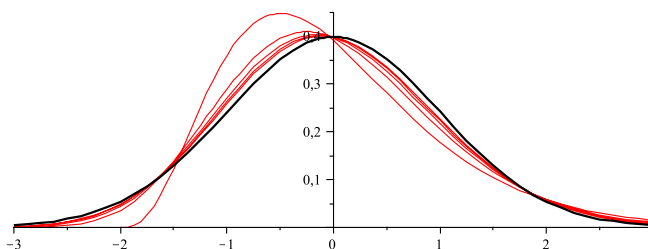


FIGURE 11 – Comparaison des graphes de f_n et de la fonction $x \rightarrow \exp(-x^2/2)/\sqrt{2\pi}$.

Le théorème limite central exprime le fait qu'une somme X_n d'un grand nombre de variables indépendantes, de même loi, et de variance finie a une distribution à peu près gaussienne. La loi des erreurs généralise ce fait là à une somme de petites variables indépendantes dont aucune n'est prépondérante. C'est en raison de cette universalité que les variables aléatoires intervenant en modélisation sont souvent supposées suivre des lois normales. Nous allons illustrer ce fait par l'exemple du tireur à la carabine.

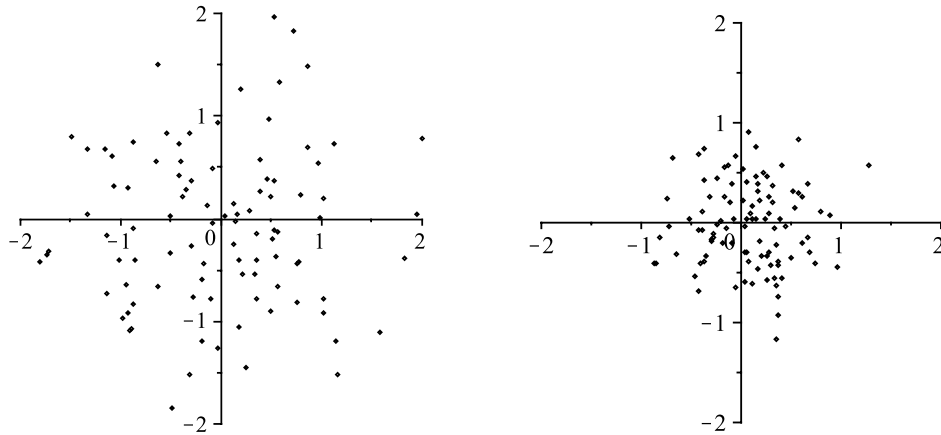


FIGURE 12 – 100 réalisations de (X, Y) où X, Y sont des gaussiennes centrées d'écart-type 1, puis d'écart-type 0, 4.

Observez par exemple la figure ci-dessus. Elle est obtenue en effectuant n réalisations de couples (X_k, Y_k) ($1 \leq k \leq n$) de variables gaussiennes indépendantes centrées. Elle correspond bien aux résultats obtenus à un tir à la carabine sur une cible. Expliquons pourquoi. Si (X_k, Y_k) sont les coordonnées du k -ième tir, X_k et Y_k sont des sommes de petites variables indépendantes (erreur de visée, tremblement, défaut de concentration, recul, perturbation par un élément extérieur comme un contrejour ou un cri,...). D'après la loi des erreurs, il est donc raisonnable de supposer que (X_k, Y_k) est un couple de variables gaussiennes. De plus, l'indépendance des tirs est exprimée par le fait que les couples (X_k, Y_k) sont indépendants entre eux. Enfin, l'écart-type de ces gaussienne mesure l'adresse du tireur : plus il est petit, et plus le tireur est adroit.