32 - Densité d'une somme de lois uniformes et indépendantes

www.daniel-saada.eu

Le but de cet article est d'établir, par deux méthodes, que la densité d'une somme de n lois uniformes et indépendantes sur $[0,1]$ est $f_n(x)=\frac{1}{(n-1)!}\sum_{k=0}^{\mathrm{int}(x)}(-1)^k C_n^k (x-k)^{n-1}$ quand x décrit l'intervalle $[0,n]$, $\mathrm{int}(x)$ désignant la partie entière de x et bien sûr $C_n^k=\binom{n}{k}=\frac{n!}{k!(n-k)!}$, les fonctions f_n faisant l'objet d'une étude approfondie.

Mes remerciements les plus vifs vont à Jean-Christophe Feauveau, Claude Morin et Alain Rémondière, qui m'ont beaucoup aidé dans l'élaboration de ce texte.

1) Transformée de Fourier et densité

Rappelons que si une variable aléatoire réelle Z a une densité f, $E[e^{itZ}]=\int_{\mathbb{R}} e^{itx} f(x)dx$ pour tout réel t. Si de plus $t \to E[e^{itZ}]$ est intégrable sur $\mathbb{R}$, alors $\frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}E[e^{itZ}]dt = f(x)$ presque partout (transformée de Fourier inverse). En d'autres termes, $x \to \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}E[e^{itZ}]dt$ est une densité de Z (si f est continue, $f(x)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}E[e^{itZ}]dt$ partout).

Soit $S_n = X_1+...+X_n$, où les X_i sont *mutuellement* indépendantes et de *même* loi : comme $e^{itS_n}=\prod_{k=1}^n e^{itX_k}$, l'indépendance des X_i donne $E[e^{itS_n}]=\left(E[e^{itX_1}]\right)^n$. Une densité g_n de S_n est donc $g_n(x)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\left(E[e^{itX_1}]\right)^n dt$ <u>si</u> $t \to \left(E[e^{itX_1}]\right)^n$ <u>est intégrable sur</u> $\mathbb{R}$.

On suppose les X_i uniformes et indépendantes sur $[-1,1]$: la transformée de Fourier de chaque X_i est $t \to \frac{\sin t}{t}$ comme il est facile de le vérifier. Pour $n \geq 2$, $\left(\frac{\sin t}{t}\right)^n$ est intégrable (au sens de Lebesgue), et donc une densité de $S_n=\sum_1^n X_i$ est

$$g_n(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin t}{t}\right)^n e^{-ixt}dt=\frac{1}{\pi}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dt.$$

On retrouve que g_n est paire (comme il se devait) et on découvre que g_n est au moins C^{n-2} (par dérivation sous le signe intégral).

Comme g_n est une densité, $\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dt$ est positive pour tout réel x.

La variable aléatoire S_n prenant ses valeurs dans $[-n,n]$,

$$\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dt$$ est nulle en dehors de $[-n,n]$.

De plus, $\int_0^{+\infty}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dtdx = \int_0^{n}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dtdx = \frac{\pi}{2}$ car l'intégrale d'une densité vaut 1.

2) Calcul de $I_n(x)=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dt$ **pour** $n \geq 2$

En intégrant par parties, $I_n(x)=\frac{1}{n-1}\int_0^{+\infty}\frac{d}{dt}\left(\sin^n(t)\cos(xt)\right)\times\frac{1}{t^{n-1}}dt$.

En intégrant $(n-1)$ fois par parties, $I_n(x)=\frac{1}{(n-1)!}\int_0^{+\infty}\left(\sin^n(t)\cos(xt)\right)^{(n-1)}\frac{dt}{t}$.

Les formules d'Euler donnent $2^{n+1}i^n(\sin t)^n\cos(xt)=\sum_{k=0}^{n}(-1)^k C_n^k(e^{it(n-2k+x)}+e^{it(n-2k-x)})$, d'où

$$2^{n+1}i\left(\sin^n(t)\cos(xt)\right)^{(n-1)}=\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}e^{it(n-2k+x)}+\sum_{k=0}^{n}(-1)^k C_n^k(n-2k-x)^{n-1}e^{it(n-2k-x)}.$$

En changeant k en $n-k$ dans le second sigma

$$2^{n+1}i\left(\sin^n(t)\cos(xt)\right)^{(n-1)}=\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}e^{it(n-2k+x)}-\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}e^{-it(n-2k+x)}$$

et $\left(\sin^n(t)\cos(xt)\right)^{(n-1)}=\frac{1}{2^n}\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}\sin\left((n-2k+x)t\right)$.

Avec le résultat classique $\int_0^{+\infty}\frac{\sin(at)}{t}dt=\frac{\pi}{2}signe(a)$, il vient

$$I_n(x)=\frac{\pi}{(n-1)!2^{n+1}}\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}signe\left(n-2k+x\right).$$

La dérivée d'ordre $(n-1)$ de $t\to(\sin t)^n e^{itx}$ s'annule en $t=0$ pour tout réel x. Or

$$2^n i^n(\sin t)^n e^{itx}=\sum_{k=0}^{n}(-1)^k C_n^k e^{it(n-2k+x)},\quad 2^n i\left(\sin^n(t)e^{ixt}\right)^{(n-1)}=\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}e^{it(n-2k+x)}$$

d'où $\sum_{k=0}^{n}(-1)^k C_n^k(n-2k+x)^{n-1}=0$ pour tout x.

Quand $|x|>n$, $n-2k+x$ est de signe constant, et alors $I_n(x)=0$ en vertu de ce qui précède.

Pour $-n<x<n$:

$$I_n(x)=\frac{\pi}{(n-1)!2^{n+1}}\left(\sum_{k=0}^{(n+x)/2}(-1)^k C_n^k(n-2k+x)^{n-1}-\sum_{k>(n+x)/2}(-1)^k C_n^k(n-2k+x)^{n-1}\right)$$

et comme la somme des deux sigmas est nulle, $I_n(x)=\frac{\pi}{(n-1)!2^n}\sum_{k=0}^{(n+x)/2}(-1)^k C_n^k(n-2k+x)^{n-1}$.

En particulier, $\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n dt=I_n(0)=\frac{\pi}{(n-1)!2^n}\sum_{k=0}^{[n/2]}(-1)^k C_n^k(n-2k)^{n-1}$, ce qui donne

$$I_1=\pi/2,\ I_2=\pi/2,\ I_3=3\pi/8, I_4=\pi/3, I_5=115\pi/384$$

ce qu'on peut vérifier avec un logiciel de calcul formel (par exemple WolframAlpha, gratuit et en ligne).

3) Densité d'une somme de lois uniformes indépendantes à valeurs dans $[0,1]$

De **1)** et **2)** on déduit la densité g_n de $S_n = \sum_1^n X_i$ quand les X_i sont dans $[-1,1]$:

$$g_n(x) = \frac{1}{(n-1)!2^n} \sum_{k=0}^{[(n+x)/2]} (-1)^k C_n^k (x+n-2k)^{n-1} \text{ sur }]-n,n[\text{ et } 0 \text{ ailleurs.}$$

Remarque 1. Pour qui veut étudier numériquement g_n, il est préférable d'utiliser

$$g_n(x) = \frac{1}{(n-1)!2^{n+1}} \sum_{k=0}^{n} (-1)^k C_n^k (n-2k+x)^{n-1} signe(n-2k+x).$$

Les $Y_i = (1+X_i)/2$ sont uniformes et indépendantes sur $[0,1]$ et

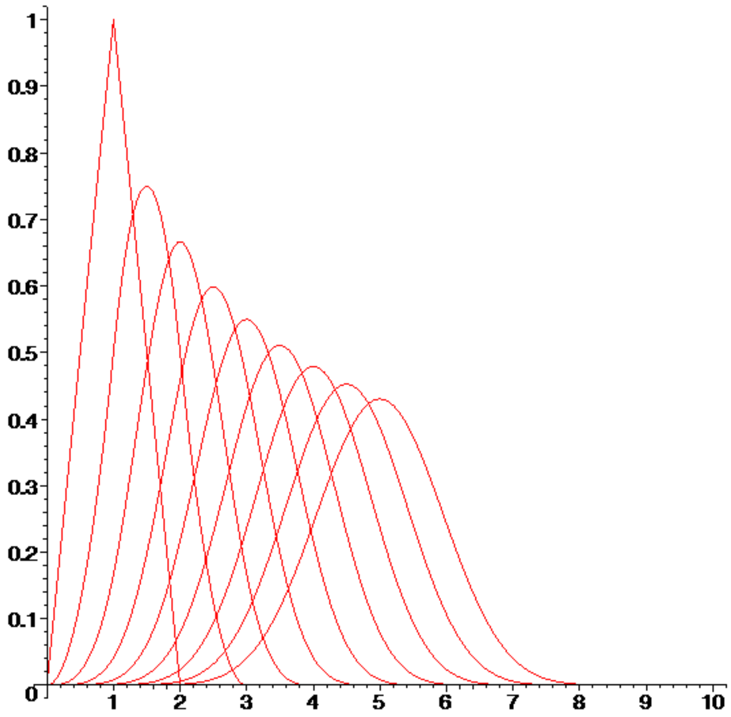
$$p\left(\sum_1^n Y_i \le x\right) = p\left(\sum_1^n X_i \le 2x-n\right).$$

En dérivant cette égalité par rapport à x, on trouve enfin la densité f_n de $\sum_1^n Y_i$ sur $[0,n]$:

$$f_n(x) = 2g_n(2x-n) = \frac{1}{(n-1)!} \sum_{k=0}^{\text{int}(x)} (-1)^k C_n^k (x-k)^{n-1}.$$

Ainsi, $f_n(x) = \dfrac{x^{n-1}}{(n-1)!}$ sur $[0,1[$; $f_n(x) = \dfrac{x^{n-1} - n(x-1)^{n-1}}{(n-1)!}$ sur $[1,2[$, etc...

On découvre que f_n est polynomiale par morceaux et donc C^∞ par morceaux sur R. On pourra vérifier que f_n est positive et d'intégrale 1 et s'assurer au moins de la continuité de f_n (dont on sait qu'elle est C^{n-2} comme g_n). Voici les graphes de f_2 à f_{10} :



La connaissance des densités permet de calculer par exemple $p(a \le S_n \le b) = \int_a^b g_n(t)dt$.

4) Une deuxième méthode de détermination de f_n

On sait que si deux variables aléatoires indépendantes X_1 et X_2 ont des densités f_1 et f_2, leur somme a pour densité $\varphi_2(x) = \int_{-\infty}^{+\infty} f_2(x-t)f_1(t)dt$. Si $X_1, X_2, ..., X_n$ sont mutuellement indépendantes, $X_1 + X_2 + ... + X_{n-1}$ est indépendante de X_n et la densité φ_n de $X_1 + X_2 + ... + X_n$ est $\varphi_n(x) = \int_{-\infty}^{+\infty} \varphi_{n-1}(x-t)f_1(t)dt$.

Quand $X_1, X_2, ..., X_n$ sont uniformes et (mutuellement) indépendantes sur $[0,1]$, les densités f_n de $\sum_1^n X_i$ sont régies par $f_{n+1}(x) = \int_{-\infty}^{+\infty} f_n(x-t)f_1(t)dt = \int_0^1 f_n(x-t)dt = \int_{x-1}^{x} f_n(t)dt$.

Il s'agit donc d'expliciter les fonctions f_n quand $f_1 = 1_{]0,1[}$.

On établit d'abord facilement par récurrence que :

a) le support de f_n est dans $]0,n[$

b) f_n est C^{n-2} sur R (pour $n \ge 2$) et C^∞ par morceaux

c) $f_n^{(k)}(x) = \sum_{i=0}^{k} (-1)^i C_k^i f_{n-k}(x-i)$ si $k \le n-2$ car la dérivation commute avec l'opérateur aux différences finies ; de même, $f_n^{(n-1)}(x) = \sum_{k=0}^{n-1} (-1)^k C_{n-1}^k f_1(x-k)$ mais pour x non entier

d) $n \ge 2$ et $k \le n-2$ impliquent $f_n^{(k)}(0) = 0$.

Sur l'intervalle $[0,1[$, la détermination de f_n est simple puisque $f_{n+1}(x) = \int_0^x f_n(t)dt$: comme

$f_1(x) = 1$ et $f_2(x) = x$, $f_n(x) = \dfrac{x^{n-1}}{(n-1)!}$.

Puisque f_n est C^{n-2} sur R (pour $n \ge 2$) et C^∞ par morceaux, alors f_n est C^{n-2} sur R et C^{n-1} par morceaux. La formule de Taylor avec reste intégral (un peu modifiée) donne alors, en vertu de **d)** :

$$f_n(x) = \sum_{k=0}^{n-2} \frac{f_n^{(k)}(0)}{k!} x^k + \int_0^x \frac{(x-t)^{n-2}}{(n-2)!} f_n^{(n-1)}(t)dt = \int_0^x \frac{(x-t)^{n-2}}{(n-2)!} f_n^{(n-1)}(t)dt .$$

On en déduit $f_n(x) = \dfrac{1}{(n-2)!} \sum_{k=0}^{n-1} (-1)^k C_{n-1}^k \int_0^x (x-t)^{n-2} f_1(t-k)dt$.

Fixons x dans $[1,n]$ et soit j la partie entière de x . Posons $u_k(x) = \int_0^x (x-t)^{n-2} f_1(t-k)dt$.

Comme $f_1(t-k) = 1$ si et seulement si $t \in]k,k+1[$, on va distinguer trois cas :

- $k \ge j+1$,alors $u_k(x) = 0$;
- $k = j$, $u_k(x) = \int_k^x (x-t)^{n-2} dt = \dfrac{(x-k)^{n-1}}{n-1}$;

- si $k \leq j-1, u_k(x) = \int_k^{k+1} (x-t)^{n-2} dt = \dfrac{(x-k)^{n-1} - (x-k-1)^{n-1}}{n-1}$.

On a donc $f_n(x) = \dfrac{1}{(n-1)!} \sum_{k=0}^{j-1} (-1)^k C_{n-1}^k \left[(x-k)^{n-1} - (x-k-1)^{n-1} \right] + \dfrac{(-1)^j}{(n-1)!} C_{n-1}^j (x-j)^{n-1}$.

$$
\begin{aligned}
(n-1)!\, f_n(x) &= \sum_{k=0}^{j-1} (-1)^k C_{n-1}^k (x-k)^{n-1} - \sum_{k=1}^{j} (-1)^{k-1} C_{n-1}^{k-1} (x-k)^{n-1} + (-1)^j C_{n-1}^j (x-j)^{n-1} \\
&= x^{n-1} + \sum_{k=1}^{j-1} (-1)^k C_n^k (x-k)^{n-1} + (-1)^j C_{n-1}^{j-1} (x-j)^{n-1} + (-1)^j C_{n-1}^j (x-j)^{n-1} \\
&= x^{n-1} + \sum_{k=1}^{j-1} (-1)^k C_n^k (x-k)^{n-1} + (-1)^j C_n^j (x-j)^{n-1} \\
&= \sum_{k=0}^{\mathrm{int}(x)} (-1)^k C_n^k (x-k)^{n-1}
\end{aligned}
$$

ce qui achève la démonstration.

Remarque 2. Les polynômes de degré $(n-1)$ qui composent f_n la rende C^{n-2} sur R. On peut se demander si ce sont les seuls. De façon précise, si E désigne l'espace vectoriel des fonctions C^{n-2} sur R qui sont nulles en dehors de $[0,n]$ et dont les restrictions à chaque intervalle $[k,k+1]$ sont des polynômes de degré $\leq (n-1)$, alors la dimension de E est 1 et f_n en est une base.

5) Étude des fonctions g_n

a) Variations de g_n

Nous avons vu que $f_{n+1}(x) = \int_{x-1}^{x} f_n(t)dt$, d'où $f_{n+1}'(x) = f_n(x) - f_n(x-1)$; pour g_n, on a

$$g_{n+1}'(x) = \frac{1}{2}\left(g_n(x+1) - g_n(x-1)\right).$$

Une récurrence immédiate prouve que si g_n croît sur R^-, alors g_{n+1} croît sur $]-\infty,-1]$. La parité de g_n entraîne que si $x \in [-1,0], 2g_{n+1}'(x) = g_n(-x-1) - g_n(x-1) \geq 0$ car $x-1 \leq -x-1 \leq 0$.

Comme $g_2(x) = (x+2)/4$ sur $[-2,0]$, on peut conclure :

g_n croît sur $[-n,0]$ et décroît sur $[0,n]$, f_n croît sur $[0,n/2]$ et décroît sur $[n/2,n]$.

Remarque 3. Comme $g_n'(x) = -\int_0^{\infty} \left(\dfrac{\sin t}{t}\right)^n t\sin(xt)dt$, l'intégrale $\int_0^{\infty} \left(\dfrac{\sin t}{t}\right)^n t\sin(xt)dt$ est donc du signe de x.

b) Monotonie du maximum de g_n

Le maximum de g_n est obtenu en 0 et $g_n(0) = \dfrac{1}{\pi}\int_0^{\infty} \left(\dfrac{\sin t}{t}\right)^n dt$. Nous prouvons que cette suite est décroissante ; pour ce faire, posons $I_n = \int_0^{\infty} \left(\dfrac{\sin t}{t}\right)^n dt$ et prouvons que

$$I_n - I_{n+1} = \int_0^\infty \left(\frac{\sin t}{t}\right)^n \left(1-\frac{\sin t}{t}\right) dt \geq 0 \text{ pour tout } n.$$

C'est évident pour n pair car $(t-\sin t)/t \geq 0$ quand t est positif.

Pour n impair, $I_n - I_{n+1} = \int_0^\pi \left(\frac{\sin t}{t}\right)^n \left(1-\frac{\sin t}{t}\right) dt + \int_\pi^\infty \left(\frac{\sin t}{t}\right)^n \left(1-\frac{\sin t}{t}\right) dt = A_n + B_n$

On majore B_n en écrivant $B_n \leq |B_n| \leq \int_\pi^\infty \frac{2}{t^n} dt = \frac{2}{(n-1)\pi^{n-1}}$; on minore A_n par[1]

$$A_n \geq \int_{\pi/4}^{\pi/2} \left(\frac{\sin t}{t}\right)^n \left(1-\frac{\sin t}{t}\right) dt \geq \left(1-\frac{\sin \pi/4}{\pi/4}\right) \int_{\pi/4}^{\pi/2} \left(\frac{\sin t}{t}\right)^n \geq \left(1-\frac{\sqrt{8}}{\pi}\right) \int_{\pi/4}^{\pi/2} \left(\frac{2}{\pi}\right)^n dt = \frac{\pi-\sqrt{8}}{4}\left(\frac{2}{\pi}\right)^n.$$

On en déduit $I_n - I_{n+1} = A_n + B_n \geq A_n - |B_n| \geq \frac{\pi-\sqrt{8}}{4}\left(\frac{2}{\pi}\right)^n - \frac{2}{(n-1)\pi^{n-1}}$

on trouve $I_n - I_{n+1} > 0$ à partir de $n=5$; mais $I_4 \geq I_5$ et $I_2 \geq I_3$ et on finit avec

$$I_1 = I_2 = \pi/2 \geq I_3 = 3\pi/8 \geq I_4 = \pi/3.$$

c) Limite et équivalent du maximum de g_n

D'abord $I_n = \int_0^\infty \left(\frac{\sin t}{t}\right)^n dt$ est équivalente à $J_n = \int_0^\pi \left(\frac{\sin t}{t}\right)^n dt$ car $K_n = \int_\pi^{+\infty} \left(\frac{\sin t}{t}\right)^n dt$ est négligeable devant J_n. En effet, $|K_n| \leq \int_\pi^{+\infty} \frac{dt}{t^n} = \frac{1}{(n-1)\pi^{n-1}}$, tandis que

$$J_n \geq \int_0^{\pi/2} \left(\frac{\sin t}{t}\right)^n dt \geq \frac{\pi}{2}\left(\frac{2}{\pi}\right)^n = \left(\frac{2}{\pi}\right)^{n-1}.$$

Un équivalent de J_n est donné par le lemme suivant bien connu (**Méthode de Laplace**).

Lemme. Soit f continue et positive sur $[0,\mu]$. On suppose $f(0)=1$, $t \in]0,\mu]$ impliquant $f(t) \in [0,1[$. Alors $J_n = \int_0^\mu f^n$ tend vers 0 ; si de plus, $1-f(t) \sim at^b$ en $t=0$, avec a et b positifs, $J_n \sim \frac{\Gamma(1+1/b)}{(na)^{1/b}}$, où Γ est la fontion gamma.

Démonstration du Lemme.

On va utiliser $\left(1-\frac{at^b}{n}\right)^n = \left(1-a\left(\frac{t}{n^{1/b}}\right)^b\right)^n \to e^{-at^b}$, aussi allons-nous faire le changement de variable $x = n^{1/b} t$: $n^{1/b} J_n = \int_0^{n^{1/b}\mu} f^n\left(\frac{x}{n^{1/b}}\right) dx$, avec $n^{1/b}\mu \to +\infty$.

Pour $x \geq 0$, il existe n au delà duquel $x \leq n^{1/b}\mu$, d'où

$\ln\left(f^n\left(\frac{x}{n^{1/b}}\right)\right) = n \ln\left(f\left(\frac{x}{n^{1/b}}\right)\right) \sim n\left(f\left(\frac{x}{n^{1/b}}\right)-1\right) \sim -ax^b$ et $f^n\left(\frac{x}{n^{1/b}}\right)$ a pour limite e^{-ax^b}.

Pour utiliser le théorème de convergence dominée, il reste à prouver que $f^n\left(\frac{x}{n^{1/b}}\right)$ est majorée par

---

[1] $t \mapsto \sin t / t$ décroît sur $[0,\pi/2]$.

une fonction intégrable sur R^+. La fonction $\dfrac{1-f(t)}{t^b}$ vaut a en $t=0$ et est >0 sur]0,µ], elle a un minimum $m>0$. Aussi,

$$f(t)\le 1-mt^b \text{ et donc } f^n\left(\frac{x}{n^{1/b}}\right)\le\left(1-m\frac{x^b}{n}\right)^n\le e^{-mx^b} \text{ (car } \ln(1-x)\le -x \text{ sur } [0,1[\text{).}$$

Comme $x\to e^{-mx^b}$ est intégrable sur R^+, $n^{1/b}J_n$ a pour limite $\int_0^{+\infty}e^{-ax^b}dx$. Or, avec le changement de variables $y=ax^b$, $\int_0^{+\infty}e^{-ax^b}dx=\dfrac{1}{ba^{1/b}}\int_0^{+\infty}y^{1/b-1}e^{-y}dy=\dfrac{1}{a^{1/b}}\dfrac{1}{b}\Gamma(1/b)=\dfrac{\Gamma(1+1/b)}{a^{1/b}}$.

On a donc bien $J_n=\int_0^{\mu}f^n\sim\dfrac{\Gamma(1+1/b)}{(na)^{1/b}}$.

La fonction $f(t)=\dfrac{\sin t}{t}$ satisfait aux conditions du lemme quand $\mu=\pi$, avec $b=2$, $a=1/6$, $m=\left(\dfrac{2}{\pi}\right)^2-\left(\dfrac{2}{\pi}\right)^3\simeq 0{,}15$. Comme $\Gamma(3/2)=\sqrt{\pi}/2$, on obtient $I_n\sim\sqrt{\dfrac{3\pi}{2n}}$ et donc

$$g_n(0)\sim\sqrt{\frac{3}{2\pi n}}.$$

Remarque 4 . $I_{10}=\dfrac{15619}{72576}\pi\approx 0{,}676$ et l'équivalent $\sqrt{\dfrac{3\pi}{2n}}$ vaut $0{,}686$ pour $n=10$.

Complément. Plus généralement, $g_n(x\sqrt{n})$ est équivalente à $\sqrt{\dfrac{3}{2\pi n}}e^{-3x^2/2}$; cet équivalent sera mis dans un contexte probabiliste en **6)**.

d) Points d'inflexion de g_n

Nous démontrons par récurrence qu'à partir du rang 4, g_n possède un unique point d'inflexion d'abscisse x_n, avec $1<x_n<n$, $g_n''<0$ sur $]0,x_n[$ et $g_n''>0$ sur $]x_n,n[$.

L' hypothèse est vérifiée pour $n=4$, avec $x_4=4/3$, puisque :

$$96g_4(x)=3x^3-12x^2+32 \text{ sur } [0,2] \text{ et } 96g_4(x)=(4-x)^3 \text{ sur } [2,4].$$

Supposons-la vérifiée à un rang $n\ge 4$.

Par hypothèse de récurrence, g_n' décroît strictement sur $[0,x_n]$ et croît strictement sur $[x_n,n]$ (en restant négative). Par continuité et stricte monotonie de g_n', il existe un unique $y<0$ tel que la droite d'ordonnée y coupe le graphe de g_n' sur $]0,n[$ en deux points d'abscisses a et b distantes de 2. Pour $c=(a+b)/2$, $2g_{n+1}''(c)=g_n'(c+1)-g_n'(c-1)=g_n'(b)-g_n'(a)=0$.

Comme $a=c-1>0$, $c>1$, et comme $b=c+1<n$, $c<n-1$. Supposons qu'il existe d sur $]1,n-1[$ tel que $g_{n+1}''(d)=0$: alors $g_n'(d+1)=g_n'(d-1)$, $d=c$, c est donc l'unique zéro de g_{n+1}'' sur $]1,n-1[$.

La fonction continue g''_{n+1} garde donc un signe constant sur chacun des intervalles $]1,c[$ et $]c,n-1[$.
Or, sur $]0,1]$, $2g''_{n+1}(x)=g'_n(x+1)-g'_n(x-1)=g'_n(x+1)+g'_n(1-x)\leq g'_n(x+1)<0$, et sur $[n-1,n+1[$, $2g''_{n+1}(x)=-g'_n(x-1)>0$. Il en résulte que $g''_{n+1}<0$ sur $]0,c[$ et $g''_n>0$ sur $]c,n+1[$, c pouvant désormais se noter x_{n+1} avec $1<x_{n+1}<n+1$.

La démonstration est achevée.

Voilà les premières valeurs de x_n :

n	4	5	6	7	8	9	10	20	50	100	200
x_n	1 ,333	1 ,382	1,519	1,619	1,718	1,812	1,901	2,635	4,115	5,797	8,181

Ces valeurs numériques indiquent que si x_n est équivalent à $C.n^a$, alors a vaut probablement (!) $1/2$. En effet, 2^a est la limite des quotients x_{2n}/x_n, et $x_{2n}/x_n=1.386,1.409,1.411$ pour $n=10,50,100$.

6) Approximation par loi normale et équivalent de x_n.

Dans ce paragraphe, on s'intéresse à la variable aléatoire $T_n=S_n/\sqrt{n}$, où, rappelons-le, $S_n=\sum_1^n X_i$, les X_i étant uniformes dans $[-1,1]$ et indépendantes.

On sait, nous allons le redémontrer, que $p\left(T_n\leq x\right)\to\dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}e^{-u^2/2\sigma^2}du$, où $\sigma=1/\sqrt{3}$ est l'écart-type commun des X_i ; de plus, la convergence est *uniforme* sur $\mathbb{R}$ en vertu d'un théorème de Dini. Appelons N la loi limite normale $N(0,1/\sqrt{3})$ et h sa densité : $h(x)=\sqrt{\dfrac{3}{2\pi}}e^{-3x^2/2}$.

Exemple. Pour $n=16$ et $x=1$, $p(S_{16}\leq 4)=0.958365$ et $p(N\leq 1)\approx 0,958368$: l'approximation est époustouflante !

La densité h_n de T_n s'obtient en dérivant sa fonction de répartition :

$$p\left(T_n\leq x\right)=p\left(S_n\leq x\sqrt{n}\right)=\int_{-\infty}^{x\sqrt{n}}g_n(t)dt\ ,\ \text{et donc}\ h_n(x)=\sqrt{n}g_n(x\sqrt{n}).$$

D'autre part, $E[e^{itT_n}]=\left(\dfrac{\sin(t/\sqrt{n})}{t/\sqrt{n}}\right)^n$ car $E[e^{itT_n}]=\left(E[e^{itX_1/\sqrt{n}}]\right)^n$, et on sait qu'alors

$$h_n(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}E[e^{itT_n}]dt=\frac{1}{\pi}\int_0^{+\infty}\cos(tx)E[e^{itT_n}]dt.$$

a) La suite des espérances $E[e^{itT_n}]$ converge vers $E[e^{itN}]$

En effet, d'une part $\left(\dfrac{\sin(t/\sqrt{n}}{t/\sqrt{n}}\right)^n \to e^{-t^2/6}$, d'autre part $E[e^{itN}] = \int_{\mathbb{R}} e^{itx} h(x)dx = e^{-t^2/6}$.

b) Les fonctions $t \to E[e^{itT_n}]$ sont majorées en valeur absolue par une fonction intégrable sur R^+.

Nous poserons $\left(\dfrac{\sin(t/\sqrt{n})}{t/\sqrt{n}}\right)^n = s_n(t)$:

- sur $[0, \pi\sqrt{n}/2]$, $0 < s_n(t) \le e^{-mt^2}$, avec $m = min\left\{\dfrac{t-\sin t}{t^3} : 0 \le t \le \pi/2\right\}$; comme

$m = \left(\dfrac{2}{\pi}\right)^2 - \left(\dfrac{2}{\pi}\right)^3 \simeq 0{,}147$, $|s_n(t)| \le e^{-0{,}14t^2}$;

- si $t > \pi\sqrt{n}/2$ et $n \ge 5$, $|s_n(t)| \le \left(\dfrac{2}{\pi}\right)^{n-4} \dfrac{n^2}{t^4}$ car $\dfrac{\sin t}{t}$ décroit sur $\left[\dfrac{\pi}{2}, \pi\right]$ et au-delà $\left|\dfrac{\sin t}{t}\right| \le \dfrac{1}{\pi}$.

Soit K le maximum de $n^2\left(\dfrac{2}{\pi}\right)^{n-4}$: il est obtenu quand $n = 7$ et vaut $33{,}03$, que l'on arrondira à $K = 34$. Comme $\pi\sqrt{n}/2 > \pi$ pour $n \ge 5$,

$$|s_n(t)| \le e^{-0{,}14t^2} \text{ sur } [0,\pi] \text{ et } |s_n(t)| \le e^{-0{,}14t^2} + 34/t^4 \text{ sur } [\pi, +\infty[$$

la fonction majorante étant bien intégrable sur R^+.

c) h_n converge simplement vers h

Comme $h_n(x) = \dfrac{1}{\pi}\int_0^{+\infty} \cos(tx) s_n(t)\, dt$, le théorème de convergence dominée donne

$$h_n(x) \to h(x) = \frac{1}{\pi}\int_0^{+\infty} e^{-t^2/6} \cos(xt) dt = \sqrt{\frac{3}{2\pi}} e^{-3x^2/2}.$$

On en déduit $g_n(x\sqrt{n})$ équivalente à $\sqrt{\dfrac{3}{2\pi n}}\, e^{-3x^2/2}$ quand n tend vers l'infini.

d) T_n converge en loi vers N

Puisque $h_n \to h$, alors $\int_R |h_n - h| \to 0$. En effet, si $k_n = \min(h, h_n)$, $0 \le k_n \le h$ et $k_n \to h$, et alors $\int k_n \to \int h = 1$ .Or, $|h - h_n| = h + h_n - 2k_n$, donc $\int |h - h_n| \to 1 + 1 - 2 = 0$.

Soit B un borélien de R : $p(T_n \in B) = \int_B h_n \to p(N \in B) = \int_B h$ car

$$|p_n(B) - p(B)| \le \int_R |h_n - h| \to 0 .$$

En appliquant à $B =]-\infty, x]$, on a la définition de la convergence en loi et cette convergence est uniforme en x .

e) h_n **converge uniformément vers** h

$|h_n(x)-h(x)|\le \frac{1}{\pi}\int_0^{+\infty}|E[e^{itT_n}]-E[e^{itN}]|\,dt$ qui tend vers 0 d'après le théorème de convergence dominée.

f) h_n'' **converge uniformément vers** $h'':x\mapsto 3\sqrt{\frac{3}{2\pi}}e^{-3x^2/2}(3x^2-1)$

On dérive deux fois $h_n(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}E[e^{itT_n}]dt$ et $h(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}E[e^{itN}]dt$:

$$h_n''(x)=\frac{-1}{2\pi}\int_{-\infty}^{+\infty}t^2e^{-itx}E[e^{itT_n}]dt \text{ et } h''(x)=\frac{-1}{2\pi}\int_{-\infty}^{+\infty}t^2e^{-itx}E[e^{itN}]dt.$$

D'où $|h_n''(x)-h''(x)|\le\frac{1}{\pi}\int_0^{+\infty}t^2|E[e^{itT_n}]-E[e^{itN}]|\,dt$, qui tend vers 0 d'après le théorème déjà cité de Lebesgue.

g) Un équivalent de x_n

$h_n(x)=\sqrt{n}\,g_n(x\sqrt{n})$ donne $h_n''(x)=n\sqrt{n}\,g_n''(x\sqrt{n})$

Dans la convergence uniforme établie en **f)**, remplaçons x par $x_n/\sqrt{n}$: comme $g_n''(x_n)=0$,

$$e^{-3x_n^2/2n}(3x_n^2/n-1)\to 0.$$

On va conclure avec le signe de g_n'', étudié en **5 d)**. Fixons un réel $a>1/\sqrt{3}$: $n\sqrt{n}\,g_n''(a\sqrt{n})$ a une limite positive et donc, pour n assez grand, $g_n''(a\sqrt{n})>0$. Il en résulte $a\sqrt{n}>x_n$ et donc $e^{-3x_n^2/2n}>e^{-3a^2/2}$ pour n assez grand, ce qui induit $3x_n^2/n-1\to 0$ et donc $x_n\sim\sqrt{\frac{n}{3}}$.

Question : (x_n) **est-elle croissante ?**

7) Compléments d'Analyse

Sans l'apport probabiliste, peut-on prouver que $g_n:x\to\frac{1}{\pi}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\cos(xt)dt$ est positive sur R, nulle en dehors de $[-n,n]$ et d'intégrale 1 ? Voici nos réponses.

a) g_n **est positive sur** R **et nulle en dehors de** $[-n,n]$

On intègre par parties $\pi g_{n+1}(x)=\int_0^{\infty}(\sin t)^{n+1}\cos(xt)\times\frac{1}{t^{n+1}}dt$:

$$\pi g_{n+1}(x)=\frac{1}{n}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\big((n+1)\cos t\cos(xt)-x\sin(xt)\sin t\big)dt.$$

Comme $(n+1)\cos t\cos(xt)-x\sin(xt)\sin t=(n+1-x)\cos(xt-t)+(n+1+x)\cos(xt+x)$ il vient

$$2ng_{n+1}(x)=(n+1-x)g_n(x-1)+(n+1+x)g_n(x+1)$$

et on valide par récurrence l'hypothèse « g_n est positive et g_n est nulle en dehors de $[-n,n]$ ».

En particulier, $\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\cos(nt)dt=0$.

b) $\int_0^{+\infty}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dtdx=\int_0^{n}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dtdx=\frac{\pi}{2}$.

On a le droit d'écrire

$$\int_0^{n}\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n \cos(xt)dtdx=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\left(\int_0^n \cos(xt)\,dx\right)dt=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\frac{\sin(nt)}{t}dt\ .$$

Posons $u_n=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\frac{\sin(nt)}{t}dt$ et intégrons par parties $\int_0^{\infty}(\sin t)^n\sin(nt)\times\frac{1}{t^{n+1}}dt$:

$$u_n=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^n\cos(nt)dt+\int_0^{\infty}\frac{(\sin t)^{n-1}\cos t\sin(nt)}{t^n}dt=\int_0^{\infty}\frac{(\sin t)^{n-1}\cos t\sin(nt)}{t^n}dt$$

En écrivant $\sin(nt)\cos t=\sin t\cos(nt)+\sin(n-1)t$, il vient $u_n=0+u_{n-1}$ et donc

$$u_n=u_{n-1}=u_1=\int_0^{\infty}\left(\frac{\sin t}{t}\right)^2 dt=\frac{\pi}{2}.$$