Démonstrations probabilistes de la Formule de Stirling

Ce problème propose deux démonstrations probabilistes de la formule de Stirling

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2n\pi}$$

Partie I: Première preuve

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle de paramètre 1. Pour $n \in \mathbb{N}$, on pose $S_n = \sum_{k=0}^{n} X_k$ et pour tout $n \geqslant 1$ on pose $S_n^* = \dfrac{S_n - n}{\sqrt{n}}$.

On admet que S_n^* converge en loi vers la loi $\mathcal{N}(0,1)$

1. Énoncer le théorème central limite
2. (a) Montrer, par récurrence sur n, que S_n est de loi $\Gamma(n+1,1)$.

   (b) En déduire que la densité de S_n^* s'écrit $g_n(x) = a_n h_n(x)$, avec
   $$a_n = \frac{n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}}{n!}$$
   et
   $$h_n(x) = \frac{1}{\sqrt{2\pi}} e^{-\sqrt{n}x} \left(1 + \frac{x}{\sqrt{n}}\right)^n \chi_{]-\sqrt{n},+\infty[}$$
3. Montrer que
   $$\lim_{n\to+\infty} \int_0^1 g_n(x)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-\frac{x^2}{2}}\,\mathrm{d}x$$
4. En utilisant le théorème de la convergence dominée, montrer que
   $$\lim_{n\to+\infty} \int_0^1 h_n(x)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-\frac{x^2}{2}}\,\mathrm{d}x$$
5. En déduire la formule de Stirling :
   $$n! \sim n^n e^{-n} \sqrt{2\pi n}$$

Partie II: Deuxième preuve

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi de poisson de paramètre 1. On pose $S_n = \sum_{k=1}^{n} X_k$ et $S_n^* = \dfrac{S_n - n}{\sqrt{n}}$

6. Soit X une variable aléatoire réelle qui suit une **loi de Poisson** de paramètre $\lambda > 0$

   (a) Caractériser la loi de X

   (b) Donner l'espérance, la variance et la fonction génératrice de X
7. Soit X et Y deux variables aléatoires indépendantes suivant les lois de poisson de paramètres respectifs λ et μ. Quelle est la loi de $X+Y$
8. Montrer que S_n est de loi $\mathcal{P}(n)$.
9. Montrer que S_n^* converge en loi vers S suivant la loi $\mathcal{N}(0,1)$. En déduire que
   $$\forall t \in \mathbb{R}, \quad \mathbb{P}(S_n^* > t) \xrightarrow[n\to+\infty]{} \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}x = \mathbb{P}(S > t)$$
10. Soit $t \in \mathbb{R}_+^*$

    (a) Montrer que $\mathbb{P}(S_n^* > t) \leqslant \dfrac{\mathbb{E}\left(S_n^{*2}\right)}{t^2}$

Démonstrations probabilistes de la Formule de Stirling

(b) En utilisant le théorème de la convergence dominée, montrer que

$$\int_0^{+\infty} \mathbb{P}\left(S_n^* > t\right) \mathrm{d}t \xrightarrow[n\to+\infty]{} \int_0^{+\infty} \mathbb{P}\left(S > t\right) \mathrm{d}t$$

11. On admet que $\displaystyle\int_0^{+\infty} \left(\int_t^{+\infty} e^{-\frac{x^2}{2}} \mathrm{d}x\right) \mathrm{d}t = \int_0^{+\infty} \left(\int_0^{x} e^{-\frac{x^2}{2}} \mathrm{d}t\right) \mathrm{d}x$. Montrer que

$$\int_0^{+\infty} \mathbb{P}\left(S > t\right) \mathrm{d}t = \frac{1}{\sqrt{2\pi}}$$

12. Soit $n \in \mathbb{N}^*$. En appliquant le TCVD à la série du terme général $\mathbb{P}\left(S_n = k\right) \chi_{\left]0, \frac{k-n}{\sqrt{n}}\right[}$, montrer que

$$\int_0^{+\infty} \mathbb{P}\left(S_n^* > t\right) \mathrm{d}t = \frac{1}{\sqrt{n}e^n} \frac{n^{n+1}}{n!}$$

13. En déduire la formule de Stirling

DÉMONSTRATIONS PROBABILISTES DE LA FORMULE DE STIRLING

Partie I: Première preuve

1. Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles définies sur un espace probabilisé (Ω, T, P) indépendantes, de même loi, possédant une espérance $m = E(X)$ et une variance $\sigma^2 = \mathbb{V}(X) > 0$. On pose

$$\overline{X_n} = \frac{\displaystyle\sum_{i=1}^{n} X_i}{n} \text{ et } Z_n = \frac{\overline{X_n} - m}{\sigma/\sqrt{n}} = \frac{\displaystyle\sum_{i=1}^{n} X_i - nm}{\sigma\sqrt{n}}.$$

Alors $(Z_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire réelle de loi $\mathcal{N}(0,1)$.

2. (a) Par récurrence sur $n \in \mathbb{N}^*$

- Pour $n = 0$, $S_0 = X_0$ est de loi $\mathcal{E}(1) = \Gamma(1,1)$
- Soit $n \in \mathbb{N}$. Supposons que S_n est de loi $\Gamma(n+1,1)$, c'est-à-dire de densité

$$s_n : x \in \mathbb{R} \longmapsto \begin{cases} \dfrac{x^n}{n!}e^{-x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

S_n et X_{n+1} sont indépendantes, donc $S_{n+1} = S_n + X_{n+1}$ est une variable aléatoire à densité de densité est définie par

$$s_{n+1} : x \in \mathbb{R} \longmapsto \int_{-\infty}^{+\infty} s_n(t)s(x-t)\,\mathrm{d}t$$

Avec $s : x \in \mathbb{R} \longmapsto \begin{cases} e^{-x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ la densité de X_{n+1}.

Les variables S_n et X_{n+1} sont positives et indépendantes, donc pour tout $x \leqslant 0$ on a $s_{n+1}(x) = 0$, et pour tout $x > 0$

$$\begin{aligned} s_{n+1}(x) &= \int_0^x s_n(t)s(x-t)\,\mathrm{d}t \\ &= \int_0^x \frac{t^n}{n!}e^{-t}e^{-x+t}\,\mathrm{d}t \\ &= \frac{x^{n+1}}{(n+1)!}e^{-x} \end{aligned}$$

Récurrence achevée.

(b) Soit $x \in \mathbb{R}$, on a :

$$\begin{aligned} \mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leqslant x\right) &= \mathbb{P}\left(S_n \leqslant n + x\sqrt{n}\right) \\ &= \int_{-\infty}^{n+x\sqrt{n}} s_n(t)\,\mathrm{d}t \end{aligned}$$

s_n est nulle sur $\mathbb{R}_-$, donc l'intégrale précédente est nulle si $x \leqslant -\sqrt{n}$.

Pour $x > -\sqrt{n}$, on obtient par la relation de Chasles et la nullité de s_n sur $\mathbb{R}_-$, puis par le changement de variable $t = n + u\sqrt{n}$: $u = \dfrac{t-n}{\sqrt{n}}$

$$\begin{aligned} \mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leqslant x\right) &= \int_0^{n+x\sqrt{n}} \frac{t^n}{n!}e^{-t}\,\mathrm{d}t \\ &= \int_{-\sqrt{n}}^{x} \frac{\left(n+u\sqrt{n}\right)^n}{n!}e^{-\left(n+u\sqrt{n}\right)}\sqrt{n}\,\mathrm{d}u \\ &= \frac{n^{n+\frac{1}{2}}e^{-n}}{n!}\int_{-\sqrt{n}}^{x}\left(1+\frac{u}{\sqrt{n}}\right)^n e^{-u\sqrt{n}}\,\mathrm{d}u \\ &= \frac{n^{n+\frac{1}{2}}e^{-n}}{n!}\int_{-\infty}^{x}\left(1+\frac{u}{\sqrt{n}}\right)^n e^{-u\sqrt{n}}\chi_{]-\sqrt{n},+\infty[}(u)\,\mathrm{d}u \end{aligned}$$

Ainsi la densité de $\dfrac{S_n - n}{\sqrt{n}}$ est la fonction

$$g_n : x \longmapsto \frac{n^{n+\frac{1}{2}}e^{-n}}{n!}\left(1+\frac{x}{\sqrt{n}}\right)^n e^{-x\sqrt{n}}\chi_{]-\sqrt{n},+\infty[}(x)$$

qui s'écrit $g_n(x) = a_n h_n(x)$, avec

$$a_n = \frac{n^{n+\frac{1}{2}}e^{-n}\sqrt{2\pi}}{n!}$$

et

$$h_n : x \longmapsto \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}x}\left(1+\frac{x}{\sqrt{n}}\right)^n \chi_{]-\sqrt{n},+\infty[}(x)$$

3. Par hypothèse $\dfrac{S_n - n}{\sqrt{n}}$ converge en loi une variable de loi $\mathcal{N}(0,1)$, donc

$$\int_0^1 g_n(x)\,\mathrm{d}x = \mathbb{P}\left(0 \leqslant \frac{S_n-n}{\sqrt{n}} \leqslant 1\right) \xrightarrow[n\to+\infty]{} \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

4.
   - (h_n) est une suite de fonctions continues sur $[0,1]$
   - Pour tout $x \in [0,1]$

$$\begin{aligned} h_n(x) &= \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}x}\left(1+\frac{x}{\sqrt{n}}\right)^n \\ &= \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}x+n\ln\left(1+\frac{x}{\sqrt{n}}\right)} \\ &= \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}x+n\left(\frac{x}{\sqrt{n}}-\frac{x^2}{2n}+\circ\left(\frac{1}{n}\right)\right)} \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}+\circ(1)} \end{aligned}$$

   (h_n) converge simplement vers $x \longmapsto \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ qui est continue sur $[0,1]$
   - Pour tout $x \in [0,1]$ et $n \in \mathbb{N}^*$, par l'inégalité de convexité de l'exponentielle : $1+\dfrac{x}{\sqrt{n}} \leqslant e^{\frac{x}{\sqrt{n}}}$

$$|h_n(x)| = \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}x}\left(1+\frac{x}{\sqrt{n}}\right)^n \leqslant \frac{1}{\sqrt{2\pi}}$$

   Et $t \longmapsto \dfrac{1}{\sqrt{2\pi}}$ est continue positive et intégrable sur le segment $[0,1]$

   D'après le théorème de convergence dominée

$$\lim_{n\to+\infty}\int_0^1 h_n(x)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

5. D'après les questions 3 et 4, $a_n \xrightarrow[n\to+\infty]{} 1$, et alors

$$n! \sim n^n e^{-n}\sqrt{2\pi n}$$

Partie II: Deuxième preuve

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi de poisson de paramètre 1. On pose $S_n = \sum_{k=1}^{n} X_k$ et $S_n^* = \dfrac{S_n - n}{\sqrt{n}}$

6. Soit X une variable aléatoire réelle qui suit une **loi de Poisson** de paramètre $\lambda > 0$
   (a) $X(\Omega) = \mathbb{N}$ et $\forall n \in \mathbb{N}$, $\mathbb{P}(X=n) = \dfrac{\lambda^n}{n!}e^{-\lambda}$

(b) $\mathbb{E}(X) = \mathbb{V}(X) = \lambda$ et sa fonction génératrice est donnée par : : $\forall t \in [-1,1]$

$$G_X(t) = \mathbb{E}\left(t^X\right) = \sum_{n=0}^{+\infty} t^n \mathbb{P}(X=n) = \sum_{n=0}^{+\infty} t^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{\lambda(t-1)}$$

7. Par indépendance,

$$G_{X+Y}(t) = E\left(t^{X+Y}\right) = E\left(t^X.t^Y\right) = E\left(t^X\right).E\left(t^Y\right) G_X(t)G_Y(t) = e^{(\lambda+\mu)(t-1)}$$

Pour tout $n \in \mathbb{N}$, on a $\mathbb{P}(X+Y=n) = \dfrac{G_{X+Y}^{(n)}(0)}{n!} = \dfrac{(\lambda+\mu)^n}{n!} e^{-(\lambda+\mu)}$. Autrement-dit $X+Y \hookrightarrow \mathcal{P}(\lambda+\mu)$

8. Par récurrence sur $n \in \mathbb{N}^*$
   - Pour $n=1$, $S_n = X_1$ suit la loi $\mathcal{P}(1)$
   - Soit $n \in \mathbb{N}^*$. Supposons que S_n suit la loi $\mathcal{P}(n)$. Comme S_n et X_{n+1} sont indépendantes, donc $S_{n+1} = S_n + X_{n+1}$ est une variable aléatoire suivant la loi $\mathcal{P}(n+1)$

   Récurrence achevée.

9. La suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes suivant la loi de poisson de paramètre 1, cette dernière admet une espérance et une variance qui valent 1.

   D'après le théorème de la limite centrée $\dfrac{\frac{S_n}{n} - 1}{1/\sqrt{n}}$ converge en loi vers S suivant la loi $\mathcal{N}(0,1)$. Soit $S_n^* = \dfrac{S_n - n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. En déduire que pour tout $t \in \mathbb{R}$

$$\mathbb{P}(S_n^* > t) = 1 - F_{S_n^*}(t) \xrightarrow[n\to+\infty]{} 1 - F_S(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-\frac{x^2}{2}}\, \mathrm{d}x = \mathbb{P}(S > t)$$

10. Soit $t \in \mathbb{R}_+^*$
   (a) Comme $[S_n^* > t] \subset \left[S_n^{*2} > t^2\right] \subset \left[S_n^{*2} \geqslant t^2\right]$, on a

$$\mathbb{P}(S_n^* > t) \geqslant \mathbb{P}\left(S_n^{*2} \geqslant^2\right) \leqslant \frac{\mathbb{E}\left(S_n^{*2}\right)}{t^2}$$

   Par inégalité de Markov appliquée à $S_n^{*2} \geqslant 0$.

   (b)
   - Pour tout $n \in \mathbb{N}^*$, l'application $t \in]0,+\infty[\longmapsto \mathbb{P}(S_n^* > t)$ est continue par morceaux sur $]0,+\infty[$. En effet

$$\begin{aligned} P(S_n^* > t) &= 1 - P(S_n^* \leqslant t) \\ &= 1 - P(S_n \leqslant t\sqrt{n} + n) \\ &= 1 - P(S_n \leqslant \left[t\sqrt{n} + n\right]) \\ &= 1 - F_{S_n}(\leqslant \left[t\sqrt{n} + n\right]) \end{aligned}$$

     F_{S_n} est continue par morceaux car $S_n(\Omega) = \mathbb{N}$ et la fonction partie entière est continue par morceaux sur $\mathbb{R}$, donc par composition de fonctions continues par morceaux l'application considérée est continue par morceaux
   - D'après la question 9, une telle suite converge simplement vers une fonction continue sur $]0,+\infty[$.
   - Pour tout $n \in \mathbb{N}^*$, on a bien, d'après la formule de Huygens kœing, $\mathbb{E}(S_n^{*2}) = \mathbb{V}(S_n^*) + \mathbb{E}(S_n^*)^2 = 1$. On conclut, d'après la question précédente que

$$\forall t \in]0,+\infty[, \quad \mathbb{P}(S_n^* > t) \leqslant \begin{cases} 1 & \text{si } t \leqslant 1 \\ \frac{1}{t^2} & \text{si } t > 1 \end{cases}$$

     Ainsi, on a majoré les termes de la suite par la fonction $\varphi : t \longmapsto \begin{cases} 1 & \text{si } t \leqslant 1 \\ \frac{1}{t^2} & \text{si } t > 1 \end{cases}$ qui est continue positive et intégrable sur $\mathbb{R}_+^*$

   Par le théorème de convergence dominée, on obtient

$$\int_0^{+\infty} \mathbb{P}(S_n^* > t)\, \mathrm{d}t \xrightarrow[n\to+\infty]{} \int_0^{+\infty} \mathbb{P}(S > t)\, \mathrm{d}t$$

11. On admet que $\displaystyle\int_0^{+\infty}\left(\int_t^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}x\right)\mathrm{d}t = \int_0^{+\infty}\left(\int_0^{x} e^{-\frac{x^2}{2}}\,\mathrm{d}t\right)\mathrm{d}x$. On a

$$\begin{aligned}
\int_0^{+\infty}\mathbb{P}(S>t)\,\mathrm{d}t &= \int_0^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\int_t^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}x\right)\mathrm{d}t\\
&= \int_0^{+\infty}\left(\int_0^{x} e^{-\frac{x^2}{2}}\,\mathrm{d}t\right)\mathrm{d}x\\
&= \int_0^{+\infty} xe^{-\frac{x^2}{2}}\,\mathrm{d}x\\
&= \frac{1}{\sqrt{2\pi}}\left[-e^{-\frac{x^2}{2}}\right]_0^{+\infty}\\
&= \frac{1}{\sqrt{2\pi}}
\end{aligned}$$

12. Soit $n\in\mathbb{N}^*$
    - Pour tout $k\in\mathbb{N}$, l'application $t\longmapsto \mathbb{P}(S_n=k)\,\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}$ est positive, continue par morceaux sur $[0,+\infty[$ et intégrable
    - La série $\displaystyle\sum_{k\geqslant 0}\mathbb{P}(S_n=k)\,\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}$ converge normalement, puisque $([S_n=k])_{k\in\mathbb{N}}$ est un système complet d'événements, donc elle converge simplement, de somme $t\longmapsto \mathbb{P}(S_n>t\sqrt{n}+n)=\mathbb{P}(S_n^*>t)=1-F_{S_n^*}(t)$ qui est continue par morceaux sur $]0,+\infty[$. En effet $\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}(t)=1$ équivaut à $k>n$ et $t<\dfrac{k-n}{\sqrt{n}}$ ou encore $k>t\sqrt{n}+n$
    - Pour $k>n$, on a

$$\begin{aligned}
\int_0^{+\infty}\mathbb{P}(S_n=k)\,\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}(t)\,\mathrm{d}t &= \int_0^{\frac{k-n}{\sqrt{n}}}\mathbb{P}(S_n=k)\,\mathrm{d}t\\
&= \frac{k-n}{\sqrt{n}}\mathbb{P}(S_n=k)\\
&= \frac{k-n}{\sqrt{n}}\frac{n^k}{k!}e^{-n}\sim\frac{e^{-n}}{\sqrt{n}}\frac{n^k}{(k-1)!}
\end{aligned}$$

   Or la série $\displaystyle\sum\frac{n^k}{(k-1)!}$ converge

Donc, d'après le théorème de convergence dominée pour les séries,

$$\begin{aligned}
\int_0^{+\infty}\mathbb{P}(S_n^*>t)\,\mathrm{d}t &= \int_0^{+\infty}\mathbb{P}\left(S_n>t\sqrt{n}+n\right)\,\mathrm{d}t\\
&= \int_0^{+\infty}\left(\sum_{k=0}^{+\infty}\mathbb{P}(S_n=k)\,\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}(t)\right)\mathrm{d}t\\
&= \sum_{k=0}^{+\infty}\int_0^{+\infty}\mathbb{P}(S_n=k)\,\chi_{\left[0,\frac{k-n}{\sqrt{n}}\right[}(t)\,\mathrm{d}t\\
&= \sum_{k=n+1}^{+\infty}\int_0^{\frac{k-n}{\sqrt{n}}}\mathbb{P}(S_n=k)\,\mathrm{d}t\\
&= \sum_{k=n+1}^{+\infty}\frac{k-n}{\sqrt{n}}\mathbb{P}(S_n=k)\\
&= \frac{1}{\sqrt{n}}\sum_{k=n+1}^{+\infty}(k-n)\,e^{-n}\frac{n^k}{k!}\\
&= \frac{1}{e^n\sqrt{n}}\sum_{k=n+1}^{+\infty}\left(\frac{n^k}{(k-1)!}-\frac{n^{k+1}}{k!}\right)\\
&= \frac{1}{e^n\sqrt{n}}\frac{n^{n+1}}{n!}\qquad\text{Par télescopage}
\end{aligned}$$

Démonstrations probabilistes de la Formule de Stirling

Ainsi

$$\int_0^{+\infty} \mathbb{P}\left(S_n^* > t\right)\,\mathrm{d}t = \frac{1}{\sqrt{n}e^n}\frac{n^{n+1}}{n!}$$

13. Enfin

$$\int_0^{+\infty} \mathbb{P}\left(S_n^* > t\right)\,\mathrm{d}t = \frac{1}{\sqrt{n}e^n}\frac{n^{n+1}}{n!} \xrightarrow[n\to+\infty]{} \frac{1}{\sqrt{2\pi}}$$

Ce qui nous donne la formule de Stirling.