Fonction de répartition et théorèmes asymptotiques

Pré-requis et notations

— On considère l'espace de probabilité $(\Omega, \mathcal{A}, P)$. Les étudiants suivant le cours de probabilité savent qu'une variable aléatoire à valeur réelle est une application $X : \Omega \to \mathbb{R}$ mesurable, c'est-à-dire :

$$\forall A \in Bor(\mathbb{R}), X^{-1}(A) \in \mathcal{A}.$$

La notion de variable aléatoire ne sera pas formellement rappelée dans ce cours. Par ailleurs, certains étudiants peuvent ne pas suivre le cours de probabilité donc, nous n'utiliserons pas le formalisme décrit ci-dessus.

— La notion de loi d'une variable aléatoire n'est également pas formellement rappelée.

— Soit $n \in \mathbb{N}$. On rappelle que $n! = n \times (n-1) \times \cdots \times 1$ et par convention $0! = 1$.

— Soit $n \in \mathbb{N}^*$ et $k \in \{0, \ldots, n\}$. Le coefficient binomiale $\binom{n}{k}$ est le nombre de combinaisons de k éléments parmi n. Plus précisément, on a :

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

En utilisant la convention que $0! = 1$ on remarque que $\binom{n}{0} = \binom{n}{n} = 1$.

— La formule du triangle de Pascal est donnée ci-dessous :

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

— Soit $A \subset \mathbb{R}$, la fonction indicatrice de l'ensemble A est définie par :

$$\forall x \in \mathbb{R}, \mathbf{1}_{x \in A} = \begin{cases} 1 \text{ si } x \in A, \\ 0 \text{ sinon} \end{cases}.$$

1 Fonction de répartition et quantile

Définition 1 (Fonction de répartition). *Soit X une variable aléatoire réelle. La fonction de répartition de la variable X est définie par :*

$$\begin{array}{rccl} F_X : & \mathbb{R} & \to & [0,1] \\ & x & \mapsto & \mathbb{P}(X \leq x) \end{array}$$

Par définition, on peut remarquer que $\mathbb{P}(x < X \leq z) = F_X(z) - F_X(x)$ pour tout $x < z$.

Proposition 1. *Une fonction de répartition admet les propriétés suivantes :*

1. *F_X est croissante.*
2. *On a les limites suivantes : $\lim_{x\to-\infty} F_X(x) = 0$ et $\lim_{x\to\infty} F_X(x) = 1$.*
3. *F_X est continue à droite en tout point $x \in \mathbb{R}$, c'est-à-dire que l'on a l'égalité :*

$$\lim_{t\to x^+} F(t) = F(x).$$

4. *F_X admet une limite à gauche en tout point $x \in \mathbb{R}$ de plus on a l'égalité*

$$\lim_{t\to x^-} F(t) = \mathbb{P}(X < x).$$

5. *F_X est continue en $x \in \mathbb{R}$ si et seulement si $\mathbb{P}(X = x) = 0$.*

Démonstration.
1) Soit $x \in \mathbb{R}, z \in \mathbb{R}$ avec $x < z$ alors $F_X(z) - F_X(x) = \mathbb{P}(x < X \leq z) \geq 0$. Par conséquent, F_X est une fonction croissante.

2) La série de terme général $(\mathbb{P}(-n-1 < X \leq -n))_{n\geq 0}$ converge, en effet,

$$\mathbb{P}(X \leq 0) = \sum_{n\geq 0} \mathbb{P}(-n-1 < X \leq -n).$$

Par conséquent le reste de cette série converge vers 0, c'est-à-dire que l'on a la limite suivante :

$$0 = \lim_{m\to+\infty} \sum_{n\geq m} \mathbb{P}(-n-1 < X \leq -n) \Rightarrow \lim_{m\to+\infty} \mathbb{P}(X \leq -m) = 0.$$

En utilisant la monotonie de la fonction F_X on en déduit que $\lim_{t\to-\infty} F_X(t) = 0$.

Un raisonnement similaire permet d'établir la seconde limite. La série de terme général $(\mathbb{P}(n < X \leq n+1))_{n\geq 0}$ converge, en effet,

$$\mathbb{P}(X > 0) = \sum_{n\geq 0} \mathbb{P}(n < X \leq n+1).$$

Ainsi le reste de cette série converge vers 0, c'est-à-dire que l'on a la limite suivante :

$$0 = \lim_{m\to+\infty} \sum_{n\geq m} \mathbb{P}(n < X \leq n+1) \Rightarrow \lim_{m\to+\infty} \mathbb{P}(X > m) = 0.$$

Par conséquent $F_X(m) = 1 - \mathbb{P}(X > m)$ tend vers 1 lorsque $m \to +\infty$. Enfin, en utilisant la monotonie de la fonction F_X, on en déduit que $\lim_{t\to+\infty} F_X(t) = 1$.

3) La série de terme général $(\mathbb{P}(x + \frac{1}{n+1} < X \leq x + \frac{1}{n}))_{n\geq 1}$ converge, en effet,

$$\mathbb{P}(x < X \leq x+1) = \sum_{n\geq 1} \mathbb{P}\left(x + \frac{1}{n+1} < X \leq x + \frac{1}{n}\right).$$

Ainsi le reste de cette série converge vers 0, c'est-à-dire que l'on a la limite suivant :

$$0 = \lim_{m\to+\infty} \sum_{n\geq m} \mathbb{P}\left(x + \frac{1}{n+1} < X \leq 1 + \frac{1}{n}\right) \Rightarrow \lim_{m\to+\infty} \mathbb{P}\left(x < X \leq x + \frac{1}{m}\right) = 0.$$

La dernière expression montre que $F_X(x + 1/m)$ converge vers $F_X(x)$. Ainsi, en utilisant la monotonie de la fonction F_X, on en déduit que $\lim_{t\to x^+} F_X(t) = F_X(x)$.

4) La série de terme général $(\mathbb{P}(x - \frac{1}{n} < X \leq x - \frac{1}{n+1}))_{n\geq 1}$ converge, en effet,

$$\mathbb{P}(x-1 < X < x) = \sum_{n\geq 1} \mathbb{P}\left(x - \frac{1}{n} < X \leq x - \frac{1}{n+1}\right).$$

Par conséquent le reste de cette série converge vers 0, c'est-à-dire que l'on a la limite suivante :

$$0 = \lim_{m\to+\infty} \sum_{n\geq m} \mathbb{P}\left(x - \frac{1}{n} < X < x - \frac{1}{n+1}\right) \Rightarrow \lim_{m\to+\infty} \mathbb{P}\left(x - \frac{1}{m} < X < x\right) = 0.$$

La dernière expression montre que $F_X(x - 1/m)$ converge vers $\mathbb{P}(X < x)$. Ainsi, en utilisant la monotonie de la fonction F_X, on en déduit que $\lim_{t\to x^-} F_X(t) = \mathbb{P}(X < x)$.

5) F_X est continue en x si et seulement si les limites à droite et à gauche de F_X en x sont égales, si et seulement si $F_X(x) = \mathbb{P}(X < x)$, si et seulement si $\mathbb{P}(X = x) = 0$. $\square$

Exemple 1. *Soit X une variable aléatoire réelle de loi uniforme sur $\{0, 1, 2\}$ vérifiant $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(X = 2) = 1/3$ alors*

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/3 & \text{si } 0 \leq x < 1 \\ 2/3 & \text{si } 1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

La courbe de cette fonction de répartition est illustrée à la Figure 1.

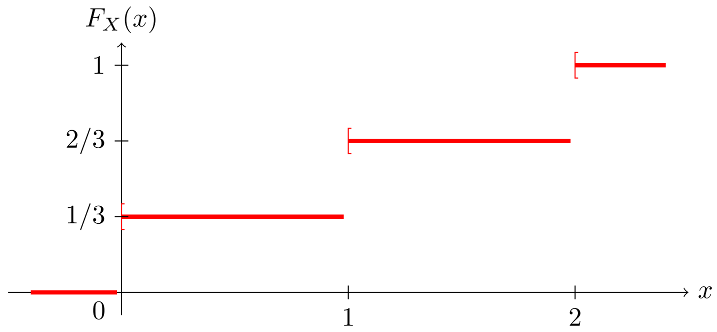

FIGURE 1 – On observe que la fonction F_X est croissante, continue à droite sur $\mathbb{R}$, admet une limite à gauche en tout point, vérifie $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$. Par exemple, $F_X(1) = 2/3$, F_X est continue à droite en 1 car $\lim_{x \to 1^+} F_X(x) = 2/3$, F_X admet une limite à gauche en 1 car $\lim_{x \to 1^-} F_X(x) = 1/3$ mais n'est pas continue à gauche.

La notion de quantile découle de la notion de fonction de répartition.

Définition 2 (Quantile)**.** *Soit X une variable aléatoire réelle de fonction de répartition F_X. Le quantile d'ordre $r \in]0, 1[$ de la variable aléatoire X est défini ci dessous :*

$$q_r = \inf\{x \in \mathbb{R} : F_X(x) \geq r\}.$$

On peut remarquer que lorsque l'équation $F_X(x) = r$ admet une unique solution alors, q_r est l'unique solution de cette équation. En particulier, si F_X est inversible alors $q_r = F_X^{-1}(r)$.

Exemple 2. *Soit X une variable aléatoire réelle de loi exponentielle de paramètre $\lambda > 0$ dont la fonction de répartition est $F_X(x) = 1 - \exp(-\lambda x)$ pour $x \geq 0$ ($F_X(x) = 0$ si $x < 0$). Le quantile d'ordre $r \in]0, 1[$ vaut $q_r = -\frac{\ln(1-r)}{\lambda}$. En particulier, pour $r = 1/2$, la médiane vaut $q_{1/2} = \frac{\ln(2)}{\lambda}$; on reconnaît le temps de demi-vie d'une loi exponentielle.*

Exemple 3. *Soit X une variable aléatoire réelle de loi uniforme sur $\{0, 1, \ldots, n\}$ vérifiant $\mathbb{P}(X = 0) = \cdots = \mathbb{P}(X = n) = 1/(n + 1)$ avec $n \in \mathbb{N}$. Si $n = 2$ la médiane $q_{0,5}$ de la variable aléatoire réelle X vaut 1. En effet :*

$$q_{0,5} = \inf\{x \in \mathbb{R} : F_X(x) \geq 0,5\} = \inf[1, +\infty[= 1.$$

Si $n = 3$ alors la fonction de répartition de X est égale à

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/4 & \text{si } 0 \leq x < 1 \\ 1/2 & \text{si } 1 \leq x < 2 \\ 3/4 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}.$$

Ainsi la médiane $q_{0,5}$ de X vaut également 1. Plus généralement, la médiane de X vaut $q_{0,5} = \lfloor n/2 \rfloor$. Notons que, lorsque n est impair, la médiane de la variable aléatoire X ne coïncide pas avec la notion de médiane (définie dans le secondaire) de la série statistique $\{0, 1, \ldots, n\}$ qui vaut $n/2$.

2 Loi binomiale et approximation asymptotique de cette loi

Nous commençons par introduire la loi de Bernoulli qui est un cas particulier, très important, de la loi binomiale lorsque le nombre d'expérience vaut 1.

2.1 Loi de Bernoulli

Définition 3. *Une variable aléatoire réelle X à valeurs dans $\{0,1\}$ suit une loi de Bernoulli. Cette loi est notée $\mathcal{B}(p)$, où $p \in [0,1]$ est le paramètre tel que $\mathbb{P}(X=1)=p$.*

Exemple mathématique : Soit X une variable aléatoire réelle et $A \subset \mathbb{R}$. La variable aléatoire $Y = \mathbf{1}_{X \in A}$, qui vaut 1 lorsque $X \in A$ et 0 sinon, suit une loi de Bernoulli $\mathcal{B}(p)$ avec $p = \mathbb{P}(X \in A)$.

Exemple en modélisation : Le 15 août 2021, le taux d'incidence de la covid-19 en Bourgogne est de 115.6 pour 100000 habitants. On tire au hasard une personne de cette région et on pose $X=1$ lorsque cette personne est malade et $X=0$ sinon. Alors, X suit une loi de Bernoulli $\mathcal{B}(p)$ avec $p = 115.6/100000$.

2.2 Loi binomiale

Définition 4. *On dit qu'une variable aléatoire réelle S à valeurs dans $\{0,\dots,n\}$ suit une loi binomiale de paramètres (n,p), notée $\mathcal{B}(n,p)$, si pour tout $k \in \{0,\dots,n\}$ on a*

$$\mathbb{P}(S=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

La Proposition 2 établit que la loi binomiale est la loi d'une somme de variables indépendantes et de même loi de Bernoulli.

Proposition 2. *Soit $X_1,\dots,X_n$ une suite de variables aléatoires indépendantes et de même loi de Bernoulli $\mathcal{B}(p)$ (avec $p \in [0,1]$) alors $S_n = X_1 + \cdots + X_n$ suit une loi $\mathcal{B}(n,p)$.*

La démonstration suivante n'utilise que des notions élémentaires. Une démonstration plus courte pourrait être donnée en utilisant la notion de fonction caractéristique (que vous allez voir en cours de probabilité).

Démonstration. Montrons par récurrence que S_n suit une loi $\mathcal{B}(n,p)$. Pour $n=1$ il est clair que S_1 suit une loi de Bernoulli $\mathcal{B}(1,p)$. Soit $n \in \mathbb{N}^*$ un entier pour lequel S_n suit une loi de Binomiale $\mathcal{B}(n,p)$. Montrons que S_{n+1} suit une loi $\mathcal{B}(n+1,p)$, c'est-à-dire montrons que pour tout entier $k \in \{0,\dots,n+1\}$ on a

$$\mathbb{P}(S_{n+1}=k) = \binom{n+1}{k} p^k (1-p)^{n+1-k}$$

Pour $k=0$ on a

$$\mathbb{P}(S_{n+1}=0) = \mathbb{P}(X_1=0,\dots,X_{n+1}=0) = (1-p)^{n+1} = \binom{n+1}{0} p^0 (1-p)^{n+1-0}.$$

Pour $k=n+1$ on a

$$\mathbb{P}(S_{n+1}=n+1) = \mathbb{P}(X_1=1,\dots,X_{n+1}=1) = p^{n+1} = \binom{n+1}{n+1} p^{n+1} (1-p)^{n+1-(n+1)}.$$

A présent, pour $k \in \{1,\dots,n\}$, on a

$$\begin{aligned}
\mathbb{P}(S_{n+1}=k) &= \mathbb{P}(S_n=k, X_{n+1}=0) + \mathbb{P}(S_n=k-1, X_{n+1}=1) \\
&= (1-p)\binom{n}{k} p^k (1-p)^{n-k} + p\binom{n}{k-1} p^{k-1}(1-p)^{n+1-k} \\
&= p^k(1-p)^{n+1-k} \underbrace{\left(\binom{n}{k} + \binom{n}{k-1}\right)}_{=\binom{n+1}{k}} = \binom{n+1}{k} p^k (1-p)^{n+1-k}.
\end{aligned}$$

□

Exemple mathématique : Soit U une variable aléatoire réelle de loi uniforme sur $]0,1[$ (voir le formulaire pour la définition d'une loi uniforme). On pose $X_i \in \{0,1\}$ le $i^{\text{ème}}$ chiffre de la partie fractionnaire de U en base 2 alors $X_1, \ldots, X_n$ est une suite de variables aléatoires indépendantes et de même de loi $\mathcal{B}(1/2)$. Ainsi, $S_n = X_1 + \cdots + X_n$ suit une loi de binomiale $\mathcal{B}(n, 1/2)$.

Exemple en modélisation : Le 15 août 2021, le taux d'incidence de la Bourgogne est de 115.6 pour 100000 habitants. On tire au hasard n personnes de cette région (avec remise) et on pose $X_i = 1$ lorsque la $i^{\text{ème}}$ personne est malade de la covid-19 et $X_i = 0$ sinon. Alors, $S_n = X_1 + \cdots + X_n$ suit loi binomiale $\mathcal{B}(n,p)$ avec $p = 115.6/100000$.

2.3 Approximation d'une loi binomiale par une loi normale

Le résultat suivant, énoncé pour la première fois par Abraham de Moivre au 18ème siècle dans le cas d'équiprobabilité et généralisé par Pierre-Simon de Laplace au début du 19ème siècle, donne une approximation de la loi binomiale par la loi normale lorsque le nombre d'expériences n tend vers $+\infty$.

Théorème 1 (De Moivre-Laplace)**.** *Soit S une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n,p)$ avec $p \in]0,1[$. Alors, la variable S centrée réduite converge en loi[1] vers la loi normale $\mathcal{N}(0,1)$ lorsque n tend vers $+\infty$ au sens suivant :*

$$\forall t \in \mathbb{R} \qquad \lim_{n\to+\infty} \mathbb{P}\left(\frac{S-np}{\sqrt{np(1-p)}} \le t\right) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx.$$

La notation classique pour cette convergence (que vous utiliserez en cours de probabilité) est :

$$\frac{S-np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Ainsi, pour n assez grand, on peut approcher la loi $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}(np, np(1-p))$. En annexe une preuve d'un cas particulier du théorème de Moivre-Laplace est donnée. Pour ce cours, la règle d'usage [2] que pour employer cette approximation est la variance de S dépasse 5 ; c'est-à-dire $np(1-p) > 5$.

Une variable aléatoire de loi binomiale est discrète à valeurs dans $\{1, \ldots, n\}$, alors qu'une variable aléatoire de loi normale est continue et à valeurs dans $\mathbb{R}$. Dans le cas d'une loi binomiale, un point a une probabilité non nulle alors que pour la loi normale, un point est un ensemble de probabilité nulle. Pour ces deux raisons, il est préférable de faire une « correction de continuité » quand on utilise l'approximation d'une loi binomiale par une loi normale :

$$\mathbb{P}(S \le k) = \mathbb{P}(S \le k + 0,5) \approx \mathbb{P}(Y \le k + 0,5),$$

où Y est une variable aléatoire de loi $\mathcal{N}(np, np(1-p))$. Lorsque le nombre d'expériences est grand ($n > 20$) mais que la variance de S est trop petite ($np(1-p) \le 5$), il est plus précis d'approcher la loi binomiale par une loi de Poisson que par une loi normale. Les fondements théoriques de cette approximation sont donnés ci-dessous.

2.4 Approximation d'une loi binomiale par une loi de Poisson

Théorème 2. *Soit S une variable aléatoire de loi $\mathcal{B}(n, p_n)$ tel que $np_n \to \lambda > 0$ lorsque n tend vers $+\infty$. Alors S converge en loi vers une loi de Poisson $\mathcal{P}(\lambda)$ (voir le formulaire pour la définition de cette loi) lorsque n tend vers $+\infty$ au sens suivant :*

$$\forall k \in \mathbb{N} \qquad \lim_{n\to+\infty} \mathbb{P}(S = k) = \exp(-\lambda)\frac{\lambda^k}{k!}.$$

Démonstration. Donnons un équivalent asymptotique de la probabilité :

$$\mathbb{P}(S = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}.$$

Asymptotiquement, lorsque n tend vers $+\infty$, on a

$$(1-p_n)^{n-k} = \exp((n-k)\ln(1-p_n)) \sim_{n\to+\infty} \exp(-\lambda)$$

1. La définition de convergence en loi sera proprement définie en cours de probabilité.
2. Il existe d'autres règles d'usages pour l'approximation d'une loi binomiale par une loi normale ; par exemple $n > 30$, $np > 5$ et $n(1-p) > 5$.

Ainsi, lorsque $k = 0$, on a $\mathbb{P}(S = 0) = (1 - p_n)^n \sim_{n\to+\infty} \exp(-\lambda)$. On suppose à présent que $k \geq 1$. On a $p_n^k \sim_{n\to+\infty} \lambda^k/n^k$; par ailleurs

$$\binom{n}{k} = \frac{n \times (n-1) \times \cdots \times (n+1-k)}{k!} \sim_{n\to+\infty} \frac{n^k}{k!}.$$

Ainsi, on obtient l'équivalent asymptotique :

$$\mathbb{P}(S = k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k} \sim_{n\to+\infty} \frac{\lambda^k}{k!} \exp(-\lambda).$$

□

Pour un nombre d'expériences n assez grand et une probabilité de succès p assez petite (resp. une probabilité d'échec $1-p$ assez petite), on peut approcher la loi du nombre de succès S (resp. du nombre d'échecs $n-S$) d'une loi $\mathcal{B}(n, p)$ par une loi de Poisson. Une règle d'usage pour utiliser cette approximation est donnée ci-dessous :

— Lorsque $n > 20$, $np(1 - p) \leq 5$ et $p < 0.5$, on approxime la loi du nombre de succès S par une loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda = np$.

— Lorsque $n > 20$, $np(1 - p) \leq 5$ et $p > 0.5$ on approxime la loi du nombre d'échecs $n - S$ par une loi $\mathcal{P}(\lambda)$ avec $\lambda = n(1 - p)$.

3 Théorèmes asymptotiques

Deux théorèmes ont une place particulière en théorie des probabilités et en statistiques : la loi des grands nombres et le théorème de la limite centrale. Ils interviennent dans l'étude de phénomènes aléatoires comportant un grand nombre de variables aléatoires réelles indépendantes de même loi.

3.1 Loi des grands nombres

Théorème 3 (Loi des grands nombres)**.** *Soit $X_1, ..., X_n$ des variables aléatoires réelles indépendantes de même loi de variance $\sigma^2 \geq 0$ et d'espérance $\mu \in \mathbb{R}$. On pose $\overline{X} = \frac{X_1+\cdots+X_n}{n}$ la moyenne aléatoire (également nommée moyenne empirique). La variable $\overline{X}$ converge en probabilité vers μ au sens suivant :*

$$\forall \epsilon > 0 \quad \mathbb{P}(|\overline{X} - \mu| > \epsilon) \underset{n\to+\infty}{\longrightarrow} 0.$$

La notation classique pour cette convergence (que vous utiliserez en cours de probabilité) est :

$$\overline{X} \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} \mu.$$

Démonstration. La preuve est basée sur l'inégalité de Bienaymé-Tchebichev. Soit Y une variable réelle dont la variance est bien définie alors

$$\mathbb{P}(|Y - \mathbb{E}(Y)| > \epsilon) \leq \frac{\mathrm{var}(Y)}{\epsilon^2}.$$

Établissons cette inégalité.

$$\mathbb{P}(|Y - \mathbb{E}(Y)| > \epsilon) = \mathbb{E}(\mathbf{1}_{|Y-\mathbb{E}(Y)|>\epsilon}) = \mathbb{E}(\mathbf{1}_{(Y-\mathbb{E}(Y))^2>\epsilon^2}) \leq \mathbb{E}\left(\frac{(Y - \mathbb{E}(Y))^2}{\epsilon^2}\right) = \frac{\mathrm{var}(Y)}{\epsilon^2}.$$

À présent, appliquons cette inégalité à la moyenne aléatoire. Clairement, $\mathbb{E}(\overline{X}) = \mu$ et $\mathrm{var}(\overline{X}) = \frac{1}{n}(\mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n)) = \sigma^2/n$. Ainsi,

$$\mathbb{P}\left(|\overline{X} - \mu| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \underset{n\to+\infty}{\longrightarrow} 0.$$

□

Le corollaire suivant parfois nommé « lemme de l'image continue » donne un raffinement de la loi des grands nombres (voir le cours de probabilité).

Corollaire 1 (Image continue)**.** *Soit $X_1, ..., X_n$ des variables aléatoires réelles indépendantes de même loi de variance $\sigma^2 \geq 0$ et d'espérance $\mu \in \mathbb{R}$; on pose $\overline{X} = \frac{X_1+\cdots+X_n}{n}$. Soit g une fonction définie sur $\mathbb{R}$ et continue en μ alors $g(\overline{X})$ converge en probabilité vers $g(\mu)$.*

3.2 Théorème de la limite centrale

Cette section vise à généraliser le théorème de Moivre-Laplace. En particulier, le théorème de la limite centrale montre que la moyenne aléatoire centrée réduite suit asymptotiquement une loi normale lorsque la taille de l'échantillon n tend vers $+\infty$. La définition de la loi normale est rappelée ci-dessous :

Définition 5. *Soit $\mu \in \mathbb{R}$ et $\sigma^2 > 0$. On dit que X suit une loi normale de paramètres (μ, σ^2), notée $\mathcal{N}(\mu, \sigma^2)$, si la loi de X a pour densité :*

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right).$$

On dit que Z *suit une loi normale centrée réduite notée* $\mathcal{N}(0,1)$, si la loi de Z a pour densité :

$$f_Z(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right).$$

Exemples mathématique :

— Soit U, V deux variables aléatoires indépendantes de loi $\mathcal{U}(0,1)$ alors, les variables aléatoires Z_1 et Z_2 définies par :
$$Z_1 = \sqrt{-2\ln(U)}\cos(2\pi V) \text{ et } Z_2 = \sqrt{-2\ln(U)}\sin(2\pi V)$$
sont indépendantes et de même loi $\mathcal{N}(0,1)$. [3]

— Soit S une variable aléatoire réelle de loi $\mathcal{B}(n,p)$ avec $p \in]0,1[$ alors, d'après le théorème de Moivre-Laplace, la variable centrée réduite : $(S-np)/\sqrt{np(1-p)}$ suit asymptotiquement, lorsque $n \to +\infty$, une loi normale $\mathcal{N}(0,1)$.

Exemple en modélisation : Les femmes américaines adultes mesurent en moyenne 163.83 cm avec un écart-type de 6.35 cm. On peut modéliser la taille aléatoire d'une femme américaine choisie au hasard par une loi normale $\mathcal{N}(163.8, 6.3^2)$.

Proposition 3. *La loi normale satisfait les propriétés suivantes :*

1. *Soit X suit une variable aléatoire ayant une loi normale $N(\mu, \sigma^2)$ alors $Z = \frac{X-\mu}{\sigma}$ suit une loi $\mathcal{N}(0,1)$.*
2. *Soit $X_1, \dots, X_n$ une suite de variables aléatoires indépendantes de loi respective $\mathcal{N}(\mu_1, \sigma_1^2), \dots, \mathcal{N}(\mu_1, \sigma_1^2)$ et $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. La variable aléatoire $Y = \sum_{i=1}^n \alpha_i X_i$ suit une loi normale dont les paramètres sont $\mu = \sum_{i=1}^p \alpha_i \mu_i$ et $\sigma^2 = \sum_{i=1}^p \alpha_i^2 \sigma_i^2$.*

Démonstration. Montrons que Z suit une loi $\mathcal{N}(0,1)$. Clairement,

$$\mathbb{P}(Z \le z) = \mathbb{P}(X \le \sigma z + \mu) = \int_{-\infty}^{\sigma z+\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

Ainsi, en posant $t = (x-\mu)/\sigma$ on en déduit que

$$\mathbb{P}(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Donc, Z suit une loi normale $\mathcal{N}(0,1)$.

La preuve de ii) n'est pas donnée. On pourrait néanmoins prouver ce résultat en calculant la fonction caractéristique de Y. □

Exemple 4 (moyenne aléatoire d'un échantillon d'une loi normale)**.** *Soit $X_1, \dots, X_n$ une suite de variables aléatoires indépendantes et de même loi $\mathcal{N}(\mu, \sigma^2)$, alors la moyenne aléatoire $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$ suit une loi normale dont les paramètres sont :*

— *L'espérance est : $\mathbb{E}(\frac{1}{n}\sum_{i=1}^n X_i) = \mu$.*

— *La variance est : $\mathrm{var}(\frac{1}{n}\sum_{i=1}^n X_i) = \sigma^2/n$.*

Ainsi, en centrant et en réduisant la moyenne aléatoire, on en déduit que $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ suit une loi normale $\mathcal{N}(0,1)$.

Dans le cas général, sans faire l'hypothèse les variables aléatoires $X_1, \dots, X_n$ sont normales, on montre néanmoins que la moyenne aléatoire centrée et réduite suit asymptotiquement une loi normale.

3. Voir la méthode de Box-Muller

Théorème 4 (théorème de la limite centrale)**.** *Soit $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes, de même loi d'espérance $\mu \in \mathbb{R}$ et de variance $\sigma^2 > 0$. La moyenne aléatoire $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$ centrée réduite converge en loi vers une loi normale $\mathcal{N}(0,1)$ au sens suivant :*

$$\forall x \in \mathbb{R} \qquad \lim_{n\to+\infty} \mathbb{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

La notation classique pour cette convergence (que vous utiliserez en cours de probabilité) est :

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(0,1).$$

Exemple 5. *Lorsque $X_1, \dots, X_n$ sont des variables aléatoires réelles indépendantes, de même loi de Bernoulli $\mathcal{B}(p)$ avec $p \in]0,1[$, dont l'espérance est p et la variance est $p(1-p)$ alors, d'après le théorème de la limite centrale on a la convergence en loi suivante*

$$\frac{\overline{X}-p}{\sqrt{p(1-p)}/\sqrt{n}} \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(0,1).$$

En remarquant que $S = \sum_{i=1}^n X_i$ suit une loi binomiale $\mathcal{B}(n,p)$ et en posant $\overline{X} = S/n$ on retrouve la formulation du théorème de Moivre-Laplace.

Le corollaire suivant parfois nommé « méthode-delta » donne un raffinement du théorème de la limite centrale.

Corollaire 2 (méthode-delta)**.** *Soit $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes, de même loi d'espérance $\mu \in \mathbb{R}$ et de variance $\sigma^2 > 0$; on pose $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$. Soit g une fonction définie sur $\mathbb{R}$ et dérivable en μ alors la convergence en loi suivante est satisfaite :*

$$\lim_{n\to+\infty} \mathbb{P}\left(\frac{g(\overline{X})-g(\mu)}{g'(\mu)\sigma/\sqrt{n}} \leq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Intuitivement, cette convergence en loi s'obtient en effectuant un développement limité à l'ordre 1 de g au voisinage de μ ; en effet, $\frac{g(\overline{X})-g(\mu)}{g'(\mu)\sigma/\sqrt{n}} \approx \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$.

Annexe : théorèmes asymptotiques pour pile ou face

Dans le cadre de la loi binomiale, Jacques Bernoulli fut le premier à établir la loi des grands nombres, c'est-à-dire la convergence de la fréquence de succès vers la probabilité du succès [1]. Lorsque l'expérience est pile ou face, les probabilités sont proportionnelles aux coefficients binomiaux ce qui facilite l'étude asymptotique de ce jeu aléatoire. La proposition suivante établit la loi des grands nombres pour le cas particulier d'équiprobabilité : $p = 1/2$ et lorsque le nombre d'expériences est pair.

Proposition 4 (Loi des grands nombres pour pile ou face)**.** *Soit S une variable aléatoire de loi $\mathcal{B}(2n, 1/2)$ avec $n \in \mathbb{N}^*$ alors pour tout $\epsilon > 0$ on a*

$$\mathbb{P}\left(\left|\frac{S}{2n} - \frac{1}{2}\right| > \epsilon\right) \leq 2\sqrt{\frac{n}{\pi}} \left(\frac{1}{1+4\epsilon}\right)^{\epsilon n - 1/2}$$

Clairement, l'inégalité donnée à la Proposition 4 permet de montrer que pour tout $\epsilon > 0$ on a

$$\lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{S}{2n} - \frac{1}{2}\right| > \epsilon\right) = 0.$$

Notons que l'inégalité donnée à la Proposition 4 permet également de donner un intervalle de fluctuation pour une loi binomiale dans le cas d'équiprobabilité. Motivé par les premiers travaux sur la loi des grands nombres par Jacques Bernoulli, Abraham de Moivre détermina un intervalle de fluctuation asymptotique d'une loi binomiale lorsque l'expérience est pile ou face [3]. Dans un formalisme moderne, la construction de cet intervalle de fluctuation repose sur la proposition suivante.

Proposition 5 (Intervalle de fluctuation asymptotique pour pile ou face)**.** *Soit S que variable aléatoire de loi $\mathcal{B}(2n,1/2)$ avec $n\in\mathbb{N}^*$ et $t\geq 0$ alors*

$$\lim_{n\to+\infty}\mathbb{P}\left(0\leq\frac{S-n}{\sqrt{n/2}}\leq t\right)=\int_0^t\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)dx.$$

En particulier, Abraham de Moivre évalue cette probabilité avec beaucoup de précision pour $t\in\{1,2,3\}$. La généralisation de cette proposition au cas où $p\in]0,1[$ est quelconque, communément appelée Théorème de Moivre-Laplace, fut établi par Pierre-Simon de Laplace [2].

Preuves des Propositions 4 et 5

Les preuves des Propositions 4 et 5, qui diffèrent légèrement des preuves historiques faîtes par Jacques Bernoulli et Abraham de Moivre, requièrent des notions mathématiques de DEUG [4]. Ces preuves ne nécessitent ni l'inégalité de Bienaymé-Tchebichev [5] ni la fonction caractéristique. Vous allez étudier la notion de fonction caractéristique ainsi que l'inégalité Bienaymé-Tchebichev plus tard en cours de probabilité. Le Lemme 1, qui donne un encadrement de la probabilité d'avoir le même nombre de piles et de faces, est utile pour prouver les Propositions 4 et 5.

Lemme 1. *Soit S une variable aléatoire de loi $\mathcal{B}(2n,1/2)$ avec $n\in\mathbb{N}^*$. La probabilité que S soit égale à son espérance est encadrée par :*

$$\frac{1}{\sqrt{\pi(n+1/2)}}\leq\mathbb{P}(S=n)\leq\frac{1}{\sqrt{\pi n}}.$$

Démonstration. La probabilité que la variable S soit égale à son espérance vaut :

$$\mathbb{P}(S=n)=\binom{2n}{n}\frac{1}{2^n}\frac{1}{2^n}=\frac{(2n)!}{4^n(n!)^2}.$$

Cette probabilité est liée à l'intégrale $I_m=\int_0^1(1-u^2)^{m/2}du$ via les formules

$$I_{2n}=\frac{1}{\mathbb{P}(S=n)(2n+1)}\text{ et }I_{2n-1}=\mathbb{P}(S=n)\frac{\pi}{2}.$$

que nous allons démontrer. Un simple calcul donne $I_0=1, I_1=\pi/4$ (aire du quart de disque). De plus, la suite $(I_m)_{m\in\mathbb{N}}$ vérifie l'équation de récurrence : $I_{m+2}=\frac{m+2}{m+3}I_m$. Montrons, via une intégration par partie, cette identité :

$$\begin{aligned}
I_{m+2} &= \int_0^1 1\times(1-u^2)^{\frac{m+2}{2}}du,\\
&= \underbrace{\left[u(1-u^2)^{\frac{m+2}{2}}\right]_0^1}_{=0}+(m+2)\int_0^1 u^2(1-u^2)^{\frac{m}{2}}du,\\
&= -(m+2)\int_0^1(1-u^2)(1-u^2)^{\frac{m}{2}}-(1-u^2)^{\frac{m}{2}}du,\\
&= (m+2)(-I_{m+2}+I_m).
\end{aligned}$$

On déduit de cette dernière égalité que $I_{m+2}=\frac{m+2}{m+3}I_m$. Par récurrence on obtient :

$$\begin{aligned}
I_{2n} &= \frac{2n\times2(n-1)\times\cdots\times2}{(2n+1)\times(2n-1)\times3}=\frac{4^n(n!)^2}{(2n+1)!}=\frac{1}{\mathbb{P}(S=n)(2n+1)},\\
I_{2n-1} &= \frac{(2n-1)\times(2n-3)\times\cdots\times3}{2n\times2(n-1)\times\cdots\times2}\frac{\pi}{2}=\frac{(2n)!}{4^n(n!)^2}\frac{\pi}{2}=\mathbb{P}(S=n)\frac{\pi}{2}.
\end{aligned}$$

4. Ce qui ne signifie pas que ces preuves sont faciles.
5. Historiquement, cette inégalité fut établie bien plus tard ; la notion de variance n'était pas connue à l'époque où Jacques Bernoulli donna sa preuve de la loi des grands nombres. Néanmoins, la preuve de la Proposition 4 qui repose sur cette inégalité est plus simple que celle que nous allons donner dans le cas particulier du jeu pile ou face.

Comme la suite $(I_m)_{m\in\mathbb{N}}$ est décroissante on en déduit l'encadrement suivant :

$$1 \leq \frac{I_{2n-1}}{I_{2n}} = (2n+1)\mathbb{P}(S=n)^2\pi/2 \leq \frac{I_{2n-2}}{I_{2n}} = \frac{2n+1}{2n}.$$

Ainsi, $1/\sqrt{\pi(n+1/2)} \leq \mathbb{P}(S=n) \leq 1/\sqrt{\pi n}$.

□

Démonstration de la Proposition 4

Démonstration de la Proposition 4. La probabilité que l'on cherche à majorer peut s'écrire comme une somme de la façon suivante :

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{S}{2n} - \frac{1}{2}\right| > \epsilon\right) &= \mathbb{P}\left(|S-n| > 2n\epsilon\right) \\
&= 2\mathbb{P}(S > n + 2n\epsilon) \\
&= 2\sum_{k=\lceil 2n\epsilon\rceil}^{n} \mathbb{P}(S = n+k)
\end{aligned}$$

On pose $r = \lceil 2n\epsilon \rceil$. Comme la suite $\mathbb{P}(S=n+k)_{k\in\{0,1,\ldots,n\}}$ est décroissante, on en déduit que

$$P\left(\left|\frac{S}{2n} - \frac{1}{2}\right| > \epsilon\right) \leq 2n\mathbb{P}(S=n+r) \tag{1}$$

Notons que $r \geq 1$ et dans le cas particulier où $r = 1$ alors $n\epsilon \leq 1/2$ ainsi, d'après le Lemme 1, on obtient l'inégalité suivante :

$$\left(\left|\frac{S}{2n} - \frac{1}{2}\right| > \epsilon\right) \leq 2n\mathbb{P}(S=n+1) \leq 2n\mathbb{P}(S=n) \leq 2\sqrt{\frac{n}{\pi}} \leq 2\sqrt{\frac{n}{\pi}}\left(\frac{1}{1+4\epsilon}\right)^{\epsilon n - 1/2}.$$

Nous supposons à présent que $r \in \{2, \ldots, n\}$. Nous allons majorer la somme décrite dans l'équation (1) en majorant $\mathbb{P}(S=n+r)/\mathbb{P}(S=n)$. Ce quotient a une expression explicite donnée ci-dessous :

$$\begin{aligned}
\frac{\mathbb{P}(S=n+r)}{\mathbb{P}(S=n)} &= \frac{(n!)^2}{(n+r)!(n-r)!}, \\
&= \frac{n\times(n-1)\times\cdots\times(n+1-r)}{(n+1)\times\cdots\times(n+r)}, \\
&= \frac{n}{n+r}\times\frac{1-1/n}{1+1/n}\times\cdots\times\frac{1-(r-1)/n}{1+(r-1)/n}.
\end{aligned} \tag{2}$$

En majorant $n/(n+r)$ par 1 et en utilisant l'inégalité suivante

$$\forall i \in \{1,\ldots,n\} \quad \frac{1-i/n}{1+i/n} \leq \frac{1}{1+2i/n}$$

on en déduit que

$$\frac{\mathbb{P}(S=n+r)}{\mathbb{P}(S=n)} \leq \prod_{i=1}^{r-1}\frac{1}{1+2i/n}.$$

Comme pour $i \in \{1,\ldots,r-1\}$ l'inégalité suivante est satisfaite

$$\frac{1}{1+2i/n}\times\frac{1}{1+2(k-i)/n} \leq \frac{1}{1+2r/n},$$

on en déduit que le carré du quotient $\mathbb{P}(S=n+r)/\mathbb{P}(S=n)$ admet la majoration suivante :

$$\left(\frac{\mathbb{P}(S=n+r)}{\mathbb{P}(S=n)}\right)^2 \leq \prod_{i=1}^{r-1}\frac{1}{1+2i/n}\prod_{i=1}^{r-1}\frac{1}{1+2(r-i)/n} \leq \left(\frac{1}{1+2r/n}\right)^{r-1}. \tag{3}$$

Ainsi, en utilisant en utilisant le Lemme 1 et l'inégalité (3), on obtient une majoration souhaitée :

$$\begin{aligned}\mathbb{P}\left(\left|\frac{X}{2n}-\frac{1}{2}\right|>\epsilon\right) &\leq 2n\mathbb{P}(S=n)\left(\frac{1}{1+2r/n}\right)^{(r-1)/2}\\ &\leq 2n\mathbb{P}(S=n)\left(\frac{1}{1+4\epsilon}\right)^{\epsilon n-1/2}\\ &\leq 2\sqrt{\frac{n}{\pi}}\left(\frac{1}{1+4\epsilon}\right)^{\epsilon n-1/2}\end{aligned}$$

□

Combiné avec le Lemme 1, les Lemmes 2 et 3 permettent d'achever la preuve de la Proposition 5.

Lemme 2. *Soit S une variable aléatoire de loi $\mathcal{B}(2n,1/2)$ avec $n\in\mathbb{N}^*$ et $k\in\{0,\ldots,\lceil n/2\rceil\}$. L'encadrement suivant :*

$$\mathbb{P}(S=n)\exp\left(-\frac{k^2}{n}\right)\frac{n}{n+k}\exp\left(\frac{k}{n}+C\frac{k^4}{n^3}\right)\leq\mathbb{P}(S=n+k)\leq\mathbb{P}(S=n)\exp\left(-\frac{k^2}{n}\right)\exp\left(\frac{k}{n}\right)\tag{4}$$

est satisfait pour une certaine constante $C\leq 0$.

Démonstration. Pour $k\in\{0,1\}$, l'encadrement (4) est satisfait. En effet, pour $k=0$ la borne inférieure et la borne supérieure de l'encadrement sont égales à $\mathbb{P}(S=n)$ et pour $k=1$ on a $\mathbb{P}(S=n+1)/\mathbb{P}(S=n)=n/(n+1)$ donc, avec $C\leq 0$, on a $\mathbb{P}(S=n)\frac{n}{n+1}\exp\left(\frac{C}{n^3}\right)\leq\mathbb{P}(S=n+1)\leq\mathbb{P}(S=n)$. On suppose à présent que que $k\geq 2$.

On pose $f(x)=\ln((1-x)/(1+x))$. On a $f(0)=0$, $f'(0)=-2$, $f''(x)\leq 0$, $f''(0)=0$ et en posant $C=\inf\{f'''(x):x\in[0;0,5]\}\leq 0$ alors pour $x\in[0;0,5]$ on a l'encadrement

$$\begin{aligned}&-2x+Cx^3\leq f(x)\leq -2x\\ \Rightarrow\quad&\exp(-2x+Cx^3)\leq\frac{1-x}{1+x}\leq\exp(-2x).\end{aligned}\tag{5}$$

D'après l'équation (2), le quotient $\mathbb{P}(S=n+k)/\mathbb{P}(S=n)$ a une expression explicite rappelée ci-dessous :

$$\frac{\mathbb{P}(S=n+k)}{\mathbb{P}(S=n)}=\frac{n}{n+k}\times\frac{1-1/n}{1+1/n}\times\cdots\times\frac{1-(k-1)/n}{1+(k-1)/n}.$$

Ainsi, en majorant $n/(n+k)$ par 1 et en utilisant la majoration donnée à l'expression (5) on obtient :

$$\mathbb{P}(S=n+k)\leq\mathbb{P}(S=n)\exp\left(-\frac{2}{n}(1+\cdots+(k-1))\right)=\mathbb{P}(S=n)\exp\left(-\frac{k^2}{n}\right)\exp\left(\frac{k}{n}\right).$$

Enfin en utilisant la minoration donnée à l'expression (5) (valable car $0\leq 1/n\leq\cdots\leq(k-1)/n\leq 1/2$) on obtient :

$$\begin{aligned}\mathbb{P}(S=n+k) &\geq \mathbb{P}(S=n)\frac{n}{n+k}\exp\left(-\frac{2}{n}(1+\cdots+(k-1))+\frac{C}{n^3}(1^3+\cdots+(k-1)^3)\right),\\ &\geq \mathbb{P}(S=n)\exp\left(-\frac{k^2}{n}\right)\frac{n}{n+k}\exp\left(\frac{k}{n}+C\frac{k^4}{n^3}\right).\end{aligned}$$

□

Lemme 3. *Soit $t\geq 0$. La limite suivante est satisfaite lorsque $n\to+\infty$:*

$$\lim_{n\to+\infty}\frac{1}{\sqrt{\pi}}\sum_{k=0}^{\left\lfloor t\sqrt{n/2}\right\rfloor}\frac{1}{\sqrt{n}}\exp\left(-\frac{k^2}{n}\right)=\int_0^t\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)dt.$$

Démonstration. Dans l'expression suivante on reconnaît une somme de Riemann de pas $1/\sqrt{n}$ et dont le nombre de pas est $\left\lfloor t\sqrt{n/2}\right\rfloor + 1$

$$\frac{1}{\sqrt{\pi}} \sum_{k=0}^{\left\lfloor t\sqrt{n/2}\right\rfloor} \frac{1}{\sqrt{n}} \exp\left(-\left(\frac{k}{\sqrt{n}}\right)^2\right).$$

Ainsi, la limite de cette somme lorsque $n \to +\infty$ est donnée par l'intégrale suivante :

$$\lim_{n\to+\infty} \frac{1}{\sqrt{\pi}} \sum_{k=0}^{\left\lfloor t\sqrt{n/2}\right\rfloor} \frac{1}{\sqrt{n}} \exp\left(-\left(\frac{k}{\sqrt{n}}\right)^2\right) = \int_0^{t/\sqrt{2}} \frac{1}{\sqrt{\pi}} \exp\left(-u^2\right) dx.$$

Enfin, en effectuant le changement de variable $x = \sqrt{2}u$ on obtient le résultat souhaité

$$\int_0^{t/\sqrt{2}} \frac{1}{\sqrt{\pi}} \exp\left(-u^2\right) du = \int_0^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

□

Démonstration de la Proposition 5

Démonstration de la Proposition 5. Les égalités suivantes sont satisfaites :

$$\begin{aligned} \mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) &= \mathbb{P}\left(n \leq S \leq n + t\sqrt{n/2}\right), \\ &= \sum_{k=0}^{\lfloor t\sqrt{n/2}\rfloor} \mathbb{P}(S = n+k). \end{aligned} \tag{6}$$

Les Lemmes 1 et 2 fournissent une majoration de la somme décrite à l'équation (6). Ainsi :

$$\mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) \leq \exp\left(\frac{t}{\sqrt{2n}}\right) \frac{1}{\sqrt{\pi}} \sum_{k=0}^{\left\lfloor t\sqrt{n/2}\right\rfloor} \frac{1}{\sqrt{n}} \exp\left(-\left(\frac{k}{\sqrt{n}}\right)^2\right).$$

Comme $\lim_{n\to+\infty} \exp(t\sqrt{2n}) = 1$, en utilisant le Lemme 3 on obtient une majoration de la limite supérieure :

$$\limsup_{n\to+\infty} \mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) \leq \int_0^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Pour n assez grand tel que $\lfloor t\sqrt{n/2}\rfloor \leq \lceil n/2 \rceil$, les Lemmes 1 et 2 fournissent également une minoration de la somme décrite à l'équation (6). Ainsi :

$$\mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) \geq \exp\left(\frac{t}{\sqrt{2n}} + C\frac{t^4}{4n^2}\right) \frac{1}{\sqrt{\pi}} \sum_{k=0}^{\left\lfloor t\sqrt{n/2}\right\rfloor} \frac{1}{\sqrt{n}} \exp\left(-\left(\frac{k}{\sqrt{n}}\right)^2\right).$$

Comme $\lim_{n\to+\infty} \exp(t\sqrt{2n} + Ct^4/(4n^2)) = 1$, en utilisant le Lemme 3 on obtient une minoration de la limite inférieure :

$$\liminf_{n\to+\infty} \mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) \geq \int_0^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Ainsi, la limite inférieure et la limite supérieure sont égales d'où

$$\lim_{n\to+\infty} \mathbb{P}\left(0 \leq \frac{S-n}{\sqrt{n/2}} \leq t\right) = \int_0^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

□

Références

[1] Jakob Bernoulli. *Ars conjectandi, opus posthumum : accedit tractatus de seriebus infinitis, et epistola Gallice scripta de ludo pilæ reticularis*. Impensis Thurnisiorum Fratrum, 1713.

[2] Pierre Simon de Laplace. *Théorie analytique des probabilités*, volume 7. Courcier, 1820.

[3] Abraham De Moivre. *The doctrine of chances : or, A method of calculating the probabilities of events in play*, volume 200. Chelsea Publishing Company, Incorporated, 1756.