Lycée Pierre-Gilles de Gennes 2024-2025

Mathématiques – TD15

Théorèmes limites en probabilité

1 Autour de la loi des grands nombres

Correction de l'exercice 1.

1. Soit $n \geq 1$, la variable aléatoire Y_n est à valeurs dans $\{0,1\}$, elle suit donc une loi de Bernoulli et son paramètre de succès est
$$p_Y = \mathbb{P}(Y_n = 1) = \mathbb{E}(Y_n) = \mathbb{E}(X_n X_{n+1}).$$
Cette dernière quantité se simplifie, par indépendance des (X_i) en
$$p_Y = \mathbb{E}(X_n)\mathbb{E}(X_{n+1}) = p^2.$$
On a donc $\mathbb{E}(Y_n) = p^2$ et $\mathbb{V}(Y_n) = p^2(1-p^2)$.
2. Soit $(i,j) \in \mathbb{N}^*$:
   - si $i = j$, $\operatorname{Cov}(Y_i, Y_j) = \mathbb{V}(Y_i) = p^2(1-p^2)$;
   - si $|i-j| > 1$, Y_i et Y_j sont indépendantes (c'est une application du lemme des coalitions !) et donc :
   $$\operatorname{Cov}(Y_i, Y_j) = 0;$$
   - si $|i-j| = 1$, par exemple $j = i+1$ (l'autre cas est symétrique, Y_i et Y_j sont a priori pas indépendantes) on a
   $$Y_iY_j = Y_iY_{i+1} = X_iX_{i+1}X_{i+1}X_{i+2} = X_iX_{i+1}^2X_{i+2} = X_iX_{i+1}X_{i+2}$$
   car $X_{i+1}^2 = X_{i+1}$.
   Par indépendance des (X_i) et la formule de Koenig-Huuygens :
   $$\operatorname{Cov}(Y_i, Y_j) = \mathbb{E}(Y_iY_j) - p^4 = p^3 - p^4 = p^3(1-p).$$
3. Soit $S_n = \dfrac{Y_1 + \cdots + Y_n}{n}$ pour $n \geq 1$. On a par linéarité de l'espérance et caractère quadratique de la variance, que
$$\mathbb{E}(S_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(Y_i) = p^2$$
et
$$\mathbb{V}(S_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathbb{V}(Y_i) + 2\sum_{1\leq i<j\leq n}\operatorname{Cov}(Y_i,Y_j)\right) = \frac{1}{n^2}\left(np^2(1-p^2) + 2(n-1)p^3(1-p)\right)$$
Il est clair que $\mathbb{V}(S_n) \xrightarrow[n\to+\infty]{} 0$ et donc, par Bienaymé-Techbycheff, si $\epsilon > 0$ est fixé,
$$0 \leq \mathbb{P}(|S_n - p^2| > \epsilon) \leq \frac{\mathbb{V}(S_n)}{\epsilon^2} \xrightarrow[n\to+\infty]{} 0.$$

Le théorème des gendarmes permet de conclure à

$$\mathbb{P}(|S_n - p^2| > \epsilon) \xrightarrow[n\to+\infty]{} 0.$$

Remarque : on ne peut pas directement appliquer la LGN à la suite $(Y_n)_n$ car elle n'est pas mutuellement indépendante.

Correction de l'exercice 2.

1. On reconnaît que S_n suit une loi binomiale $\mathcal{B}(n,p)$ car les tirages sont indépendants. On compte le nombre de succès dans n tirages indépendants, un succès étant le tirage d'un 6, avec probabilité $p = \dfrac{1}{6}$.

   Plus précisément, si on nomme X_k la variable de Bernoulli de paramètre $\dfrac{1}{6}$ valant 1 si un 6 est tiré au tirage k, 0 sinon, on a

$$S_n = \sum_{k=1}^{n} X_k.$$

   — Pour l'espérance, on a

$$\mathbb{E}(Y_n) = \mathbb{E}(e^{\frac{1}{n}S_n}) = \mathbb{E}(e^{\frac{1}{n}\sum_{k=1}^n X_k}) = \mathbb{E}(\prod_{k=1}^{n} e^{\frac{1}{n}X_k}) \overset{\text{indep.}}{=} \prod_{k=1}^{n} \mathbb{E}(e^{\frac{1}{n}X_k}) = (1-p+pe^{\frac{1}{n}})^n$$

   — Pour la variance, on a, sur la même trame,

$$\mathbb{E}(Y_n^2) = \mathbb{E}(e^{\frac{2}{n}S_n}) = \mathbb{E}(e^{\frac{2}{n}\sum_{k=1}^n X_k}) = \mathbb{E}(\prod_{k=1}^{n} e^{\frac{2}{n}X_k}) \overset{\text{indep.}}{=} \prod_{k=1}^{n} \mathbb{E}(e^{\frac{2}{n}X_k}) = (1-p+pe^{\frac{2}{n}})^n$$

   et donc, par Koenig-Huyghens :

$$\mathbb{V}(Y_n) = \mathbb{E}(Y_n^2) - \mathbb{E}(Y_n)^2 = (1-p+pe^{\frac{2}{n}})^n - (1-p+pe^{\frac{1}{n}})^{2n}.$$

   Remarque : on peut faire ces calculs sans avoir recours à cette décomposition de S_n, en utilisant la formule de transfert pour la loi binomiale et le binôme de Newton. Par exemple, pour l'espérance, on a :

$$\mathbb{E}(Y_n) = \mathbb{E}(e^{\frac{1}{n}S_n}) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{\frac{k}{n}} = \sum_{k=0}^{n} \binom{n}{k} (e^{\frac{1}{n}}p)^k (1-p)^{n-k} \overset{\text{(Newton)}}{=} \left(1-p+pe^{\frac{1}{n}}\right)^n.$$

2. Il s'agit de démontrer que

$$\mathbb{E}(Y_n) = (1-p+pe^{\frac{1}{n}})^n \xrightarrow[n\to+\infty]{} e^p.$$

   La difficulté dans l'expression $(1-p+pe^{\frac{1}{n}})^n$ est que $(1-p+pe^{\frac{1}{n}}) \to 1$ et $n \to +\infty$. On a donc affaire à une indétermination du type "$1^{+\infty}$".

   En passant sous forme exponentielle et en utilisant les DL

$$e^x - 1 \underset{x\to 0}{\sim} x \quad \text{et} \quad \ln(1+x) \underset{x\to 0}{\sim} (x)$$

on a :

$$\mathbb{E}(Y_n) = (1-p+pe^{\frac{1}{n}})^n = \exp\left(n\ln(1+p(e^{\frac{1}{n}}-1))\right)$$

où

$$n\ln(1+p(e^{\frac{1}{n}}-1)) \underset{n\to+\infty}{\sim} np(e^{\frac{1}{n}}-1) \underset{n\to+\infty}{\sim} p.$$

Donc

$$\mathbb{E}(Y_n) \xrightarrow[n\to+\infty]{} e^p.$$

De la même manière (remplacer les $\dfrac{1}{n}$ par $\dfrac{2}{n}$), on a

$$\mathbb{E}(Y_n^2) \xrightarrow[n\to+\infty]{} e^{2p}$$

et donc

$$\mathbb{V}(Y_n) = \mathbb{E}(Y_n^2) - \mathbb{E}(Y_n)^2 \xrightarrow[n\to+\infty]{} e^{2p} - (e^p)^2 = 0.$$

Finalement, soit $\epsilon > 0$, on a pour tout $n \in \mathbb{N}$:

$$|Y_n - e^p| \leq |Y_n - \mathbb{E}(Y_n)| + |\mathbb{E}(Y_n) - e^p|$$

Soit n_0 tel que $\forall n \geq n_0$, $|\mathbb{E}(Y_n) - e^p| < \dfrac{1}{2}\epsilon$.
En raisonnant par contraposée, on a alors pour $n \geq n_0$

$$|Y_n - e^p| \geq \epsilon \Rightarrow |Y_n - \mathbb{E}(Y_n)| > \frac{1}{2}\epsilon$$

et donc l'inégalité de probabilités :

$$0 \leq \mathbb{P}(|Y_n - e^p| \geq \epsilon) \leq \mathbb{P}(|Y_n - \mathbb{E}(Y_n)| > \frac{1}{2}\epsilon)$$

et enfin, par l'inégalité de Bienaymé-Tchebycheff, on a

$$0 \leq \mathbb{P}(|Y_n - e^p| \geq \epsilon) \leq \frac{\mathbb{V}(Y_n)}{(\frac{1}{2}\epsilon)^2} \xrightarrow[n+\infty]{} 0.$$

En passant au complémentaire,

$$\mathbb{P}(|Y_n - e^p| < \epsilon) \xrightarrow[n+\infty]{} 1.$$

Correction de l'exercice 3.

1. On remarque que :

$$\begin{aligned}
\Sigma_n^2 &= \frac{1}{N-1}\left(\sum_{i=1}^N X_i^2 - 2M_N\sum_{i=1}^N X_i + NM_n^2\right) \\
&= \frac{1}{N-1}\left(\sum_{i=1}^N X_i^2 - 2NM_N^2 + NM_n^2\right) \\
&= \frac{1}{N-1}\left(\sum_{i=1}^N X_i^2 - NM_N^2\right).
\end{aligned}$$

Or, pour tout $i \in [\![1, N]\!]$, X_i possède un moment d'ordre 2 donc X_i^2 possède une espérance. Par linéarité, on en déduit que $\sum_{i=1}^{N} X_i^2$ possède une espérance donnée par :

$$\mathbb{E}(\sum_{i=1}^{N} X_i^2) = \sum_{i=1}^{N} E(X_i^2) = N\mathbb{E}(X^2).$$

Par ailleurs, on a :

$$M_N^2 = \frac{1}{N^2}\left(\sum_{i=1}^{N} X_i^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{i-1} X_iX_j\right).$$

Par indépendance, pour tout $i \neq j$ la variable X_iX_j possède une espérance donnée par :

$$\mathbb{E}(X_iX_j) = \mathbb{E}(X_i)\mathbb{E}(X_j) = \mu^2.$$

Par linéarité, on en déduit que M_N^2 possède une espérance donnée par :

$$\begin{aligned}
\mathbb{E}(M_N^2) &= \frac{1}{N^2}\left(\sum_{i=1}^{N} \mathbb{E}(X_i^2) + 2\sum_{i=1}^{N}\sum_{j=1}^{i-1} \mathbb{E}(X_iX_j)\right) \\
&= \frac{1}{N^2}\left(N\mathbb{E}(X^2) + 2\sum_{i=1}^{N}\sum_{j=1}^{i-1} \mu^2\right) \\
&= \frac{1}{N^2}\left(N\mathbb{E}(X^2) + 2\sum_{i=1}^{N}(i-1)\mu^2\right) \\
&= \frac{1}{N^2}\left(N\mathbb{E}(X^2) + N(N-1)\mu^2\right).
\end{aligned}$$

Finalement, par linéarité, Σ_N^2 possède une espérance donnée par :

$$\begin{aligned}
\mathbb{E}(\Sigma_N^2) &= \frac{1}{N-1}\left(\mathbb{E}(\sum_{i=1}^{N} X_i^2) - N\mathbb{E}(M_N^2)\right) \\
&= \frac{1}{N-1}\left(N\mathbb{E}(X^2) - \frac{1}{N}(N\mathbb{E}(X^2) + N(N-1)\mu^2)\right) \\
&= \mathbb{E}(X^2) - \mu^2 \\
&= \mathbb{V}(X).
\end{aligned}$$

2. Soit $\epsilon > 0$, par l'inégalité de Bienaymé-Tchebycheff, on a

$$\mathbb{P}(|\Sigma_N^2 - \mathbb{E}(\Sigma_N^2)| > \epsilon) \leq \frac{\mathbb{V}(\Sigma_N^2)}{\epsilon^2}$$

et donc par la question précédente et la formule admise, il vient, pour tout $N \geq 3$,

$$0 \leq \mathbb{P}(|\Sigma_N^2 - \mathbb{V}(X))| > \epsilon) \leq \frac{\mu_4}{N\epsilon^2}$$

Comme $\frac{\mu_4}{N\epsilon^2} \xrightarrow[N\to+\infty]{} 0$, par le théorème des gendarmes, il vient que

$$\mathbb{P}(|\Sigma_N^2 - \mathbb{V}(X))| > \epsilon) \xrightarrow[N\to+\infty]{} 0.$$

3. Le calcul est assez pénible. On a, en posant $\mathbb{V}(X)^2 = \sigma^4$:

$$\mathbb{V}(\Sigma_N^2) = \mathbb{E}(\Sigma_N^4) - \mathbb{E}(\Sigma_N^2)^2 = \mathbb{E}(\Sigma_N^4) - \sigma^4$$

et on peut supposer (invariance par centrage) que $\mathbb{E}(X) = \mathbb{E}(X_n) = 0$. Par ailleurs,

$$\begin{aligned}
\mathbb{E}(\Sigma_N^4) &= \left(\frac{N}{N-1}\right)^2 \mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N (X_n - M_N)^2\right)^2\right) \\
&= \left(\frac{N}{N-1}\right)^2 \mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N X_n^2 - M_N^2\right)^2\right) \\
&= \left(\frac{N}{N-1}\right)^2 \left[\mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N X_n^2\right)^2\right) - 2\mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N X_n^2\right) M_N^2\right) + \mathbb{E}\left(M_N^4\right)\right].
\end{aligned}$$

On a

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N X_n^2\right)^2\right) &= \frac{1}{N^2}\left[\sum_{n=1}^N \mathbb{E}(X_n^4) + 2\sum_{1\leq m<n\leq N} \mathbb{E}(X_n^2 X_m^2)\right] \\
&= \frac{1}{N^2}\left[\sum_{n=1}^N m_4 + 2\sum_{1\leq m<n\leq N} \sigma^4\right] \\
&= \frac{1}{N} m_4 + \frac{N-1}{N}\sigma^4.
\end{aligned}$$

D'autre part :

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{1}{N}\sum_{n=1}^N X_n^2\right) M_N^2\right) &= \frac{1}{N^2}\mathbb{E}\left(\left(\sum_{m=1}^N X_m\right)^2 X_1^2\right) \\
&= \frac{1}{N^2}\mathbb{E}\left(\sum_{m=1}^N X_m^2 X_1^2\right) + \frac{2}{N^2}\sum_{1\leq n<m\leq N} \mathbb{E}\left(X_m X_n X_1^2\right) \\
&= \frac{1}{N^2}\mathbb{E}\left(X_1^4\right) + \frac{1}{N^2}\sum_{m=2}^N \mathbb{E}\left(X_m^2 X_1^2\right) \\
&= \frac{1}{N^2} m_4 + \frac{N-1}{N^2}\sigma^4
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{E}\left(M_N^4\right) &= \frac{1}{N^4}\mathbb{E}\left(\left(\sum_{n=1}^{N} X_n\right)^4\right) \\
&= \frac{1}{N^4}\left[\sum_{\#\{i,j,k,\ell\}=1} \mathbb{E}(X_iX_jX_kX_\ell) + \sum_{\#\{i,j,k,\ell\}=2} \mathbb{E}(X_iX_jX_kX_\ell) + \underbrace{\sum_{\#\{i,j,k,\ell\}=3} \mathbb{E}(X_iX_jX_kX_\ell)}_{=0}\right. \\
&\qquad \left. + \underbrace{\sum_{\#\{i,j,k,\ell\}=4} \mathbb{E}(X_iX_jX_kX_\ell)}_{=0}\right] \\
&= \frac{1}{N^3}m_4 + \frac{1}{N^4}\binom{4}{2}\binom{N}{2}\sigma^4 \\
&= \frac{1}{N^3}m_4 + \frac{3(N-1)}{N^3}\sigma^4
\end{aligned}$$

On en déduit finalement :

$$\begin{aligned}
\mathbb{E}(\Sigma_N^4) &= \frac{N^2}{(N-1)^2}\left[\left(\frac{1}{N}m_4 + \frac{N-1}{N}\sigma^4\right) - 2\left(\frac{1}{N^2}m_4 + \frac{N-1}{N^2}\sigma^4\right) + \left(\frac{1}{N^3}m_4 + \frac{3(N-1)}{N^3}\sigma^4\right)\right] \\
&= \frac{1}{N}m_4 + \frac{N^2-2N+3}{N(N-1)}\sigma^4 \\
&= \frac{1}{N}m_4 + \frac{-N+3}{N(N-1)}\sigma^4 + \sigma^4
\end{aligned}$$

et finalement, comme attendu,

$$\mathbb{V}(\Sigma_N^2) = \frac{1}{N}m_4 - \frac{N-3}{N(N-1)}\sigma^4$$

2 Convergence en loi

Correction de l'exercice 4.

1. La fonction de répartition des X_n est la fonction F définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, \quad F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}.$$

   (a) Soient $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$. On a par indépendance :

$$\begin{aligned}
\mathbb{P}(\max(X_1,\dots,X_n) \leq x) = \mathbb{P}([X_1 \leq x] \cap \dots \cap [X_n \leq x]) &= \mathbb{P}(X_1 \leq x) \times \dots \times \mathbb{P}(X_n \leq x) \\
&= F(x)^n \\
&= \begin{cases} 0 & \text{si } x < 0 \\ x^n & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}.
\end{aligned}$$

On en déduit donc :

$$\lim_{n\to+\infty} \mathbb{P}(\max(X_1,\ldots,X_n) \leq x) = \lim_{n\to+\infty} \begin{cases} 0 & \text{si } x < 0 \\ x^n & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases} = \begin{cases} 0 & \text{si } x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Soit X une variable aléatoire suivant la loi certaine de paramètre 1. Alors sa fonction de répartition est la fonction F_X donnée par :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}.$$

Ainsi :

$$\forall x \in \mathbb{R}, \quad \lim_{n\to+\infty} F_{\max(X_1,\ldots,X_n)}(x) = F_X(x).$$

Donc la suite $(\max(X_1,\ldots,X_n))_{n\in\mathbb{N}^*}$ converge en loi vers X.

(b) Soient $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$. On a par indépendance :

$$\begin{aligned} \mathbb{P}(\min(X_1,\ldots,X_n) > x) = \mathbb{P}([X_1 > x] \cap \cdots \cap [X_n > x]) &= \mathbb{P}(X_1 > x) \times \cdots \times \mathbb{P}(X_n > x) \\ &= (1 - F(x))^n. \end{aligned}$$

Ainsi :

$$\begin{aligned} F_{\min(X_1,\ldots,X_n)}(x) = 1 - \mathbb{P}(\min(X_1,\ldots,X_n) > x) &= 1 - (1 - F(x))^n \\ &= \begin{cases} 0 & \text{si } x < 0 \\ 1 - (1-x)^n & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}. \end{aligned}$$

On en déduit donc :

$$\lim_{n\to+\infty} F_{\min(X_1,\ldots,X_n)}(x) = \lim_{n\to+\infty} \begin{cases} 0 & \text{si } x < 0 \\ 1 - (1-x)^n & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases} = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}$$

Soit X une variable aléatoire suivant la loi certaine de paramètre 0. Alors sa fonction de répartition est la fonction F_X donnée par :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}.$$

Ainsi pour tout $x \in \mathbb{R}^*$ c'est-à-dire pour tout x en lequel F_X est continue , on a : $\lim\limits_{n\to+\infty} F_{\min(X_1,\ldots,X_n)}(x) = F_X(x)$. Donc la suite $(\min(X_1,\ldots,X_n))_{n\in\mathbb{N}^*}$ converge en loi vers X.

2. Soit X une variable de loi $\mathcal{B}(e^{-1})$. Comme les variables X_n, $n \in \mathbb{N}^*$, et X sont à valeurs dans $\mathbb{Z}$, on va utiliser la caractérisation de la convergence en loi pour les variables discrètes. Soient $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}$.

$$\mathbb{P}(X_n = k) = \begin{cases} 0 & \text{si } k \notin \{0,1\} \\ \left(1 - \dfrac{1}{n}\right)^n & \text{si } k = 1 \\ 1 - \left(1 - \dfrac{1}{n}\right)^n & \text{si } k = 0 \end{cases}.$$

Or, par équivalent usuel :

$$\ln\left(1-\frac{1}{n}\right) \underset{n\to+\infty}{\sim} -\frac{1}{n}.$$

Par compatibilité avec le quotient on en déduit donc :

$$n\ln\left(1-\frac{1}{n}\right) \underset{n\to+\infty}{\sim} -1.$$

Ainsi, $\lim\limits_{n\to+\infty} n\ln\left(1-\frac{1}{n}\right) = -1$. Par continuité de la fonction exponentielle en 0 :

$$\lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^n = \lim_{n\to+\infty} e^{n\ln\left(1-\frac{1}{n}\right)} = e^{-1}.$$

On en déduit donc :

$$\forall k\in\mathbb{Z},\quad \lim_{n\to+\infty}\mathbb{P}(X_n=k)=\mathbb{P}(X=k).$$

Ainsi $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.

3. La fonction de répartition de X_n est la fonction F_n définie sur $\mathbb{R}$ par :

$$\forall x\in\mathbb{R},\quad F_n(x)=\begin{cases} 0 & \text{si } x<0 \\ \dfrac{x}{1-\frac{1}{n}} & \text{si } x\in\left[0,1-\dfrac{1}{n}\right] \\ 1 & \text{si } x>1-\dfrac{1}{n}\end{cases}.$$

Soit $x\in\mathbb{R}$.

- Si $x<0$ alors pour tout $n\in\mathbb{N}^*$, $F_n(x)=0$. Donc $\lim\limits_{n\to+\infty}F_n(x)=0$.
- Si $x\in[0,1[$ alors il existe un rang N tel que :

$$\forall n\geq N,\quad x\in\left[0,1-\frac{1}{n}\right[.$$

  Donc pour tout $n\geq N$, $F_n(x)=\dfrac{x}{1-\frac{1}{n}}$ et $\lim\limits_{n\to+\infty}F_n(x)=x$.
- Si $x\geq 1$ alors pour tout $n\in\mathbb{N}^*$, $F_n(x)=1$ donc $\lim\limits_{n\to+\infty}F_n(x)=1$.

Donc :

$$\forall x\in\mathbb{R},\quad \lim_{n\to+\infty}F_n(x)=\begin{cases}0 & \text{si } x<0\\ x & \text{si } x\in[0,1[\\ 1 & \text{si } x\geq 1\end{cases}$$

Ainsi, si X suit la loi uniforme sur $[0,1]$ alors :

$$\forall x\in\mathbb{R},\quad \lim_{n\to+\infty}F_n(x)=F_X(x).$$

Donc la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.

Correction de l'exercice 5.

1. Soit $n \in \mathbb{N}^*$. On a :

$$\forall x \in \mathbb{R}, \quad F_{X_n}(x) = \int_{-\infty}^{x} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-nt^2} dt.$$

   En effectuant le changement de variable $s = \sqrt{n}t$ on obtient pour tout $x \in \mathbb{R}$:

$$F_{X_n}(x) = \int_{-\infty}^{\sqrt{n}x} \frac{1}{\sqrt{2\pi}} e^{-s^2} ds = \Phi(\sqrt{n}x).$$

2. Soit $x \in \mathbb{R}$.
   - Si $x > 0$ alors $\lim\limits_{n\to+\infty} \sqrt{n}x = +\infty$. Or, $\lim\limits_{t\to+\infty} \Phi(t) = 1$ donc par composition des limites, on trouve :

$$\lim_{n\to+\infty} F_{X_n}(x) = \lim_{n\to+\infty} \Phi(\sqrt{n}x) = 1.$$

   - Si $x < 0$ alors $\lim\limits_{n\to+\infty} \sqrt{n}x = -\infty$. Or, $\lim\limits_{t\to-\infty} \Phi(t) = 0$ donc par composition des limites, on trouve :

$$\lim_{n\to+\infty} F_{X_n}(x) = \lim_{n\to+\infty} \Phi(\sqrt{n}x) = 0.$$

3. Soit X suivant une loi certaine de paramètre 0. D'après la question précédente, pour tout $x \in \mathbb{R}^*$, c'est-à-dire pour tout réel x en lequel F_X est continue, on a :

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x).$$

   La suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.

Correction de l'exercice 6.

1. Soit $x \in \mathbb{R}$.

$$F_{Y_n}(x) = \mathbb{P}(X_n \leq nx) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{\lfloor nx \rfloor + 1}{n+1} & \text{si } x \in [0, 1[\\ 1 & \text{si } x \geq 1 \end{cases}.$$

2. Si $x < 0$ alors $\lim\limits_{n\to+\infty} F_{Y_n}(x) = 0$.
   Si $x > 1$ alors $\lim\limits_{n\to+\infty} F_{Y_n}(x) = 1$.
   Si $x \in [0, 1]$. On a

$$\frac{nx+1}{n+1} \leq \frac{\lfloor nx \rfloor + 1}{n+1} \leq \frac{nx+2}{n+1}.$$

   Donc par encadrement on obtient $\lim\limits_{n\to+\infty} F_{Y_n}(x) = x$.

3. D'après la question précédente :

$$\forall x \in \mathbb{R}, \quad \lim_{n\to+\infty} F_{Y_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0, 1] \\ 1 & \text{si } x > 1 \end{cases} = F_X(x)$$

   où $X \hookrightarrow \mathcal{U}([0, 1])$. Ainsi $Y_n \xrightarrow{\mathcal{L}} X$ où $X \hookrightarrow \mathcal{U}([0, 1])$.

Correction de l'exercice 7.

1. Pour tout $k \in \mathbb{N}$ on a

$$\mathbb{P}(X_n = k) = \begin{cases} 1 - \left(\dfrac{1}{2} + \dfrac{1}{2n}\right) & \text{si } k = 0 \\ \dfrac{1}{2} + \dfrac{1}{2n} & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases}$$

donc

$$\lim_{n \to +\infty} \mathbb{P}(X_n = k) = \begin{cases} \dfrac{1}{2} & \text{si } k = 0 \\ \dfrac{1}{2} & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases}.$$

Ainsi $X_n \xrightarrow{\mathcal{L}} X$ où $X \hookrightarrow \mathcal{B}\left(\dfrac{1}{2}\right)$.

2. Pour tout $n \in \mathbb{N}^*$, la variable X_n est à valeurs dans $\mathbb{N}$ de même qu'une variable X de loi certaine de paramètre 0. On va donc utiliser la caractérisation de la convergence en loi pour les variables aléatoires à valeurs entières.
Soit $k \in \mathbb{N}$.

— Si $k = 0$ alors pour tout $n \in \mathbb{N}^*$ on a $\mathbb{P}(X_n = k) = e^{-\frac{1}{n}}$. Donc on a

$$\lim_{n \to +\infty} \mathbb{P}(X_n = 0) = 1 = P(X = 0).$$

— Si $k > 0$ alors pour tout $n \in \mathbb{N}^*$ on a $\mathbb{P}(X_n = k) = \dfrac{1}{n^k k!} e^{-\frac{1}{n}}$. Donc on a

$$\lim_{n \to +\infty} \mathbb{P}(X_n = k) = 0 = \mathbb{P}(X = k).$$

Ainsi, pour tout $k \in \mathbb{N}$ on a : $\lim\limits_{n \to +\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$.

Donc $X_n \xrightarrow{\mathcal{L}} X$.

3 Théorème central limite

Correction de l'exercice 8.

1. Chacune des variables X, Y, Z compte le nombre de succès d'une répétition de 10 épreuves de Bernoulli indépendantes de paramètre $\dfrac{1}{2}$.
Par indépendance de X, Y et Z, $X + Y + Z$ compte donc le nombre de succès d'une répétition de 30 épreuves de Bernoulli indépendantes de paramètre $\dfrac{1}{2}$ donc suit la loi $\mathcal{B}\left(30, \dfrac{1}{2}\right)$.

2. Soit $(X_n)_{n \in \mathbb{N}^*}$ des variables aléatoires suivant la loi $\mathcal{B}\left(\dfrac{1}{2}\right)$. D'après le TCL on sait que $\left(\dfrac{\sum_{i=1}^n X_i - \frac{n}{2}}{\sqrt{\frac{n}{4}}}\right)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire X suivant une

loi $\mathcal{N}(0,1)$. Ainsi, on a l'approximation suivant pour tout $-\infty \leq a \leq b \leq +\infty$:

$$\mathbb{P}\left(a \leq 2\frac{\sum_{i=1}^{30} X_i - 15}{\sqrt{30}} \leq b\right) \simeq \Phi(b) - \Phi(a).$$

On a :

$$\mathbb{P}\left(a \leq 2\frac{\sum_{i=1}^{30} X_i - 15}{\sqrt{30}} \leq b\right) = \mathbb{P}\left(a\frac{\sqrt{30}}{2} + 15 \leq \sum_{i=1}^{30} X_i \leq b\frac{\sqrt{30}}{2} + 15\right).$$

Or, $\displaystyle\sum_{i=1}^{30} X_i$ et $X+Y+Z$ ont la même loi donc :

$$\begin{aligned}\mathbb{P}\left(a \leq 2\frac{\sum_{i=1}^{30} X_i - 15}{\sqrt{30}} \leq b\right) &= \mathbb{P}\left(a\frac{\sqrt{30}}{2} + 15 \leq X+Y+Z \leq b\frac{\sqrt{30}}{2} + 15\right)\\ &= \mathbb{P}\left(a\frac{\sqrt{30}}{6} + 5 \leq R \leq b\frac{\sqrt{30}}{6} + 5\right).\end{aligned}$$

En prenant $a = -\sqrt{\dfrac{6}{5}}$ et $b = +\infty$ on obtient :

$$\mathbb{P}(R \geq 4) = \mathbb{P}\left(a \leq 2\frac{\sum_{i=1}^{30} X_i - 15}{\sqrt{30}}\right) \simeq 1 - \Phi(a) = \Phi(-a) \simeq 0.86.$$

Correction de l'exercice 9. D'après le TCL on sait que $\left(\dfrac{\sum_{i=1}^{n} T_i - 3n}{\sqrt{\frac{n}{2}}}\right)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X suivant une loi $\mathcal{N}(0,1)$. Ainsi, on a l'approximation suivant pour tout $-\infty \leq a \leq b \leq +\infty$:

$$\mathbb{P}\left(a < \frac{\sum_{i=1}^{1000} T_i - 3000}{\sqrt{500}} \leq b\right) \simeq \Phi(b) - \Phi(a).$$

Or on a :

$$\begin{aligned}\mathbb{P}\left(a < \frac{\sum_{i=1}^{1000} T_i - 3000}{\sqrt{500}} \leq b\right) &= \mathbb{P}\left(a\sqrt{500} + 3000 < \sum_{i=1}^{1000} T_i \leq b\sqrt{500} + 3000\right)\\ &= \mathbb{P}\left(a\frac{\sqrt{5}}{100} + 3 < S \leq b\frac{\sqrt{5}}{100} + 3\right).\end{aligned}$$

En prenant $a = -\sqrt{5}$ et $b = \sqrt{5}$ on obtient :

$$\mathbb{P}(2.95 < S \leq 3.05) \simeq \Phi(\sqrt{5}) - \Phi(-\sqrt{5}) = 2\Phi(\sqrt{5}) - 1 \simeq 0.956.$$

Correction de l'exercice 10.

1. Par stabilité par somme des lois de Poisson, S_n étant la somme de n variables aléatoires indépendantes de loi $\mathcal{P}(1)$, on sait que S_n suit la loi $\mathcal{P}(n)$.
   En conséquence, $\mathbb{E}(S_n) = n = \mathbb{V}(S_n)$.

2. Soit $n \in \mathbb{N}^*$.
$$u_n = \sum_{i=1}^{n} e^{-n}\frac{n^i}{i!} = \sum_{i=1}^{n} \mathbb{P}(S_n = i) = \mathbb{P}\left(\bigcup_{i=1}^{n}[S_n = i]\right) = \mathbb{P}(1 \leq S_n \leq n).$$
Comme S_n est à valeurs dans $\mathbb{N}$ on en déduit :
$$u_n = \mathbb{P}(S_n \leq n) - \mathbb{P}(S_n = 0) = \mathbb{P}(S_n \leq n) - e^{-n}.$$

3. D'après le TCL, on sait que :
$$\lim_{n\to+\infty} \mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = \Phi(0) = \frac{1}{2},$$
où Φ est la fonction de répartition d'une variable aléatoire de loi normale centrée réduite. Or, pour tout $n \in \mathbb{N}^*$ on a :
$$\mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = \mathbb{P}(S_n \leq n).$$
Ainsi, on en déduit :
$$\lim_{n\to+\infty} \mathbb{P}\left(S_n \leq n\right) = \frac{1}{2}.$$
Enfin, comme $\lim\limits_{n\to+\infty} e^{-n} = 0$ on trouve :
$$\lim_{n\to+\infty} u_n = \lim_{n\to+\infty} \left(\mathbb{P}\left(S_n \leq n\right) - e^{-n}\right) = \frac{1}{2}.$$

Correction de l'exercice 11. D'après le TCL on sait que $\left(\frac{S_n - \frac{1}{2}n}{\sqrt{\frac{n}{12}}}\right)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X suivant une loi $\mathcal{N}(0,1)$.
Ainsi, pour tout $-\infty \leq a \leq b \leq +\infty$:
$$\lim_{n\to+\infty} \mathbb{P}\left(a < \frac{S_n - \frac{1}{2}n}{\sqrt{\frac{n}{12}}} \leq b\right) = \Phi(b) - \Phi(a).$$
Or on a :
$$\mathbb{P}\left(a < \frac{S_n - \frac{1}{2}n}{\sqrt{\frac{n}{12}}} \leq b\right) = \mathbb{P}\left(a\sqrt{\frac{n}{12}} + \frac{n}{2} < S_n \leq b\sqrt{\frac{n}{12}} + \frac{n}{2}\right).$$
En prenant $a = -\sqrt{12}$ et $b = \sqrt{12}$ on obtient :
$$\mathbb{P}\left(a < \frac{S_n - \frac{1}{2}n}{\sqrt{\frac{n}{12}}} \leq b\right) = \mathbb{P}\left(\frac{n}{2} - \sqrt{n} < S_n \leq \frac{n}{2} + \sqrt{n}\right)$$
Donc la limite existe et vaut :
$$\Phi(\sqrt{12}) - \Phi(-\sqrt{12}) = 2\Phi(\sqrt{12}) - 1$$
où Φ est la fonction de répartition de la loi normale centrée réduite.

Correction de l'exercice 12.

1. Montrons cela par récurrence sur $n \in \mathbb{N}^*$.

   — **Initialisation :** c'est l'hypothèse faite sur X.

   — **Hérédité :** soit $n \in \mathbb{N}^*$ et supposons que pour tout famille $(X_k)_{k\in[\![0,2^n]\!]}$ de variables mutuellement indépendantes ayant la loi de X, la variable $\dfrac{\sum_{k=1}^{2^n} X_k}{\sqrt{2^n}}$ suit la loi de X.

   On considère $(X_k)_{k\in[\![0,2^{n+1}]\!]}$ et on souhaite montrer que $\dfrac{\sum_{k=1}^{2^{n+1}} X_k}{\sqrt{2^{n+1}}}$ suit la loi de X.

   On applique l'hypothèse de récurrence à $(X_k)_{k\in[\![0,2^n]\!]}$ et $(X_k)_{k\in[\![2^n+1,2^{n+1}]\!]}$ pour obtenir que :
$$X_1 = \frac{\sum_{k=1}^{2^n} X_k}{\sqrt{2^n}} \quad \text{et} \quad X_2 = \frac{\sum_{k=2^n+1}^{2^{n+1}} X_k}{\sqrt{2^n}}$$
   suivent la loi de X.

   Mais alors, d'après l'énoncé :
$$\frac{\sum_{k=1}^{2^{n+1}} X_k}{\sqrt{2^{n+1}}} = \frac{X_1 + X_2}{\sqrt{2}}$$
   suit la loi de X. Cela conclut l'hérédité

   — **Conclusion :** par récurrence, le résultat souhaité est démontré.

2. Soit $x \in \mathbb{R}$. D'après la question précédente, pour tout $n \in \mathbb{N}^*$, on a :
$$F_X(x) = \mathbb{P}\left(\frac{\sum_{k=1}^{2^n} X_k}{\sqrt{2^n}} \leq x\right) = \mathbb{P}\left(\frac{\sum_{k=1}^{2^n} X_k}{\sigma\sqrt{2^n}} \leq \frac{x}{\sigma}\right)$$
$$\xrightarrow[n\to+\infty]{} F_N\left(\frac{x}{\sigma}\right)$$
   où $N \hookrightarrow \mathcal{N}(0,1)$ d'après le TCL.

   Ainsi :
$$\forall x \in \mathbb{R}, \quad F_X(x) = F_{\sigma N}(x)$$
   donc X suit la loi de σN c'est-à-dire la loi $\mathcal{N}(0,\sigma^2)$.

4 Tests

Correction de l'exercice 13. On pose $n = 746977$ et on considère chaque naissance comme une variable aléatoire X_i de même loi que X.
On peut faire l'hypothèse que $(X_1,\dots,X_n)$ sont indépendantes.
D'après le TCL, on a pour $t > 0$:
$$\mathbb{P}\left(-t \leq \sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \leq t\right) \simeq \Phi(t) - \Phi(-t) = 2\Phi(t) - 1$$
où Φ est la fonction de répartition de la loi normale centrée réduite.

On veut $2\Phi(t) - 1 = 0.95$ soit $t \simeq 1.96$ et on obtient :

$$\mathbb{P}\left(-t \leq \sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \leq t\right) \simeq 0.95.$$

Sous l'hypothèse $H_0 : p = \dfrac{1}{2}$ on a donc :

$$\mathbb{P}\left(-t \leq 2\sqrt{n}\left(\overline{X}_n - \frac{1}{2}\right) \leq t\right) \simeq 0.95.$$

Or ici :

$$\overline{X}_n = \frac{364392}{n}$$

ce qui donne :

$$2\sqrt{n}\left(\overline{X}_n - \frac{1}{2}\right) \simeq -21.0499.$$

On ne tombe donc pas dans l'intervalle $[-t, t]$ de non rejet et on rejette donc l'hypothèse.

Correction de l'exercice 14. La variable X du dosage de l'activité de l'enzyme de l'individu testé suit une loi normale $\mathcal{N}(\mu, \sigma^2)$ de paramètres inconnus et on souhaite tester l'hypothèse :

$$H_0 : \mu = 10.7.$$

On ne connait pas la variance donc on va utiliser la valeur approchée S_n^2 (ici $n = 8$ car on a 8 mesures) :

$$M_n = 7.95 \quad ; \quad S_n \simeq 3.29.$$

D'après le TCL on sait que sous l'hypothèse H_0 :

$$\lim_{n\to+\infty} \mathbb{P}(-1.96 \leq \sqrt{n}\frac{M_n - 10.7}{S_n} \leq 1.96) \simeq 0.95.$$

Ici, on a :

$$\sqrt{n}\frac{|M_n - 10.7|}{S_n} \simeq 2.364173756627205$$

donc on rejette l'hypothèse H_0.
Attention cependant, ici l'échantillon est trop petit pour que le TCL soit réellement pertinent et il faudrait plutôt utiliser un test de Student (voir TP7). La variable $\sqrt{n}\dfrac{M_n - 10.7}{S_n}$ suit la loi de Student à 7 degrés.
Avec la commande Python, on trouve la valeur t telle que sous l'hypothèse H_0 :

$$\mathbb{P}(t \leq \sqrt{n}\frac{M_n - 10.7}{S_n} \leq t) \simeq 0.95.$$

On obtient :

`stats.t.ppf(1-0.05/2,df=7)` $= 2.3646242510102993.$

Ainsi l'hypothèse H_0 n'est pas rejetée avec ce test.