Correction des Travaux dirigés de Probabilités et Statistiques - Convergences, LGN et TCL

Julian Tugaut*

*Si vous trouvez des erreurs de français ou de mathématiques ou bien si vous avez des questions et/ou des suggestions, envoyez-moi un mail à julian.tugaut@univ-st-etienne.fr

Table des matières

Exercice 70

Énoncé

Soit X une variable aléatoire réelle qui suit la loi exponentielle de paramètre 1. On pose $X_n := X\mathbb{1}_{[0;n[}(X) + e^{2n}\mathbb{1}_{[n;+\infty[}(X)$.

1. Vérifier que $(X_n)_n$ converge en probabilité vers X.

2. Calculer $\mathbb{E}(X_n)$.

3. Est-il vrai que $\mathbb{E}(X_n)$ converge vers $\mathbb{E}(X)$?

Correction

Correction du 1)

Par définition, si $X < n$ alors $X_n = X$ c'est-à-dire $\{X < n\} \subset \{X_n = X\}$. Par contraposée, on a $\{X_n \neq X\} \subset \{X \geq n\}$. En particulier, $\mathbb{P}(X_n \neq X) \leq \mathbb{P}(X \geq n)$. Or, $\mathbb{P}(X \geq n) = \int_n^\infty e^{-t}dt = e^{-n}$.

Également, pour tout $\epsilon > 0$, on remarque $\mathbb{P}(|X - X_n| \geq \epsilon) \leq \mathbb{P}(X_n \neq X) \leq e^{-n} \longrightarrow 0$, quand n tend vers l'infini. Conséquemment, $\lim\limits_{n\to\infty} \mathbb{P}(|X - X_n| \geq \epsilon) = 0$ ce qui signifie que X_n converge vers X en probabilité.

Correction du 2)

Pour calculer, on remarque $X_n = f_n(X)$ où f_n est la fonction définie par

$$f_n(x) := \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x < n \\ e^{2n} & \text{si } x \geq n \end{cases}.$$

On a donc

$$\begin{aligned}
\mathbb{E}[X_n] &= \mathbb{E}[f_n(X)] \\
&= \int_{\mathbb{R}} f_n(x)e^{-x}H(x)dx \\
&= \int_0^\infty f_n(x)e^{-x}dx \\
&= \int_0^n xe^{-x}dx + \int_n^\infty e^{2n}e^{-x}dx \\
&= \left[-(x+1)e^{-x}\right]_0^n + e^{2n}\left[-e^{-x}\right]_n^\infty \\
&= 1 - (n+1)e^{-n} + e^n\,.
\end{aligned}$$

On trouve ainsi $\mathbb{E}[X_n] = 1 - (n+1)e^{-n} + e^n$.

Correction du 3)

On remarque $\lim\limits_{n\to\infty} \mathbb{E}[X_n] = +\infty \neq 1 = \mathbb{E}[X]$.

Remarque 1

Si $(X_n)_n$ convergeait dans L^1, on aurait $\lim_{n\to+\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

Pourtant, X comme X_n sont dans L^1. Néanmoins, on ne peut pas majorer uniformément (par rapport à n) $|X_n| = X_n$ par une variable aléatoire dans L^1 si bien que la convergence en probabilité n'implique pas la convergence dans L^1.

Remarque 2

Comme la suite $(X_n)_n$ converge en probabilité vers X, on en déduit qu'elle converge en loi et donc $\mathbb{E}[f(X_n)]$ tend vers $\mathbb{E}[f(X)]$ pour toute fonction f continue et bornée. Ici, comme la fonction $x \mapsto x$ n'est pas bornée, on ne dispose pas de la limite $\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$.

Remarque 3

On pouvait prouver la convergence en probabilité dans la première question en établissant la convergence presque sûre. En effet, pour tout $x \in [0; +\infty[$, il existe n tel que $x < n$. Par exemple, on peut considérer la partie entière de $x + 1$. Ainsi, pour $\omega \in \Lambda := \{\omega \in \Omega \ : \ X(\omega) < +\infty\}$, on a $X_{n(\omega)}(\omega) = X(\omega)$ pour $n(\omega)$ assez grand. L'ensemble Λ étant de probabilité 1, la preuve de la convergence presque sûre s'ensuit. On en déduit la convergence en probabilité.

Exercice 71

Énoncé

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes telles que $\mathbb{P}_{X_n} = \frac{1}{2}\delta_{-\sqrt{n}} + \frac{1}{2}\delta_{\sqrt{n}}$. On note $S_n := \sum_{k=1}^n X_k$.
Démontrer que $\frac{S_n}{n^{\frac{3}{2}}}$ converge vers 0 en probabilité.

Correction

Les variables aléatoires X_k sont centrées, avec $\mathrm{Var}[X_k] = k$. Comme elles sont indépendantes, on a

$$\mathrm{Var}[S_n] = \sum_{k=1}^n \mathrm{Var}[X_k] = \sum_{k=1}^n k = \frac{n(n+1)}{2}\,.$$

Maintenant, d'après l'inégalité de Bienaymé-Tchebychev, on a pour tout $\epsilon > 0$:

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n^{\frac{3}{2}}}\right| \geq \epsilon\right) &\leq \frac{\mathrm{Var}[S_n]}{\epsilon^2 n^3} \\ &\leq \frac{n(n+1)}{2\epsilon^2 n^3} \\ &\leq \frac{1}{n\epsilon^2} \longrightarrow 0\,.\end{aligned}$$

Cela implique que $\frac{S_n}{n^{\frac{3}{2}}}$ converge en probabilité vers 0.

Exercice 72

Énoncé

Soit $(p_n)_n$ une suite de réels dans $]0;1[$ telle que $\lim\limits_{n\infty} np_n = \lambda > 0$. Soit $(X_n)_n$ une suite de variables aléatoires réelles telle que pour tout n, X_n suit la loi binomiale $\mathcal{B}(n, p_n)$. Soit Y une variable aléatoire de loi de Poisson de paramètre λ. Montrer que $(X_n)_n$ converge en loi vers Y.

Correction

On sait que la fonction caractéristique de X_n est

$$\varphi_{X_n}(t) = \left(p_n e^{it} + (1 - p_n)\right)^n .$$

Or, p_n tend vers 0. On réécrit donc comme suit :

$$\varphi_{X_n}(t) = \left(1 + p_n(e^{it} - 1)\right)^n .$$

Il s'agit ici d'une forme indéterminée du type 1^∞. Pour la résoudre, on utilise un développement limité :

$$\log\left\{1 + p_n\left(e^{it} - 1\right)\right\} = p_n\left(e^{it} - 1\right) + o(p_n) .$$

Donc :

$$\varphi_{X_n}(t) = \exp\left\{np_n\left(e^{it} - 1\right) + o(np_n)\right\} \longrightarrow \exp\left\{\lambda\left(e^{it} - 1\right)\right\} = \varphi_Y(t) ,$$

ce qui achève la preuve.

Remarque

On a utilisé la détermination principale du logarithme, ce qui nécessiterait - pour plus de rigueur - de faire de l'analyse complexe. Néanmoins, on peut aussi obtenir la convergence de la fonction caractéristique sans passer par le logarithme complexe.

En effet, on peut écrire

$$\begin{aligned}
\varphi_{X_n}(t) &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p_n^k \left(e^{it} - 1\right)^k \\
&= \sum_{k=0}^{\infty} \frac{n!}{k!(n-k)!} p_n^k \left(e^{it} - 1\right)^k \mathbb{1}_{k \leq n} =: \sum_{k=0}^{\infty} a_n(k) .
\end{aligned}$$

Prouvons maintenant que pour tout $k \in \mathbb{N}$, $\lim_{n\to\infty} a_n(k) = \frac{\lambda^k}{k!}(e^{it} - 1)^k$. Pour cela, on utilise la formule de Stirling : $n!$ est équivalent à $n^n e^{-n}\sqrt{2\pi n}$ quand n tend vers l'infini.

De même, $(n-k)!$ est équivalent, pour k fixé, à $(n-k)^{n-k}e^{-(n-k)}\sqrt{2\pi(n-k)}$. Également, on sait que p_n est équivalent à $\frac{\lambda}{n}$ quand n tend vers l'infini. On peut ainsi dire que $a_n(k)$ est équivalent à

$$\frac{1}{k!}\frac{n^n e^{-n}\sqrt{2\pi n}}{(n-k)^{n-k}e^{-(n-k)}\sqrt{2\pi(n-k)}}\frac{\lambda^k}{n^k}(e^{it}-1)^k = \frac{\lambda^k(e^{it}-1)^k}{k!}\underbrace{\left(\frac{n}{n-k}\right)^n}_{=\left(1-\frac{k}{n}\right)^{-n}\longrightarrow e^k} e^{-k}\underbrace{(n-k)^k\frac{1}{n^k}}_{=\left(1-\frac{k}{n}\right)^k\longrightarrow 1}\underbrace{\sqrt{\frac{n}{n-k}}}_{\longrightarrow 1}$$

$$\longrightarrow \frac{\lambda^k(e^{it}-1)^k}{k!}\,.$$

Pour conclure, il faut montrer l'existence d'une suite b telle que $|a_n(k)| \leq b(k)$ pour tout $n, k \in \mathbb{N}$ et telle que $\sum_{k=0}^{\infty} b(k) < \infty$.

Or, comme $\lim_{n\to\infty} np_n = \lambda$, il existe n_0 tel que si $n \geq n_0$ alors $0 \leq p_n \leq \frac{2\lambda}{n}$. Par conséquent, si n est assez grand, on peut écrire

$$\begin{aligned}
|a_n(k)| &\leq \frac{1}{k!}(2\lambda)^k\frac{n!}{(n-k)!}\frac{1}{n^k}\underbrace{\left|e^{it}-1\right|^k}_{\leq 2^k}\\
&\leq \frac{(4\lambda)^k}{k!}\underbrace{\prod_{p=1}^{k}\left(\frac{n-k+p}{n}\right)}_{\leq 1}\\
&\leq \frac{(4\lambda)^k}{k!} =: b(k)\,.
\end{aligned}$$

Or, $\sum_{k=0}^{\infty} b(k) = e^{4\lambda} < \infty$. Par conséquent, on a la convergence de $\varphi_{X_n}(t) = \sum_{k=0}^{\infty} a_n(k)$ vers $\sum_{k=0}^{\infty}\frac{\lambda^k(e^{it}-1)^k}{k!} = e^{\lambda(e^{it}-1)} = \varphi_Y(t)$.

Autre correction

On peut aussi prouver la convergence de $\mathbb{P}(X_n = k)$ vers $\mathbb{P}(Y = k)$ en utilisant la formule de Stirling. En effet, pour tout $n \geq k$:

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \binom{n}{k}p_n^k(1-p_n)^{n-k}\\
&= \frac{n!}{k!(n-k)!}\left(\frac{p_n}{1-p_n}\right)^k(1-p_n)^n\\
&\sim \frac{1}{k!}\frac{n^n e^{-n}\sqrt{2\pi n}}{(n-k)^{n-k}e^{-n}e^k\sqrt{2\pi(n-k)}}\left(\frac{p_n}{1-p_n}\right)^k(1-p_n)^n\\
&\sim \frac{(n-k)^k p_n^k}{k!}\underbrace{\left(1-\frac{k}{n}\right)^{-n}}_{\longrightarrow e^k} e^{-k}\sqrt{\frac{n}{n-k}}\frac{1}{(1-p_n)^k}\underbrace{(1-p_n)^n}_{\longrightarrow e^{-\lambda}}\\
&\longrightarrow \frac{\lambda^k}{k!}e^{-\lambda} = \mathbb{P}(Y = k)\,.
\end{aligned}$$

On en déduit la convergence de $F_{X_n}(t)$ vers $F_Y(t)$ en tout point de continuité t de F_Y. Ainsi, X_n converge en loi vers Y.

Une troisième correction

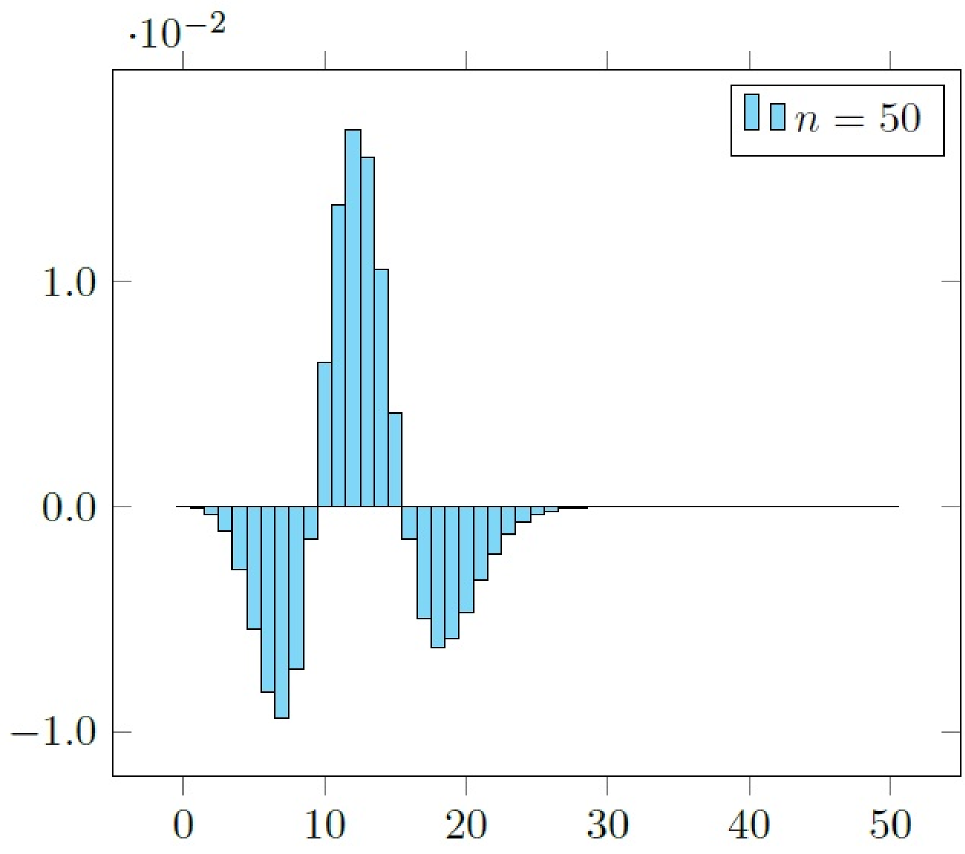
On peut se servir de la convergence de $\mathbb{P}(X_n = k)$ vers $\mathbb{P}(Y = k)$ prouvée précédemment pour obtenir directement la convergence de $\mathbb{E}[f(X_n)]$ vers $\mathbb{E}[f(Y)]$ pour toute fonction f continue et bornée. En effet, on a

$$\mathbb{E}[f(X_n)] = \sum_{k=0}^{\infty} \underbrace{\mathbb{P}(X_n = k)}_{\longrightarrow \mathbb{P}(Y=k)} f(k) \, .$$

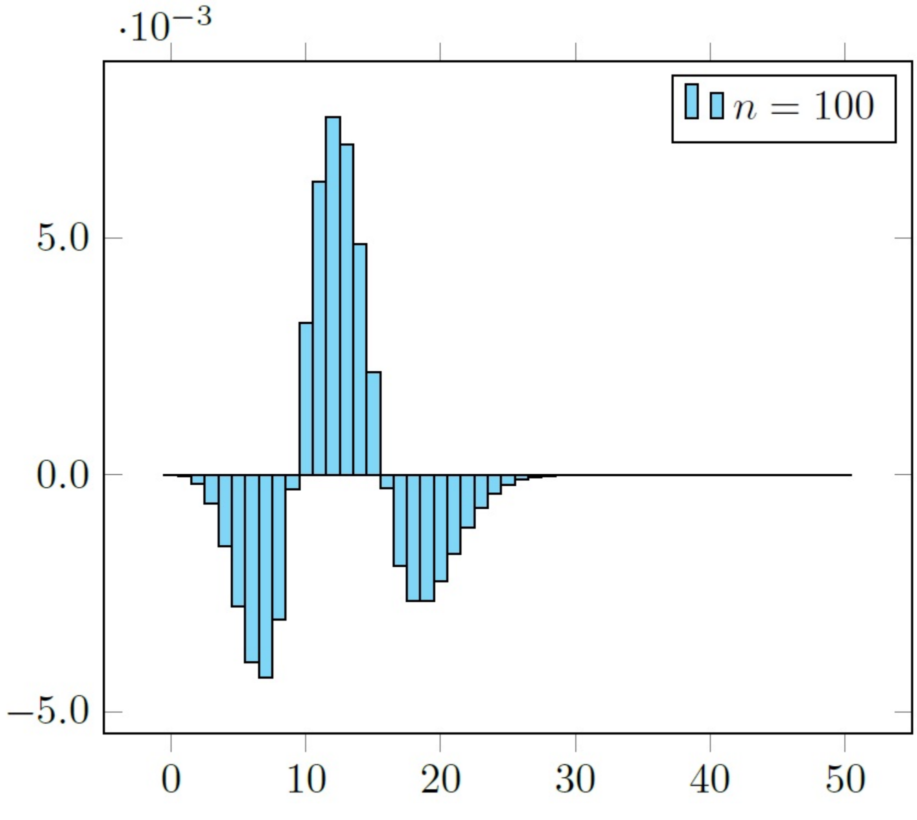
Il suffit pour conclure de montrer que la quantité $|\mathbb{P}(X_n = k) f(k)|$ est majorée pour tout $n, k \in \mathbb{N}$ par $b(k)$ où la suite b est sommable. Pour ce faire, on utilise des arguments similaires à ceux employés dans la remarque.

Illustration

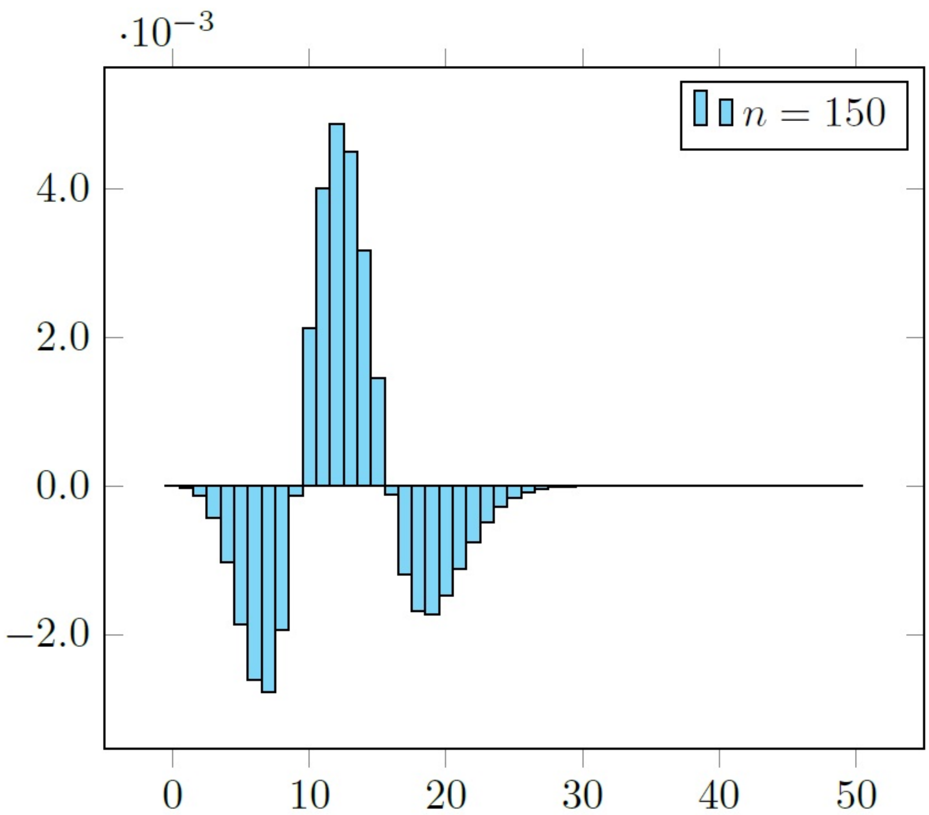
On fournit, pour $\lambda = 12$ une représentation graphique de la différence entre la distribution $\mathcal{P}(\lambda)$ et la loi binomiale de paramètres n et $p_n = \frac{\lambda}{n}$. Pour $n = 50$:



Pour $n = 100$:


$\cdot 10^{-3}$
$n = 100$
5.0
0.0
−5.0
0
10
20
30
40
50

Pour $n = 150$:


$\cdot 10^{-3}$
$n = 150$
4.0
2.0
0.0
−2.0
0
10
20
30
40
50

Exercice 73

Énoncé

Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes et de même loi $\mathcal{N}(0;1)$. Déterminer la limite en loi de la suite $(Y_n)_n$ avec

$$Y_n := \frac{1}{n}\sum_{k=1}^{n} \sqrt{k} X_k\,.$$

Rappels

Pour déterminer la limite en loi d'une suite de variables aléatoires réelles, on passe dans le domaine fréquentiel en calculant la fonction caractéristique et la limite de celle-ci. Également, on rappelle que la fonction caractéristique de la loi normale de paramètres m et σ^2 est $\psi(u) = \exp\left\{imu - \frac{\sigma^2}{2}u^2\right\}$.

Remarque

Il est ici important de montrer que la fonction limite est continue en 0 et de valeur égale à 1.

Correction

On calcule la fonction caractéristique de Y_n. On rappelle que l'on a

$$\varphi_{X_n}(t) := \mathbb{E}\left\{\exp\left(itX_n\right)\right\} = e^{-\frac{t^2}{2}}\,.$$

Par définition, il vient

$$\varphi_{Y_n}(t) = \mathbb{E}\left\{\exp\left(it\sum_{k=1}^{n}\frac{\sqrt{k}}{n}X_k\right)\right\}\,.$$

L'indépendance des variables aléatoires X_k nous donne

$$\varphi_{Y_n}(t) = \prod_{k=1}^{n}\mathbb{E}\left\{\exp\left(it\frac{\sqrt{k}}{n}X_k\right)\right\} = \prod_{k=1}^{n}\varphi_{X_k}\left(t\frac{\sqrt{k}}{n}\right)\,.$$

Puis, comme $\varphi_{X_k}(t) = e^{-\frac{t^2}{2}}$, on a

$$\varphi_{Y_n}(t) = \prod_{k=1}^{n} e^{-\frac{t^2}{2}\frac{k}{n^2}} = e^{-t^2\frac{\sum_{k=1}^{n}k}{2n^2}} \longrightarrow e^{-\frac{t^2}{4}} = e^{-\frac{\frac{1}{2}}{2}t^2}\,.$$

Cette fonction vaut 1 en 0 et elle est continue. Il s'ensuit que $(Y_n)_n$ converge en loi vers la loi normale centrée et de variance égale à $\frac{1}{2}$.

Autre correction

On peut aussi utiliser la densité de probabilité. En effet, comme les variables aléatoires réelles sont indépendantes et suivent une loi normale, on sait que Y_n suit une loi normale de paramètres m_n et σ_n^2. Évaluons les paramètres. D'abord, comme $\mathbb{E}[X_k] = 0$, il s'ensuit $m_n = 0$. Puis, on a

$$\sigma_n^2 = \mathrm{Var}(Y_n) = \frac{1}{n^2}\sum_{k=1}^{n} k\mathrm{Var}(X_k) = \frac{n+1}{2n} \longrightarrow \frac{1}{2} = \sigma_\infty\,,$$

quand n tend vers l'infini. On a donc

$$f_{Y_n}(y) = \frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left\{-\frac{y^2}{2\sigma_n^2}\right\} \longrightarrow \frac{1}{\sqrt{2\pi\sigma_\infty^2}}\exp\left\{-\frac{y^2}{2\sigma_\infty^2}\right\} =: f(y)\,,$$

pour tout $y \in \mathbb{R}$, lorsque n tend vers l'infini. Ici, f est la densité de probabilité de la loi normale centrée et de variance égale à $\frac{1}{2}$. Par conséquent, $(Y_n)_n$ converge en loi vers la loi normale centrée et de variance égale à $\frac{1}{2}$.

Exercice 74

Énoncé

Soit $(X_n)_n$ une suite de variables aléatoires telle que, pour tout n, X_n suit la loi uniforme sur l'ensemble fini

$$\mathbb{A}_n := \left\{0; \frac{1}{n}; \frac{2}{n}; \cdots; \frac{n-1}{n}\right\} .$$

Montrer que $(X_n)_n$ converge en loi vers une loi uniforme sur l'intervalle $[0;1]$.

Correction

On calcule la fonction caractéristique de X_n :

$$\varphi_{X_n}(t) := \mathbb{E}\left\{\exp\left(itX_n\right)\right\} = \frac{1}{n}\sum_{k=0}^{n-1} e^{it\frac{k}{n}} .$$

Cette quantité converge vers $\int_0^1 e^{itx}dx$ (en tant que limite d'une somme de Riemann), ce qui est la fonction caractéristique en t de la loi uniforme sur $[0;1]$. La preuve est ainsi achevée.

Remarque

On peut calculer $\varphi_{X_n}(t)$ (pour $t \neq 2\pi np$) sans utiliser les sommes de Riemann :

$$\varphi_{X_n}(t) = \frac{1}{n}\sum_{k=0}^{n-1} e^{it\frac{k}{n}} = \frac{1}{n}\frac{1-e^{it}}{1-e^{\frac{it}{n}}} \longrightarrow \frac{e^{it}-1}{it} ,$$

quand n tend vers l'infini. Et, on reconnaît l'expression de la fonction caractéristique en $t \neq 2\pi np$ de la loi uniforme sur $[0;1]$. La preuve est ainsi achevée.

Autre correction

Soit f une fonction continue et bornée (comme on travaille sur $[0;1]$, elle est nécessairement bornée en fait). Alors, on a

$$\mathbb{E}[f(X_n)] = \frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \longrightarrow \int_0^1 f(x)dx ,$$

en utilisant la somme de Riemann. Or, $\mathbb{E}[f(Y)] = \int_0^1 f(x)dx$ où Y suit la loi uniforme sur $[0;1]$. Ceci permet de conclure.

Exercice 75

Énoncé

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes de loi $\mathcal{N}(m, \sigma^2)$. On note $Y_n := \frac{1}{n}\sum_{k=1}^{n} X_k^2$.
Montrer que la suite $(Y_n)_n$ converge presque sûrement et déterminer sa limite.

Correction

On a $\mathbb{E}[X_1^2] = \mathrm{Var}[X_1] + \mathbb{E}[X_1]^2 = \sigma^2 + m^2$. De plus, les variables aléatoires X_n^2 sont mutuellement indépendantes. Puis, comme X_1 suit une loi normale, son moment d'ordre 8 est fini si bien que l'on a $\mathbb{E}\left((X_1^2)^4\right) < \infty$. Conséquemment, on peut utiliser la loi forte des grands nombres avec la suite $(X_k^2)_k$et l'on obtient la convergence presque sûre de Y_n vers $\mathbb{E}[X_1^2] = m^2 + \sigma^2$.

Exercice 76

Énoncé

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes de loi $\mathcal{E}\,(1)$. On note $S_n := \sum_{k=1}^{n} X_k$.
Montrer que $\frac{S_n - n}{\sqrt{n}}$ converge en loi vers $\mathcal{N}\,(0,1)$.

Correction

On a $\mathbb{E}[X_n] = 1$ et $\mathrm{Var}[X_n] = 1$ donc il s'agit d'une simple application du théorème central limite.

Remarque 1

On peut démontrer cette convergence en loi sans utiliser le théorème central limite. En effet, on peut calculer facilement la densité de probabilité de S_n.

Démontrons que pour tout $n \in \mathbb{N}^*$, $f_{S_n}(x) = \frac{x^{n-1}}{(n-1)!}e^{-x}H(x)$. Pour ce faire, procédons à une récurrence.

D'abord, si $n = 1$, on a bien $f_{S_1}(x) = f_{X_1}(x) = e^{-x}H(x)$.

Supposons maintenant la propriété vraie au rang $n \in \mathbb{N}^*$. Alors, l'indépendance des variables aléatoires implique $f_{S_{n+1}}(x) = \left(f_{S_n} * f_{X_{n+1}}\right)(x)$ d'où

$$\begin{aligned}
f_{S_{n+1}}(x) &= \int_{\mathbb{R}} f_{X_{n+1}}(x-y) f_{S_n}(y)dy \\
&= H(x)\int_0^x e^{-(x-y)}\frac{y^{n-1}}{(n-1)!}e^{-y}dy \\
&= H(x)e^{-x}\int_0^x \frac{y^{n-1}}{(n-1)!}dy \\
&= \frac{x^n}{n!}e^{-x}H(x)\,.
\end{aligned}$$

On peut ainsi conclure que l'on a bien $f_{S_n}(x) = \frac{x^{n-1}}{(n-1)!}e^{-x}H(x)$ pour tout $n \geq 1$.

Désormais, on regarde la loi de $\frac{S_n - n}{\sqrt{n}} =: Y_n$. On utilise pour cela la méthode habituelle :

$$\begin{aligned}
F_{Y_n}(y) &= \mathbb{P}\left(Y_n \leq y\right) \\
&= \mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leq y\right) \\
&= \mathbb{P}\left(S_n \leq n + \sqrt{n}y\right) \\
&= F_{S_n}\left(n + \sqrt{n}y\right)\,.
\end{aligned}$$

Cette fonction est bien continue sur $\mathbb{R}$ et dérivable presque partout. Il s'ensuit que Y_n est à densité. De plus, sa densité est

$$f_{Y_n}(y) := \frac{d}{dy}F_{Y_n}(y) = \sqrt{n}f_{S_n}\left(n+\sqrt{n}y\right) = \sqrt{n}\frac{(n+\sqrt{n}y)^{n-1}}{(n-1)!}e^{-n-\sqrt{n}y}H\left(n+\sqrt{n}y\right) .$$

Soit maintenant $y \in \mathbb{R}$ quelconque. Alors, pour n assez grand, $n+\sqrt{n}y > 0$ d'où $H(n+\sqrt{n}y) = 1$. On a donc, pour y fixé et pour n assez grand :

$$f_{Y_n}(y) = \sqrt{n}\frac{(n+\sqrt{n}y)^{n-1}}{(n-1)!}e^{-n-\sqrt{n}y} .$$

On utilise maintenant la formule de Stirling selon laquelle $n!$ est équivalent en l'infini à $n^n e^{-n}\sqrt{2\pi n}$. D'où

$$\begin{aligned}
f_{Y_n}(y) &\sim \sqrt{n}n^{n-1}\left(1+\frac{y}{\sqrt{n}}\right)^{n-1}\frac{1}{(n-1)^{n-1}e^{-(n-1)}\sqrt{2\pi(n-1)}}e^{-n}e^{-\sqrt{n}y} \\
&\sim \underbrace{\frac{\sqrt{n}}{\sqrt{2\pi(n-1)}}}_{\longrightarrow\frac{1}{\sqrt{2\pi}}}\ \underbrace{\frac{n^{n-1}}{(n-1)^{n-1}e}}_{=\frac{1}{e\left(1-\frac{1}{n}\right)^{n-1}}\longrightarrow 1}\ e^{-\sqrt{n}y}\left(1+\frac{y}{\sqrt{n}}\right)^{n-1} \\
&\sim \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}y}\left(1+\frac{y}{\sqrt{n}}\right)^{n-1} = \frac{1}{\sqrt{2\pi}}e^{-\sqrt{n}y}\left(1+\frac{y}{\sqrt{n}}\right)^{n}\left(1+\frac{y}{\sqrt{n}}\right)^{-1} .
\end{aligned}$$

On utilise maintenant un raisonnement classique :

$$\begin{aligned}
\left(1+\frac{y}{\sqrt{n}}\right)^n &= \exp\left[n\left(\frac{y}{\sqrt{n}}-\frac{1}{2}\left(\frac{y}{\sqrt{n}}\right)^2+o\left(\frac{1}{n}\right)\right)\right] \\
&= \exp\left[\sqrt{n}y-\frac{y^2}{2}+o(1)\right] .
\end{aligned}$$

Il s'ensuit

$$\begin{aligned}
f_{Y_n}(y) &\sim \frac{1}{\sqrt{2\pi}}\underbrace{e^{-\sqrt{n}y}\exp\left[\sqrt{n}y-\frac{y^2}{2}+o(1)\right]}_{\longrightarrow e^{-\frac{y^2}{2}}}\underbrace{\left(1+\frac{y}{\sqrt{n}}\right)^{-1}}_{\longrightarrow 1} \\
&\longrightarrow \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} ,
\end{aligned}$$

ce qui achève de prouver la convergence en loi de $\frac{S_n-n}{\sqrt{n}}$ vers la loi normale centrée réduite.

Remarque 2

On peut aussi prouver la convergence en loi avec la fonction caractéristique. En effet, les variables aléatoires étant indépendantes, il s'ensuit

$$\varphi_{S_n}(u) = \varphi_{X_1}(u)^n = (1 - iu)^{-n} \ .$$

Or,

$$\varphi_{\frac{S_n - n}{\sqrt{n}}}(u) = e^{-i\sqrt{n}u}\varphi_{S_n}\left(\frac{u}{\sqrt{n}}\right) \ ,$$

puis un développement limité à l'ordre deux achève la preuve.

Exercice 77

Énoncé

Un logiciel doit faire un calcul comportant cinquante nombres. Il arrondit chacun de ces nombres à l'entier le plus proche et effectue leur somme.
Si les erreurs d'arrondi individuelles sont distribuées uniformément sur l'intervalle $\left[-\frac{1}{2};\frac{1}{2}\right]$, calculer la probabilité pour que la somme obtenue ait un écart de plus de trois par rapport à la somme exacte.

Correction

On suppose d'abord que les erreurs d'arrondi sont indépendantes. On pourrait calculer la loi de la somme de cinquante variables aléatoires suivant la loi uniforme mais ce serait extrêmement fastidieux. De plus, on note que $50 > 30$. On peut donc utiliser la théorème central limite. L'espérance de U_i est 0 pour tout i. Et, la variance de U_i est $\frac{1}{12}$. Par conséquent, le TCL nous dit que $Y := \frac{\sum_{i=1}^{50} U_i}{\sqrt{50\times\frac{1}{12}}}$ suit approximativement la loi normale centrée réduite.

On cherche

$$\begin{aligned}
\mathbb{P}\left(|U_1+\cdots+U_{50}|>3\right) &= \mathbb{P}\left(|Y| > \frac{3\sqrt{6}}{5}\right) \\
&\approx \mathbb{P}\left(|Y|>1.47\right) \\
&\approx 2\left(1-\Phi(1.47)\right) \\
&\approx 2-2\times 0.9292 \\
&\approx 0.1416\,.
\end{aligned}$$

Remarque

Le calcul de la loi de $U_1+\cdots+U_{50}$ peut se faire par le produit de convolution de 50 fois la fonction indicatrice sur $\left[-\frac{1}{2};\frac{1}{2}\right]$. Mais, nous allons éviter.

Néanmoins, on peut utiliser la transformée de Fourier. En effet, on sait que la fonction caractéristique de U_i est $\varphi_U(t) = 2\frac{\sin(\frac{u}{2})}{u}$ donc la fonction caractéristique de $U_1+\cdots+U_{50}$ est $2^{50}\left(\frac{\sin(\frac{u}{2})}{u}\right)^{50}$. Cette fonction est bien sommable donc on peut utiliser le théorème d'inversion et la densité de probabilité de la somme de 50 variables aléatoires suivant la loi uniforme sur $\left[-\frac{1}{2};\frac{1}{2}\right]$ est

$$x \mapsto \frac{1}{2\pi}\int_{\mathbb{R}} e^{iux}2^{50}\left(\frac{\sin(\frac{u}{2})}{u}\right)^{50} du\,.$$

Évidemment, le calcul de l'intégrale n'est pas immédiat.

On peut en fait donner une expression “simple” de la densité de probabilité de la somme de 50 variables aléatoires suivant la loi uniforme sur $\left[-\frac{1}{2};\frac{1}{2}\right]$. Pour ce faire, on remarque $\mathbb{1}_{\left[-\frac{1}{2};\frac{1}{2}\right]} = \left(\delta_{-\frac{1}{2}} - \delta_{\frac{1}{2}}\right) * H$. Ainsi, la densité de probabilité recherchée est :

$$\left(\delta_{-\frac{1}{2}} - \delta_{\frac{1}{2}}\right)^{*50} * H^{*50}\,.$$

Or, on peut montrer par récurrence, voir la Remarque sur l'Exercice 76, que l'on a : $H^{*50}(x) = \frac{x^{49}}{49!}H(x)$. Puis, en utilisant le binôme de Newton :

$$\left(\delta_{-\frac{1}{2}} - \delta_{\frac{1}{2}}\right)^{*50} = \sum_{k=0}^{50}\binom{50}{k}(-1)^k\delta_{25-k}\,.$$

Par conséquent, on obtient :

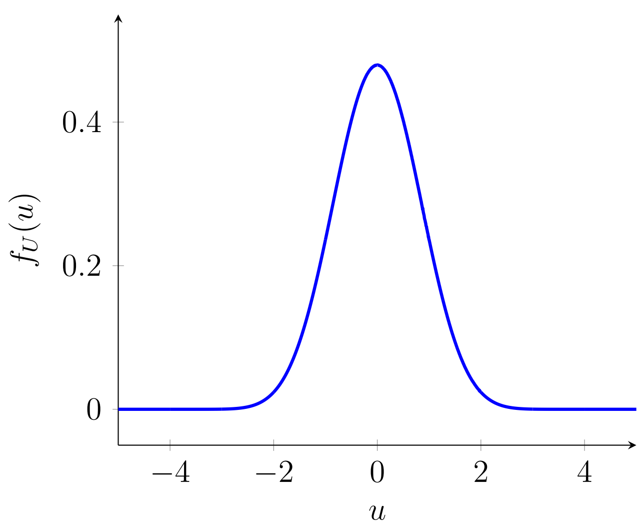
$$f_{U_1+\cdots+U_{50}}(x) = \sum_{k=0}^{50}\binom{50}{k}(-1)^k\frac{(x+k-25)^{49}}{49!}H(x+k-25)\,.$$

On peut aussi donner explicitement

$$f_{U_1+\cdots+U_8}(x) = \sum_{k=0}^{8}\binom{8}{k}(-1)^k\frac{(x+k-4)^{7}}{7!}H(x+k-4)\,,$$

dont voici une représentation graphique :

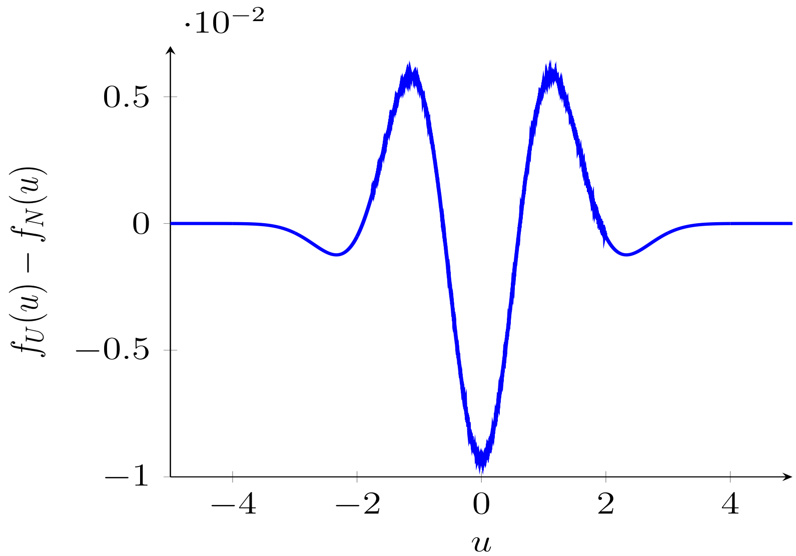
Figure 1 – Densité de probabilité de la somme



Par ailleurs, d'après le TCL, si 8 est suffisamment grand (normalement, on a besoin de $n \geq 30$), on s'attend à ce que $U_1 + \cdots + U_8$ suive approximativement la loi normale centrée et de variance $\frac{8}{12} = \frac{2}{3}$.

Voici une illustration graphique de la différence entre les deux lois :

Figure 2 – Différence entre les densités de probabilité des deux lois



On observe en particulier que $|f_{S_8}(x) - f_X(x)| < \frac{1}{100}$ pour tout $x \in [-4; 4]$. Ainsi, bien que 8 soit plus petit que 30, on peut appliquer le TCL.

Exercice 78

Énoncé

Une entreprise fabrique des briquets. Elle envoie à ses revendeurs des cartons comportant 30 briquets. Néanmoins, l'emballage est fait à la main si bien que seuls 95% des cartons contiennent effectivement 30 briquets. 3% en contiennent 28 et 2% en contiennent 31.

1. En notant X le nombre de briquets que contient un carton pris au hasard à la sortie de l'usine, donner la loi de X, son espérance et sa variance σ^2.
2. Un magasin a commandé 12 000 briquets c'est-à-dire 400 cartons. Calculer la probabilité que le magasin en reçoive moins de 12 000. On utilisera l'approximation $\frac{1}{\sigma} \approx 2.6875$.

Correction

Correction du 1)

La loi est donnée dans l'énoncé : $\mathbb{P}_X = \frac{3}{100}\delta_{28} + \frac{95}{100}\delta_{30} + \frac{2}{100}\delta_{31}$. On a directement :

$$\begin{aligned}
\mathbb{E}(X) &= \frac{3\times 28 + 95\times 30 + 2\times 31}{100} \\
&= \frac{3\times(30-2) + 95\times 30 + 2\times(30+1)}{100} \\
&= 30 + \frac{3\times(-2) + 2\times 1}{100} \\
&= 30 - \frac{4}{100}\,.
\end{aligned}$$

Numériquement, $\mathbb{E}[X] = 29.96$.
Et, de même :

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{3\times 28^2 + 95\times 30^2 + 2\times 31^2}{100} - \left(30 - \frac{4}{100}\right)^2 \\
&= \frac{3\times(30-2)^2 + 95\times 30^2 + 2\times(30+1)^2}{100} - 30^2 - \frac{16}{10^4} + \frac{240}{100} \\
&= \frac{3\times 30^2 + 3\times(-120+4) + 95\times 30^2 + 2\times 30^2 + 2\times(60+1)}{100} - 30^2 + \frac{240}{100} - \frac{16}{10^4} \\
&= 30^2 + \frac{-348+122}{100} - 30^2 + \frac{240}{100} - \frac{16}{10^4} \\
&= \frac{14}{100} - \frac{16}{10^4} \\
&= \frac{1400-16}{10^4} \\
&= \frac{1384}{10^4}\,.
\end{aligned}$$

Numériquement, $\mathrm{Var}(X) = 0.1384$.

Correction du 2)

Le nombre de briquets que va recevoir le magasin est donné par $Y := X_1 + \cdots + X_{400}$ où X_i est le nombre de briquets dans le carton numéro i. Les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi $\mathbb{P}_X = \frac{3}{100}\delta_{28} + \frac{95}{100}\delta_{30} + \frac{2}{100}\delta_{31}$.

On cherche $\mathbb{P}(Y < 12\ 000)$. On va procéder à une approximation via le TCL (ce qui est possible car $400 > 30$) :

$$\begin{aligned}
\mathbb{P}(Y < 12\ 000) &= \mathbb{P}\left(\underbrace{\frac{Y - \mathbb{E}(Y)}{\sqrt{\mathrm{Var}(Y)}}}_{=:Z} < \frac{12\ 000 - \mathbb{E}(Y)}{\sqrt{\mathrm{Var}(Y)}}\right) \\
&= \mathbb{P}\left(Z < \frac{12\ 000 - 400 \times \left(30 - \frac{4}{100}\right)}{\sqrt{400 \times \sigma^2}}\right) \\
&= \mathbb{P}\left(Z < \frac{16}{20\sigma}\right) \\
&\approx \mathbb{P}\left(Z < \frac{4}{5} \times 2.6875\right) \\
&= \mathbb{P}(Z < 2.15)\,.
\end{aligned}$$

En effet, $4 \times 2.6875 = 2 \times 5.375 = 10.75$ et $\frac{10.75}{5} = 2.15$. D'après le TCL, Z suit approximativement la loi normale. Il vient : $\mathbb{P}(Y < 12\ 000) \approx \Phi(2.15) = 0.9842$.

Remarque

Comme $\mathbb{E}[Y] < 12\ 000$, on pouvait s'attendre à ce que cette probabilité soit supérieure à $\frac{1}{2}$. On pourrait se demander à quel point le magasin se fait arnaquer. Calculons par exemple

$$\begin{aligned}
\mathbb{P}(Y < 11\ 975) &= \mathbb{P}(Z < \frac{-9}{20\sigma}) \\
&\approx \mathbb{P}(Z < -1.21) \\
&\approx 1 - \Phi(1.21) \\
&\approx 0.1131\,.
\end{aligned}$$

Ainsi, il y a environ 11% de chance que le magasin ait 25 briquets de moins que prévu.

Exercice 79 (*)

Énoncé

Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes telle que, pour tout n, X_n est de fonction de répartition F_n définie par

$$F_n(x) := \left(1 - \frac{1}{n+x}\right) \mathbb{1}_{\mathbb{R}_+^*}(x)\,.$$

On note $S_n := \sum_{k=1}^n X_k$ et $Y_n := \frac{S_n}{n}$.
Démontrer que la suite $(X_n)_n$ converge vers 0 en probabilité mais que la suite $(Y_n)_n$ ne converge pas vers 0 en probabilité.

Correction

Les variables aléatoires X_n sont presque sûrement positives. On a, pour tout $\epsilon > 0$,

$$\mathbb{P}\left(|X_n| \geq \epsilon\right) = 1 - F_n(\epsilon) = \frac{1}{n+\epsilon}\,.$$

Conséquemment, $\lim\limits_{n\to\infty} \mathbb{P}\left(|X_n| \geq \epsilon\right) = 0$, c'est-à-dire que la suite $(X_n)_n$ converge vers 0 en probabilité.
Par ailleurs, si l'on pose $M_n := \max\limits_{1\leq k\leq n} X_k$, les X_n étant toutes positives presque sûrement, on a $Y_n \geq \frac{M_n}{n}$, presque sûrement. Par conséquent, pour tout $\epsilon > 0$:

$$\mathbb{P}\left(Y_n \geq \epsilon\right) \geq \mathbb{P}\left(\frac{M_n}{n} \geq \epsilon\right)\,.$$

Les variables aléatoires X_n étant indépendantes, on a pour tout $x > 0$:

$$\begin{aligned}
\mathbb{P}\left(M_n \leq x\right) &= \mathbb{P}\left(\bigcap_{k=1}^n \{X_k \leq x\}\right) \\
&= \prod_{k=1}^n \mathbb{P}\left(X_k \leq x\right) \\
&= \prod_{k=1}^n \left(1 - \frac{1}{k+x}\right) \\
&\leq \left(1 - \frac{1}{n+x}\right)^n\,.
\end{aligned}$$

Il en résulte que

$$1 - \left(1 - \frac{1}{n(\epsilon+1)}\right)^n \leq \mathbb{P}\left(Y_n \geq \epsilon\right)\,.$$

Puis, en passant à la limite inférieure, on obtient

$$0 < 1 - \exp\left(-\frac{1}{\epsilon+1}\right) \leq \liminf_n \mathbb{P}\left(Y_n \geq \epsilon\right)\,,$$

ce qui prouve que la suite $(Y_n)_n$ ne converge pas vers 0 en probabilité.

Remarque

Si la suite numérique **déterministe** $(x_n)_n$ converge vers l alors la suite $(S_n)_n$ définie par $S_n := \frac{1}{n}\sum_{k=1}^{n} x_k$ converge également vers l. Il s'agit du lemme de Césaro.

Ici, on vient de montrer que ce lemme ne s'étend pas aux variables aléatoires pour la convergence en probabilité.

Exercice 80 (*)

Énoncé

On tente d'organiser le planning des entretiens de recrutement dans une école d'ingénieurs. On a constaté que 10% des personnes convoquées ne viennent pas à l'entretien. 250 personnes ont été convoquées pour les entretiens. On note X le nombre de personnes effectivement présentes pour passer l'entretien.

1. Déterminer la loi exacte de X.
2. Calculer de façon approchée $\mathbb{P}(X \leq 230)$. On prendra l'approximation $\sqrt{10} \approx 3.15$.

Correction

Correction du 1)

Posons $X_i := \begin{cases} 1 & \text{si la personne } i \text{ vient} \\ 0 & \text{sinon} \end{cases}$.

Les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre 0.9. Alors, $X := X_1 + \cdots X_{250}$ est une loi binomiale de paramètres 250 et 0.9.

Correction du 2)

On utilise le théorème central de la limite dans cette question. On a $\mathbb{E}(X_i) = 0.9$ et $\mathrm{Var}(X_i) = 0.9 \times (1 - 0.9) = 0.09$. On pose $Y := \frac{X - 250 \times 0.9}{\sqrt{0.09 \times 250}}$. D'après le TCL, Y suit approximativement la loi normale centrée réduite vu que $250 > 30$.

On a ainsi :

$$
\begin{aligned}
\mathbb{P}(X \leq 230) &= \mathbb{P}(Y \leq \frac{230 - 225}{1.5\sqrt{10}}) \\
&= \mathbb{P}(Y \leq \frac{\sqrt{10}}{3}) \\
&\approx \mathbb{P}(Y \leq 1.05)\,.
\end{aligned}
$$

D'après le TCL, Z suit approximativement la loi normale. Il vient : $\mathbb{P}(X \leq 230) \approx \Phi(1.05) = 0.8531$.

Exercice 81 (*)

Énoncé

Une société d'assurance A doit assurer 100 véhicules identiques, dont la valeur est de 10 000 euros. Sur un an, la probabilité pour qu'un véhicule soit accidenté et irréparable est de $p = 0.01$. Les accidents des voitures sont supposés indépendants. On suppose que A doit payer le 31 décembre sur son fonds de roulement tous les sinistres de l'année.

1. À combien doit s'élever le fonds de roulement si on veut que A puisse indemniser tous les sinistres dans 99% des cas ?

2. Une autre société d'assurance, la société B, doit effectuer le même travail que A dans les mêmes conditions (mais sur 100 autres véhicules). Les sociétés A et B projettent de fusionner. Elles assureraient les 200 véhicules. Calculer la valeur du fonds de roulement que devrait posséder la nouvelle société (dans les mêmes conditions que précédemment).

3. La fusion permet-elle de réaliser des économies ?

Correction

Correction du 1)

Posons $n := 100$. Pour tout $i \in [\![1; n]\!]$, on se donne X_i la variable aléatoire qui vaut 1 si la voiture est accidentée et irréparable dans l'année et 0 sinon. Les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre $p := 0.01$.

Le nombre de sinistres en une année est $\sum_{i=1}^n X_i$. D'après le Théorème Central de la Limite, $Y_n := \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}}$ suit approximativement la loi normale centrée réduite.

On cherche x tel que $\mathbb{P}(10\ 000 \sum_{i=1}^n X_i \leq x) \geq 0.99$. On transforme comme suit :

$$\begin{aligned}
\mathbb{P}(10\ 000 \sum_{i=1}^n X_i \leq x) &= \mathbb{P}\left(Y_n \leq \frac{\frac{x}{10\ 000} - np}{\sqrt{np(1-p)}}\right) \\
&\approx \Phi\left(\frac{\frac{x}{10\ 000} - np}{\sqrt{np(1-p)}}\right) .
\end{aligned}$$

On veut donc trouver x tel que $\frac{\frac{x}{10\ 000} - np}{\sqrt{np(1-p)}} \geq \Phi^{-1}(0.99) \approx 2.3266$. Il vient :

$$\begin{aligned}
x &\geq 10\ 000 \left(np + 2.3266 \times \sqrt{np(1-p)}\right) \\
&= 10\ 000 \left(100 \times 0.01 + 2.3266 \times 10 \times \frac{1}{10} \times \frac{\sqrt{99}}{10}\right) \\
&\approx 33\ 149.38 .
\end{aligned}$$

Le fonds de roulement doit donc être d'au moins 33 149.38 euros.

Correction du 2)

On refait les mêmes calculs avec $n := 200$ et l'on obtient

$$\begin{aligned} x &\geq 10\ 000 \left(200 \times 0.01 + 2.3266 \times 10\sqrt{2} \times \frac{1}{10} \times \frac{\sqrt{99}}{10} \right) \\ &\approx 52\ 738.16\,. \end{aligned}$$

Le fonds de roulement doit donc être d'au moins 52 738.16 euros.

Correction du 3)

On remarque qu'ensemble, les deux compagnies auront besoin d'un fonds de roulement inférieur à la somme des deux fonds qu'elles avaient chacune de leur côté vu que $52\ 738.16 < 66\ 298.76 = 2 \times 33\ 149.38$.

Sur ce plan, la fusion est donc avantageuse.

Exercice 82 (*)

Énoncé

Une machine doit percer une pièce métallique. Mais ce perçage est très technique et la probabilité qu'elle réussisse à faire le trou n'est que de $\frac{1}{2}$. La machine traite n pièces, de façon indépendante. On note p_n la proportion de pièces correctes :

$$p_n := \frac{\#\left\{\text{pièces correctement percées}\right\}}{n}\,.$$

1. Calculer p_∞, la limite de p_n pour n tendant vers l'infini.
2. Soit $\epsilon > 0$. Combien de pièces la machine doit-elle traiter pour être sûr à 95% que p_n soit proche à ϵ près de la probabilité limite p_∞ ?

Correction

Correction du 1)

Pour tout $i \in [\![1; n]\!]$, on se donne X_i la variable aléatoire qui vaut 1 si la pièce est bien percée et 0 sinon. Les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre $p := \frac{1}{2}$.

Ainsi, $p_n = \frac{\sum_{i=1}^n X_i}{n}$. Or, les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi $\mathcal{B}\left(\frac{1}{2}\right)$. En particulier, $\mathbb{E}(X_i) = \frac{1}{2}$ et $\mathrm{Var}(X_i) = \frac{1}{4} < +\infty$. De fait, en application de la loi forte des grands nombres, on en déduit que p_n converge presque sûrement vers $\mathbb{E}(X_1) = \frac{1}{2}$.

Correction du 2)

On cherche ici un entier n tel que

$$\mathbb{P}\left(\left|p_n - \frac{1}{2}\right| \leq \epsilon\right) \geq 0.95\,.$$

Ainsi, on veut n tel que

$$\mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i - \frac{n}{2}}{\frac{1}{2}\sqrt{n}}\right| \leq \frac{\sqrt{n}\epsilon}{\frac{1}{2}}\right) \geq 0.95\,.$$

On applique alors le TCL par indépendance des variables aléatoires (et car les variables aléatoires admettent un moment d'ordre 2 fini). De fait, la variable aléatoire $Y_n := \frac{\sum_{i=1}^n X_i - \frac{n}{2}}{\frac{1}{2}\sqrt{n}}$ suit approximativement la loi normale centrée réduite.

On veut donc n qui satisfasse $2\Phi(2\sqrt{n}\epsilon) - 1 \geq 0.95$ soit $2\sqrt{n}\epsilon \geq 1.645$. Il vient $n \geq \left(\frac{0.8225}{\epsilon}\right)^2$.

Exercice 83 (*)

Énoncé

On considère une fonction f positive sur l'intervalle $[a;b]$ avec $b > a$. On suppose de plus qu'il existe une constante $M > 0$ telle que $f(x) \leq M$ pour tout $x \in [a;b]$. On note $\mathcal{S}$ la surface sous la courbe de f entre a et b. On réalise l'expérience suivante : on tire de façon indépendante n points $A_1 := (X_1, Y_1), \cdots, A_n := (X_n, Y_n)$ au hasard dans le rectangle $[a;b] \times [0;M]$. En d'autres termes, les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi $\mathcal{U}_{[a;b]}$ et les variables aléatoires Y_i sont indépendantes et identiquement distribuées suivant la loi $\mathcal{U}_{[0;M]}$. De plus, X_i est indépendante de Y_j pour tout $1 \leq i, j \leq n$.

1. Quelle est la densité de probabilité du couple (X_1, Y_1) ?
2. Calculer la probabilité que le point A_1 soit dans la surface $\mathcal{S}$.
3. On pose $p_n := \frac{\#\{i \in [\![1;n]\!] \,:\, A_i \in \mathcal{S}\}}{n}$. Montrer que p_n converge presque sûrement vers la quantité $\frac{\int_a^b f(x)dx}{M(b-a)}$.
4. En déduire une façon de calculer une valeur approchée de $\int_a^b f(x)dx$.
5. On prend $a = 1$, $b = 2$ et $M = 4$. Quelle valeur donner à n pour être sûr à 95% d'obtenir une approximation de $\int_a^b f(x)dx$ à 10^{-2} près par la méthode précédente ?

Correction

Correction du 1)

Comme X_1 et Y_1 sont indépendantes, on a

$$f_{(X_1,Y_1)}(x, y) = f_{X_1}(x) f_{Y_1}(y) = \frac{1}{(b-a)M} \mathbb{1}_{[a;b]\times[0;M]}(x, y) .$$

Correction du 2)

Par définition :

$$\begin{aligned}\mathbb{P}(A_1 \in \mathcal{S}) &= \iint_{\mathcal{S}} f_{(X_1,Y_1)}(x, y) dx dy \\ &\iint_{\mathcal{S}} \frac{1}{(b-a)M} \mathbb{1}_{[a;b]\times[0;M]}(x, y) dx dy .\end{aligned}$$

Or, par définition de la surface, on a $\mathcal{S} = \{(x, y) \in [a;b] \times [0;M] \ : \ 0 \leq y \leq f(x)\}$. Il s'ensuit

$$
\begin{aligned}
\mathbb{P}(A_1 \in \mathcal{S}) &= \frac{1}{(b-a)M} \int_{x=a}^{x=b} \int_{y=0}^{y=f(x)} dydx \\
&= \frac{1}{(b-a)M} \int_a^b f(x)dx \\
&= \frac{\text{Aire de } \mathcal{S}}{(b-a)M} \,.
\end{aligned}
$$

Correction du 3)

On pose $X_i := 1$ si $A_i \in \mathcal{S}$ et $X_i = 0$ sinon. Par définition, les variables aléatoires X_i sont indépendantes et identiquement distribuées suivant la loi $\mathcal{B}\left(\frac{\text{Aire de } \mathcal{S}}{(b-a)M}\right)$.

Or, $p_n = \frac{1}{n}\sum_{i=1}^n X_i$. Donc, en application de la loi forte des grands nombres, on en déduit que p_n converge presque sûrement vers $\mathbb{E}(X_1) = \frac{\text{Aire de } \mathcal{S}}{(b-a)M}$.

Correction du 4)

Il suffit de tirer un grand nombre de points A_i et l'on aura une estimation de la quantité $\frac{\text{Aire de } \mathcal{S}}{(b-a)M}$ Il suffit ensuite de multiplier par $M(b-a)$.

Correction du 5)

Posons $I_n := M(b-a)p_n$. D'après la question 4, on sait que I_n converge presque sûrement vers $\int_a^b f(x)dx$. On cherche ici un entier n tel que

$$
\mathbb{P}\left(\left|I_n - \int_a^b f(x)dx\right| \leq 10^{-2}\right) \geq 0.95 \,.
$$

Or, on dispose de :

$$
\begin{aligned}
\mathbb{P}\left(\left|I_n - \int_a^b f(x)dx\right| \leq 10^{-2}\right) &= \mathbb{P}\left(\left|p_n - \frac{\int_a^b f(x)dx}{M(b-a)}\right| \leq \frac{10^{-2}}{M(b-a)}\right) \\
&= \mathbb{P}\left(\left|\underbrace{\frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{\sqrt{n\text{Var}(X_1)}}}_{=:Y_n}\right| \leq \frac{10^{-2}\sqrt{n}}{M(b-a)\sqrt{\text{Var}(X_1)}}\right) \,.
\end{aligned}
$$

Or, d'après le TCL, la variable aléatoire Y_n suit approximativement la loi normale centrée réduite dès que $n \geq 30$. On cherche ainsi n tel que $2\Phi\left(\frac{10^{-2}\sqrt{n}}{M(b-a)\sqrt{\text{Var}(X_1)}}\right) - 1 \geq 0.95$. On en déduit

$$
\frac{10^{-2}\sqrt{n}}{M(b-a)\sqrt{\text{Var}(X_1)}} \geq 1.96 \,,
$$

d'où $n \geq 1.96^2 \times 400^2 \times \mathrm{Var}(X_1) = 614\ 656\,(p - p^2)$ où $p := \frac{\int_a^b f(x)dx}{M(b-a)}$. Ainsi, il semblerait que nous soyons coincés. Heureusement, on remarque $p - p^2 \leq \frac{1}{4}$ pour tout $p \in [0;1]$. Conséquemment, il suffit de prendre $n \geq 614\ 656 \times \frac{1}{4} = 153\ 664$.