DS7 (vB) - Concours blanc - Correction (`ESSEC I 2023`)

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidat·es sont invité·es à ***encadrer*** *dans la mesure du possible les résultats de leurs calculs.*
Aucun document n'est autorisé. ***L'utilisation de toute calculatrice et de tout matériel électronique*** *est* ***interdite****. Seule l'utilisation d'une règle graduée est autorisée.*
Si au cours de l'épreuve, un candidat ou une candidate repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

On suppose, et c'est valable pour toute l'épreuve, que la librairie suivante est importée sous son alias habituel :

- `import numpy as np`
- `import numpy.linalg as al`
- `import numpy.random as rd`
- `import matplotlib.pyplot as plt`
- `import pandas as pd`

On s'intéresse dans ce sujet à la méthode de Stein, introduite par Charles Stein (1920/2016) en 1972, dont les développements et applications sont nombreux.

Les parties 1 et 2 concernent la justification de la méthode, elles sont indépendantes.

Dans la partie 3, on s'intéresse à l'estimation en un point d'une densité d'une loi de probabilité. Cette partie peut être traitée indépendamment des deux premières parties.

Dans la partie 4, on met en œuvre la méthode de Stein, vue dans les parties 1 et 2, pour établir des convergences « uniformes » en loi et on démontre le résultat admis dans la partie 3. Cette partie est indépendante de la partie 3 à l'exception de sa dernière question.

Dans tout le problème :

- les variables aléatoires considérées sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
- si X est une variable aléatoire, $\mathbb{E}(X)$ et $\mathbb{V}(X)$ désignent respectivement, lorsqu'elles existent, l'espérance et la variance de X.
- W désigne l'ensemble des fonctions h de classe $\mathcal{C}^1$ sur $\mathbb{R}$ telles que :
$$\forall x \in \mathbb{R},\ \left|h'(x)\right| \leqslant 1$$
- N est une variable aléatoire qui suit la loi normale $(0, 1)$.
- on admet que si X est une variable aléatoire possédant une espérance et $h \in W$, $\mathbb{E}(h(X))$ existe. On note en particulier c_h l'espérance de $h(N)$.
- On note Φ la fonction de répartition de la loi normale $(0, 1)$ définie par, pour tout $x \in \mathbb{R}$, $\Phi(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}\, dt$. On rappelle que c'est la primitive sur $\mathbb{R}$, qui vaut $\frac{1}{2}$ en 0, de la fonction $\varphi : t \mapsto \dfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}$.

Partie 1 - Transformation de Stein

Soit $h \in W$. On définit sur $\mathbb{R}$, la fonction $\theta : x \mapsto \dfrac{\Phi(x)}{\varphi(x)}$ et la fonction f_h par,
$$f_h : x \mapsto \theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\, dt + \theta(x) \int_{x}^{+\infty} h'(t)(1 - \Phi(t))\, dt$$
lorsque ces intégrales convergent.

L'objectif principal de cette partie est d'obtenir, pour X une variable aléatoire admettant une espérance, une expression de $\mathbb{E}(h(X)) - \mathbb{E}(h(N))$ qui ne fait pas intervenir N directement.

1. a) Montrer que pour tout $x \geqslant 0$ et $t \in [x, +\infty[$, $0 \leqslant x\varphi(t) \leqslant t\varphi(t)$. En déduire que :
$$\forall x \geqslant 0,\ 0 \leqslant x(1 - \Phi(x)) \leqslant \varphi(x)$$
(on remarquera que pour tout $t \in \mathbb{R}$, $\varphi'(t) = -t\varphi(t)$)

Commentaire

La remarque faite par l'énoncé peut sembler anodine, mais la formule $\varphi'(t) = -t\varphi(t)$ sera utilisée à de nombreuses reprises dans cette partie pour ne pas avoir à revenir aux formules explicites de $\varphi(t)$ et $\varphi'(t)$. La difficulté n'est donc pas de démontrer cette formule, mais bien d'y penser à chaque fois qu'elle sera utile.

Démonstration.
Soit $x \geqslant 0$ et soit $t \in [x, +\infty[$.

$$0 \leqslant x \leqslant t$$

donc $$0 \leqslant x\varphi(t) \leqslant t\varphi(t) \qquad \textit{(car } \varphi(t) \geqslant 0,\ \varphi \textit{ étant une densité)}$$

Soit $B \geqslant x$. Par croissance de l'intégrale, les bornes étant rangées dans l'ordre croissant ($x \leqslant B$) :

$$0 \leqslant \int_x^B x\varphi(t)\ dt \leqslant \int_x^B t\varphi(t)\ dt$$

donc $$0 \leqslant x \int_x^B \varphi(t)\ dt \leqslant -\int_x^B \varphi'(t)\ dt \qquad \textit{(car } \varphi'(t) = -t\varphi(t)\textit{)}$$

donc $$0 \leqslant x \int_x^B \varphi(t)\ dt \leqslant -\left[\ \varphi(t)\ \right]_x^B$$

donc $$0 \leqslant x \int_x^B \varphi(t)\ dt \leqslant \varphi(x) - \varphi(B)$$

Or, comme l'intégrale $\displaystyle\int_x^{+\infty} \varphi(t)\ dt$ converge (φ étant une densité), on peut écrire par relation de Chasles :

$$\lim_{B \to +\infty} \int_x^B \varphi(t)\ dt = \int_x^{+\infty} \varphi(t)\ dt = \int_{-\infty}^{+\infty} \varphi(t)\ dt - \int_{-\infty}^{x} \varphi(t)\ dt = 1 - \Phi(x)$$

De plus :

$$\lim_{B \to +\infty} \varphi(B) = \lim_{B \to +\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{B^2}{2}} = 0$$

donc, par passage à la limite lorsque B tend vers $+\infty$, on obtient bien :

$$\boxed{\forall x \geqslant 0,\ 0 \leqslant x(1 - \Phi(x)) \leqslant \varphi(x).}$$

Commentaire

- Tout d'abord, remarquons que la deuxième partie de la question peut se traiter même sans avoir fait la première. Aux concours, il est autorisé d'admettre une partie de la question et de traiter la suite.
- L'énoncé peut donner des indications sur la manière de résoudre les questions. Il peut être intéressant de repérer en première lecture de l'énoncé les questions pour lesquelles la méthode de résolution est précisée. En particulier, on pourra repérer celles commençant par « En déduire », « Montrer par récurrence », « Par intégration par parties », etc
- L'égalité

  $$1 - \Phi(x) = \int_x^{+\infty} \varphi(t)\ dt$$

  est classique et à connaître (elle est vue en cours au moment de démontrer $\Phi(-x) = 1 - \Phi(x)$). C'est la clé de la résolution de cette question : en faisant apparaître une intégrale, on comprend qu'on passera de la première inégalité à la seconde par intégration. Cela paraît d'autant plus pertinent que :
  - × la partie de droite de l'inégalité fait apparaître la fonction φ dont la dérivée est la fonction $t \mapsto -t\varphi(t)$
  - × l'encadrement $0 \leqslant x\varphi(t) \leqslant t\varphi(t)$ a précisément été démontré pour $t \geqslant x$

$\square$

b) Procéder de façon analogue pour montrer que : $\forall x \leqslant 0,\ -\varphi(x) \leqslant x\Phi(x) \leqslant 0$.

Démonstration.
Soit $x \leqslant 0$ et soit $t \in\ ]-\infty, x]$.

$$\begin{aligned} & t \leqslant x \leqslant 0 \\ \text{donc} \qquad & t\varphi(t) \leqslant x\varphi(t) \leqslant 0 \end{aligned}$$

Soit $A \leqslant x$. Par croissance de l'intégrale, les bornes étant rangées dans l'ordre croissant ($A \leqslant x$), on a

$$\begin{aligned} & \int_A^x t\varphi(t)\ dt \leqslant \int_A^x x\varphi(t)\ dt \leqslant 0 \\ \text{donc} \qquad & -\int_A^x \varphi'(t)\ dt \leqslant x\int_A^x \varphi(t)\ dt \leqslant 0 \\ \text{donc} \qquad & -\left[\ \varphi(t)\ \right]_A^x \leqslant x\int_A^x \varphi(t)\ dt \leqslant 0 \\ \text{donc} \qquad & \varphi(A) - \varphi(x) \leqslant x\int_A^x \varphi(t)\ dt \leqslant 0 \end{aligned}$$

Or, comme l'intégrale $\displaystyle\int_{-\infty}^x \varphi(t)\ dt$ converge (φ étant une densité), on peut écrire par définition de Φ :

$$\lim_{A\to-\infty} \int_A^x \varphi(t)\ dt = \int_{-\infty}^x \varphi(t)\ dt = \Phi(x)$$

De plus,

$$\lim_{A\to-\infty} \varphi(A) = \lim_{A\to-\infty} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{A^2}{2}} = 0$$

donc, par passage à la limite lorsque A tend vers $-\infty$, on obtient bien :

$$\boxed{\forall x \leqslant 0,\ -\varphi(x) \leqslant x\Phi(x) \leqslant 0.}$$

$\square$

c) En déduire à l'aide d'une intégration par parties, pour tout x réel, la convergence des intégrales qui suivent et montrer que :

$$\int_{-\infty}^x \Phi(t)\ dt = x\Phi(x) + \varphi(x) \quad \text{et} \quad \int_x^{+\infty} (1-\Phi(t))\ dt = -x(1-\Phi(x)) + \varphi(x) \qquad (R_1)$$

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

- Remarquons tout d'abord que les fonctions Φ et $1-\Phi$ sont continues sur $\mathbb{R}$ donc $\displaystyle\int_{-\infty}^x \Phi(t)\ dt$ est impropre seulement en $-\infty$ et $\displaystyle\int_x^{+\infty} (1-\Phi(t))\ dt$ est impropre seulement en $+\infty$.
- Soit $A \leqslant 0$. On procède par intégration par parties en posant :

$$\left|\begin{array}{ll} u'(t) = 1 & u(t) = t \\ v(t) = \Phi(t) & v'(t) = \varphi(t) \end{array}\right.$$

Commentaire

L'énoncé nous indique la marche à suivre : il faut faire une intégration par parties. Or, l'intégrale $\displaystyle\int_{-\infty}^{x} \Phi(t)\ dt$ ne fait apparaître qu'un seul terme. Il n'y a donc pas le choix : il faut dériver Φ et faire apparaître un 1 en facteur pour le primitiver.

Cette intégration par parties est valide car les fonctions u et v sont de classe $\mathcal{C}^1$ sur le segment $[A, x]$.

$$\begin{aligned}\int_{A}^{x} \Phi(t)\ dt &= \left[\ t\Phi(t)\ \right]_{A}^{x} - \int_{A}^{x} t\varphi(t)\ dt \\ &= x\Phi(x) - A\,\Phi(A) + \int_{A}^{x} \varphi'(t)\ dt \qquad \textit{(car } \varphi'(t) = -t\varphi(t)\textit{)} \\ &= x\Phi(x) - A\,\Phi(A) + \varphi(x) - \varphi(A)\end{aligned}$$

Il s'agit de montrer la convergence de $\displaystyle\int_{-\infty}^{x} \Phi(t)\ dt$. Pour cela, nous allons montrer que tous les termes de droite admettent une limite lorsque A tend vers $-\infty$.

× Tout d'abord :

$$\lim_{A \to -\infty} \varphi(A) = 0$$

× Ensuite, puisque $A \leqslant 0$, on a, d'après la question ***1.b)*** : $-\varphi(A) \leqslant A\,\Phi(A) \leqslant 0$.
Par théorème d'encadrement, on en déduit :

$$\lim_{A \to -\infty} A\,\Phi(A) = 0$$

$$\boxed{\text{Ainsi : pour tout } x \text{ réel, } \int_{-\infty}^{x} \Phi(t)\ dt \text{ converge et } \int_{-\infty}^{x} \Phi(t)\ dt = x\Phi(x) + \varphi(x).}$$

- Soit $B \geqslant 0$. On procède par intégration par parties en posant :

$$\left|\begin{array}{ll} u'(t) = 1 & u(t) = t \\ v(t) = 1 - \Phi(t) & v'(t) = -\varphi(t) \end{array}\right.$$

Cette intégration par parties est valide car les fonctions u et v sont de classe $\mathcal{C}^1$ sur le segment $[x, B]$.

$$\begin{aligned}\int_{x}^{B} (1 - \Phi(t))\ dt &= \left[\ t(1 - \Phi(t))\ \right]_{x}^{B} - \int_{x}^{B} -t\varphi(t)\ dt \\ &= B(1 - \Phi(B)) - x(1 - \Phi(x)) - \int_{x}^{B} \varphi'(t)\ dt \qquad \textit{(car } \varphi'(t) = -t\varphi(t)\textit{)} \\ &= B(1 - \Phi(B)) - x(1 - \Phi(x)) + \varphi(x) - \varphi(B)\end{aligned}$$

Il s'agit maintenant de montrer la convergence de $\displaystyle\int_{x}^{+\infty} (1 - \Phi(t))\ dt$. Pour cela, nous allons montrer que tous les termes de droite admettent une limite lorsque B tend vers $+\infty$.

× Tout d'abord :

$$\lim_{B \to +\infty} \varphi(B) = 0$$

× Ensuite, puisque $B \geqslant 0$, on a, d'après la question ***1.a)*** : $0 \leqslant B(1-\Phi(B)) \leqslant \varphi(B)$.
Par théorème d'encadrement, on en déduit :

$$\lim_{B \to +\infty} B(1-\Phi(B)) = 0$$

$$\boxed{\text{Ainsi : pour tout } x \text{ réel, } \int_x^{+\infty} (1-\Phi(t))\, dt \text{ converge et } \int_x^{+\infty} (1-\Phi(t))\, dt = -x(1-\Phi(x)) + \varphi(x).}$$

□

2. a) Montrer que pour tous réels x et y,

$$|h(x)-h(y)| \leqslant |x-y|, \quad \text{puis que } |h(x)| \leqslant |x| + |h(0)|$$

Démonstration.
Comme $h \in W$:

× la fonction h est de classe $\mathcal{C}^1$ sur $\mathbb{R}$,

× pour tout $x \in \mathbb{R}$, $|h'(x)| \leqslant 1$.

Ainsi, par inégalité des accroissements finis, on obtient :

$$\boxed{\forall (u,v) \in \mathbb{R}^2,\ |h(u)-h(v)| \leqslant |u-v|}$$

Soit $x \in \mathbb{R}$. D'après l'inégalité précédente appliquée en $u = x$ et $v = 0$:

$$|h(x)-h(0)| \leqslant |x|$$

d'où, par inégalité triangulaire,

$$|h(x)| = |(h(x)-h(0)) + h(0)| \leqslant |h(x)-h(0)| + |h(0)| \leqslant |x| + |h(0)|$$

$$\boxed{\text{On a bien : } \forall x \in \mathbb{R}, |h(x)| \leqslant |x| + |h(0)|}$$

□

b) Pour tout x réel, justifier la convergence de $\displaystyle\int_{-\infty}^{x} h'(t)\Phi(t)\, dt$ et montrer que :

$$\int_{-\infty}^{x} h'(t)\Phi(t)\, dt = h(x)\Phi(x) - \int_{-\infty}^{x} h(t)\varphi(t)\, dt$$

On admet de même que, $\displaystyle\int_x^{+\infty} h'(t)(1-\Phi(t))\, dt$ *converge et que,*

$$\int_x^{+\infty} h'(t)(1-\Phi(t))\, dt = -h(x)(1-\Phi(x)) + \int_x^{+\infty} h(t)\varphi(t)\, dt$$

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

- Montrons que l'intégrale $\displaystyle\int_{-\infty}^{x} h'(t)\Phi(t)\, dt$ converge absolument, ce qui justifiera sa convergence.

× Tout d'abord, $h \in W$ donc la fonction $t \mapsto h'(t)\Phi(t)$ est continue sur $\mathbb{R}$. On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{x} h'(t)\Phi(t)\,dt$ est seulement impropre en $-\infty$.

× Soit $t \in \,]-\infty, x]$. Puisque $h \in W$, on a : $|h'(t)| \leqslant 1$. De plus : $\Phi(t) \geqslant 0$. On en déduit :

$$\left|h'(t)\Phi(t)\right| = \left|h'(t)\right| \Phi(t) \leqslant \Phi(t)$$

× Les fonctions $t \mapsto |h'(t)\Phi(t)|$ et $t \mapsto \Phi(t)$ sont continues sur $\mathbb{R}$ donc sur $]-\infty, x]$.

× L'intégrale $\displaystyle\int_{-\infty}^{x} \Phi(t)\,dt$ converge (d'après la question ***1.c)***).

Par théorème de comparaison pour les intégrales généralisées de fonctions continues et positives, il vient que

$$\boxed{\int_{-\infty}^{x} h'(t)\Phi(t)\,dt \text{ est (absolument) convergente.}}$$

- Soit $A \leqslant 0$. On procède ensuite par intégration par parties en posant

$$\left| \begin{array}{ll} u'(t) = h'(t) & u(t) = h(t) \\ v(t) = \Phi(t) & v'(t) = \varphi(t) \end{array} \right.$$

Cette intégration par parties est valide car les fonctions u et v sont de classe $\mathcal{C}^1$ sur le segment $[A, x]$.

$$\begin{aligned} \int_{A}^{x} h'(t)\Phi(t)\,dt &= \left[\ h(t)\Phi(t)\ \right]_{A}^{x} - \int_{A}^{x} h(t)\varphi(t)\,dt \\ &= h(x)\Phi(x) - h(A)\Phi(A) - \int_{A}^{x} h(t)\varphi(t)\,dt \end{aligned}$$

Montrons : $\displaystyle\lim_{A \to -\infty} h(A)\Phi(A) = 0$.
Tout d'abord :

$$\begin{aligned} |h(A)\Phi(A)| &= |h(A)|\,\Phi(A) \\ &\leqslant (|A| + |h(0)|)\,\Phi(A) && \textit{(d'après la question \textbf{2.a)})} \\ &\leqslant -A\Phi(A) + |h(0)|\,\Phi(A) && \textit{(car } A \leqslant 0\textit{)} \end{aligned}$$

De plus,

× Φ est une fonction de répartition donc : $\displaystyle\lim_{A \to -\infty} \Phi(A) = 0$.

× D'après la question ***1.b)***, $\displaystyle\lim_{A \to -\infty} A\Phi(A) = 0$.

× On en déduit : $\displaystyle\lim_{A \to -\infty} -A\Phi(A) + |h(0)|\,\Phi(A) = 0$.

$$\boxed{\text{Par théorème d'encadrement : } \lim_{A \to -\infty} h(A)\Phi(A) = 0.}$$

Remarquons maintenant que :

$$\int_{A}^{x} h(t)\varphi(t)\,dt = h(x)\Phi(x) - h(A)\Phi(A) - \int_{A}^{x} h'(t)\Phi(t)\,dt$$

et tous les termes de droite admettent une limite finie lorsque A tend vers $-\infty$.

Ainsi, $\displaystyle\int_{-\infty}^{x} h(t)\varphi(t)\,dt$ converge et on peut écrire, par passage à la limite lorsque A tend vers $-\infty$:

$$\boxed{\text{pour tout } x \text{ réel, } \int_{-\infty}^{x} h'(t)\Phi(t)\ dt = h(x)\Phi(x) - \int_{-\infty}^{x} h(t)\varphi(t)\ dt.}$$

Commentaire

Tout passage à la limite doit être justifié par l'existence des limites de chacun des termes en présence. Ici, il ne faut pas oublier de démontrer la convergence de l'intégrale impropre $\displaystyle\int_{-\infty}^{x} h(t)\varphi(t)\ dt$ (c'est-à-dire l'existence de $\displaystyle\lim_{A\to-\infty} \int_{A}^{x} h(t)\varphi(t)\ dt$), même si l'énoncé ne nous le demande pas explicitement.

□

c) En déduire que, pour tout x réel :

$$-\int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt = c_h - h(x)$$

Démonstration.
D'après la question précédente :

$$\begin{aligned}
& -\int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \\
&= -\left(h(x)\Phi(x) - \int_{-\infty}^{x} h(t)\varphi(t)\ dt\right) - h(x)(1-\Phi(x)) + \int_{x}^{+\infty} h(t)\varphi(t)\ dt \\
&= \cancel{-h(x)\Phi(x)} + \int_{-\infty}^{x} h(t)\varphi(t)\ dt - h(x) + \cancel{h(x)\Phi(x)} + \int_{x}^{+\infty} h(t)\varphi(t)\ dt \\
&= \int_{-\infty}^{+\infty} h(t)\varphi(t)\ dt - h(x) \qquad \textit{(par relation de Chasles)}
\end{aligned}$$

Remarquons maintenant que φ est une densité de N et que la fonction h est continue sur $\mathbb{R}$. D'après le théorème de transfert :

× La variable aléatoire $h(N)$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} h(t)\varphi(t)\ dt$ est absolument convergente.

× En cas de convergence : $\mathbb{E}(h(N)) = \displaystyle\int_{-\infty}^{+\infty} h(t)\varphi(t)\ dt$.

Or, d'après l'énoncé, $h(N)$ admet une espérance. On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} h(t)\varphi(t)\ dt$ converge absolument (on savait déjà qu'elle était convergente) et que :

$$\int_{-\infty}^{+\infty} h(t)\varphi(t)\ dt = \mathbb{E}(h(N)) = c_h$$

$$\boxed{\text{Finalement : } \forall x \in \mathbb{R},\ -\int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt = c_h - h(x).}$$

Commentaire

Remarquons que le théorème de transfert n'est pas utilisé dans cette question de la manière habituelle. D'habitude, on utilise ce théorème pour **démontrer** qu'une variable aléatoire de la forme $g(X)$ admet une espérance. Ici, on **l'admet**.

□

3. ***a)*** Établir que pour tout $x \in \mathbb{R}$,

$$\theta'(x) = 1 + x\theta(x)$$
$$\theta''(x) = x + (1+x^2)\theta(x)$$
$$\theta(-x)\Phi(x) = \theta(x)(1-\Phi(x))$$

Démonstration.
La fonction θ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ comme quotient de deux fonctions de classe $\mathcal{C}^2$ sur $\mathbb{R}$ dont le dénominateur ne s'annule pas.
Soit $x \in \mathbb{R}$.

- D'une part,

$$\begin{aligned}
\theta'(x) &= \frac{\varphi(x)^2 - \Phi(x)\varphi'(x)}{\varphi(x)^2} \\
&= \frac{\varphi(x)^2 + \Phi(x)\, x\varphi(x)}{\varphi(x)^2} && \textit{(car } \varphi'(x) = -x\varphi(x)\textit{)} \\
&= 1 + x\frac{\Phi(x)}{\varphi(x)} \\
&= 1 + x\theta(x)
\end{aligned}$$

  et donc

$$\begin{aligned}
\theta''(x) &= \theta(x) + x\theta'(x) \\
&= \theta(x) + x(1 + x\theta(x)) \\
&= \theta(x) + x + x^2\theta(x) \\
&= x + (1+x^2)\theta(x)
\end{aligned}$$

- D'autre part,

$$\begin{aligned}
\theta(-x)\Phi(x) &= \frac{\Phi(-x)}{\varphi(-x)}\Phi(x) \\
&= \frac{\Phi(-x)}{\varphi(x)}\Phi(x) && \textit{(car } \varphi \textit{ est paire)} \\
&= \Phi(-x)\frac{\Phi(x)}{\varphi(x)} \\
&= (1-\Phi(x))\theta(x) && \textit{(par propriété de } \Phi\textit{)}
\end{aligned}$$

□

b) En déduire que f_h est une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$ qui vérifie, pour tout x réel :

$$f_h'(x) - xf_h(x) = c_h - h(x)$$

Pourquoi peut-on alors affirmer que f_h est de classe $\mathcal{C}^2$ sur $\mathbb{R}$?

Démonstration.
Soit $x \in \mathbb{R}$. Rappelons que :

$$f_h(x) = \theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt$$

Ainsi, par relation de Chasles :

$$\begin{aligned} f_h(x) = \theta(-x) &\left(\int_{-\infty}^{0} h'(t)\Phi(t)\ dt + \int_{0}^{x} h'(t)\Phi(t)\ dt \right) \\ &+ \theta(x) \left(\int_{x}^{0} h'(t)(1-\Phi(t))\ dt + \int_{0}^{+\infty} h'(t)(1-\Phi(t))\ dt \right) \end{aligned}$$

× Notons $g_1 : x \mapsto \displaystyle\int_0^x h'(t)\Phi(t)\ dt$.
La fonction $t \mapsto h'(t)\Phi(t)$ est continue sur $\mathbb{R}$ donc g_1 est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ par théorème fondamental de l'analyse et, pour tout $x \in \mathbb{R}$:

$$g_1'(x) = h'(x)\Phi(x)$$

× De manière analogue, la fonction $g_2 : x \mapsto \displaystyle\int_x^0 h'(t)(1-\Phi(t))\ dt$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et, pour tout $x \in \mathbb{R}$:

$$g_2'(x) = -h'(x)(1-\Phi(x))$$

Avec ces notations, il vient :

$$f_h(x) = \theta(-x) \left(\int_{-\infty}^{0} h'(t)\Phi(t)\ dt + g_1(x) \right) + \theta(x) \left(\int_{0}^{+\infty} h'(t)(1-\Phi(t))\ dt + g_2(x) \right)$$

$$\boxed{\text{Ainsi, par somme et produit de fonctions de classe } \mathcal{C}^1 \text{ sur } \mathbb{R},\ f_h \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathbb{R}.}$$

En utilisant les questions ***3.a)*** et ***2.c)***, on obtient

$$\begin{aligned} f_h'(x) &= -\theta'(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(-x) g_1'(x) \\ &\quad + \theta'(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt + \theta(x) g_2'(x) \\ &= -\theta'(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(-x) h'(x)\Phi(x) \\ &\quad + \theta'(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt - \theta(x) h'(x)(1-\Phi(x)) \\ &= h'(x) \left(\cancel{\theta(-x)\Phi(x) - \theta(x)(1-\Phi(x))} \right) \\ &\quad - (1 - x\theta(-x)) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + (1 + x\theta(x)) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \\ &= - \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \\ &\quad + x\theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + x\theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \\ &= c_h - h(x) + x f_h(x) \end{aligned}$$

$$\boxed{\text{Ainsi, pour tout } x \text{ réel, } f_h'(x) - x f_h(x) = c_h - h(x).}$$

D'après ce qui précède, f_h' est de la forme $f_h' = c_h - h + g \times f_h$ où
- $\times$ c_h est une constante,
- $\times$ $h \in W$ donc h est de classe $\mathcal{C}^1$ sur $\mathbb{R}$,
- $\times$ f_h est de classe $\mathcal{C}^1$ sur $\mathbb{R}$,
- $\times$ $g : x \mapsto x$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Ainsi, f_h' est de classe $\mathcal{C}^1$ sur $\mathbb{R}$,

$$\boxed{\text{et donc } f_h \text{ est de classe } \mathcal{C}^2 \text{ sur } \mathbb{R}.}$$

Commentaire

Dans cette question, il s'agit de dériver des fonctions de la forme

$$g : x \mapsto \int_{-\infty}^{x} \gamma(t)\ dt$$

où γ est continue sur $\mathbb{R}$. Le théorème fondamental de l'analyse ne permet pas de le faire directement. En effet, ce théorème assure que la fonction

$$g_1 : x \mapsto \int_{c}^{x} \gamma(t)\ dt$$

est dérivable lorsque c est une constante **réelle**. Afin de rester dans le cadre du cours, nous utilisons la relation de Chasles pour écrire :

$$\forall x \in \mathbb{R},\ g(x) = \int_{-\infty}^{0} \gamma(t)\ dt + g_1(x)$$

avec $c = 0$ (nous avons le choix). L'intégrale $\displaystyle\int_{-\infty}^{0} \gamma(t)\ dt$ étant une constante, elle ne pose plus de difficultés techniques pour la dérivation.

$\square$

c) En conclure que, si X est une variable aléatoire admettant une espérance,

$$|\mathbb{E}(h(X)) - \mathbb{E}(h(N))| = \left|\mathbb{E}(f_h'(X) - X f_h(X))\right|$$

Démonstration.

- Tout d'abord, puisque $h \in W$ et puisque X et N admettent chacune une espérance, il vient (d'après le résultat admis en début d'énoncé) que $\mathbb{E}(h(X))$ et $\mathbb{E}(h(N))$ existent.
- Montrons que $\mathbb{E}(f_h'(X) - X f_h(X))$ existe.
  D'après la question précédente, pour tout $x \in \mathbb{R}$,

  $$f_h'(x) - x f_h(x) = c_h - h(x)$$

  On en déduit que

  $$f_h'(X) - X f_h(X) = c_h - h(X)$$

  Or, c_h est une constante et $h(X)$ admet une espérance (d'après le premier point) donc $c_h - h(X)$ admet une espérance par transformation affine.

$$\boxed{\text{On en déduit que } \mathbb{E}(f_h'(X) - X f_h(X)) \text{ existe.}}$$

- On peut alors écrire

$$\begin{aligned}
\left|\mathbb{E}(f_h'(X) - X f_h(X))\right| &= |\mathbb{E}(c_h - h(X))| \\
&= |\mathbb{E}(c_h) - \mathbb{E}(h(X))| && \textit{(par linéarité de l'espérance)} \\
&= |c_h - \mathbb{E}(h(X))| && \textit{(car } c_h \textit{ est une constante)} \\
&= |\mathbb{E}(h(N)) - \mathbb{E}(h(X))| && \textit{(par définition de } c_h\textit{)}
\end{aligned}$$

$$\boxed{\text{Ainsi, on a bien : } |\mathbb{E}(h(X)) - \mathbb{E}(h(N))| = |\mathbb{E}(f_h'(X) - X f_h(X))|.}$$

Commentaire

On retiendra qu'il faut toujours démontrer l'existence de l'espérance d'une variable aléatoire avant de se lancer dans le calcul de celle-ci.

□

4. Majoration de $|f_h|$.

a) Montrer, en utilisant les égalités (R_1), que pour tout x réel :

$$\theta(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} (1 - \Phi(t))\ dt = 1$$

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

$$\begin{aligned}
&\theta(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} (1 - \Phi(t))\ dt \\
&= \theta(-x)\Big(x\Phi(x) + \varphi(x)\Big) + \theta(x)\Big(-x(1 - \Phi(x)) + \varphi(x)\Big) && \textit{(d'après les égalités } R_1\textit{)} \\
&= x\theta(-x)\Phi(x) + \theta(-x)\varphi(x) - x\theta(x)(1 - \Phi(x)) + \theta(x)\varphi(x) \\
&= x\,\cancel{(\theta(-x)\Phi(x) - \theta(x)(1 - \Phi(x)))} + (\theta(-x) + \theta(x))\varphi(x)
\end{aligned}$$

Or,

$$\begin{aligned}
\theta(-x) + \theta(x) &= \frac{\Phi(-x)}{\varphi(-x)} + \frac{\Phi(x)}{\varphi(x)} \\
&= \frac{\Phi(-x)}{\varphi(x)} + \frac{\Phi(x)}{\varphi(x)} && \textit{(car } \varphi \textit{ est paire)} \\
&= \frac{1 - \Phi(x)}{\varphi(x)} + \frac{\Phi(x)}{\varphi(x)} && \textit{(par propriété de } \Phi\textit{)} \\
&= \frac{1}{\varphi(x)}
\end{aligned}$$

Ce qui donne, en multipliant par $\varphi(x)$: $(\theta(-x) + \theta(x))\varphi(x) = 1$.

$$\boxed{\text{Finalement : pour tout } x \text{ réel, } \theta(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} (1 - \Phi(t))\ dt = 1.}$$

□

b) En déduire que pour tout x réel : $|f_h(x)| \leqslant 1$.

Démonstration.
Soit $x \in \mathbb{R}$. Tout d'abord, par définition de f_h :

$$|f_h(x)| = \left| \theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right|$$

En appliquant l'inégalité triangulaire sur les sommes de réels puis sur les intégrales (les bornes étant rangées dans l'ordre croissant), et en utilisant la propriété $|ab| = |a|\,|b|$ sur les réels, on obtient :

$$\begin{aligned}
& \left| \theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right| \\
\leqslant\ & \left| \theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt \right| + \left| \theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right| && \textit{(par inégalité triangulaire sur les réels)} \\
\leqslant\ & |\theta(-x)| \left| \int_{-\infty}^{x} h'(t)\Phi(t)\ dt \right| + |\theta(x)| \left| \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right| && \textit{(valeur absolue d'un produit)} \\
\leqslant\ & |\theta(-x)| \int_{-\infty}^{x} \left| h'(t)\Phi(t) \right|\ dt + |\theta(x)| \int_{x}^{+\infty} \left| h'(t)(1-\Phi(t)) \right|\ dt && \textit{(par inégalité triangulaire sur les intégrales)} \\
\leqslant\ & |\theta(-x)| \int_{-\infty}^{x} \left| h'(t) \right| |\Phi(t)|\ dt + |\theta(x)| \int_{x}^{+\infty} \left| h'(t) \right| |1-\Phi(t)|\ dt && \textit{(valeur absolue d'un produit)}
\end{aligned}$$

Remarquons maintenant que $\theta(-x) \geqslant 0$, $\theta(x) \geqslant 0$, et, pour tout $t \in \mathbb{R}$, $0 \leqslant \Phi(t) \leqslant 1$.
On en déduit que :

$$|f_h(x)| \leqslant \theta(-x) \int_{-\infty}^{x} \left| h'(t) \right| \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} \left| h'(t) \right| (1-\Phi(t))\ dt$$

× On sait que $h \in W$, et donc, pour tout $t \in \mathbb{R}$, $|h'(t)| \leqslant 1$. Il suit que, pour tout $t \in \mathbb{R}$:

$$\left| h'(t) \right| \Phi(t) \leqslant \Phi(t) \qquad \text{et} \qquad \left| h'(t) \right| (1-\Phi(t)) \leqslant 1-\Phi(t)$$

× Par croissance de l'intégrale, les bornes étant rangées dans l'ordre croissant, on obtient :

$$\int_{-\infty}^{x} \left| h'(t) \right| \Phi(t)\ dt \leqslant \int_{-\infty}^{x} \Phi(t)\ dt \qquad \text{et} \qquad \int_{x}^{+\infty} \left| h'(t) \right| (1-\Phi(t))\ dt \leqslant \int_{x}^{+\infty} (1-\Phi(t))\ dt$$

D'où

$$|f_h(x)| \leqslant \theta(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} (1-\Phi(t))\ dt$$

$$\boxed{\text{Ainsi, d'après la question } \textbf{\textit{4.a)}}, \text{ pour tout } x \text{ réel, } |f_h(x)| \leqslant 1.}$$

□

5. *Majoration de* $|f_h''|$.

a) Montrer que pour tout x réel :

$$\theta''(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} (1-\Phi(t))\ dt = 1$$

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

$$\begin{aligned}
&\theta''(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} (1-\Phi(t))\ dt \\
&= \Big(-x + \big(1+(-x)^2\big)\,\theta(-x)\Big) \int_{-\infty}^{x} \Phi(t)\ dt \\
&\qquad + \Big(x + (1+x^2)\theta(x)\Big) \int_{x}^{+\infty} (1-\Phi(t))\ dt && \textit{(d'après la question } \textbf{\textit{3.a)}}\textit{)} \\
&= x \left(-\int_{-\infty}^{x} \Phi(t)\ dt + \int_{x}^{+\infty} (1-\Phi(t))\ dt \right) \\
&\qquad + (1+x^2) \left(\theta(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} (1-\Phi(t))\ dt \right) \\
&= x(-x\Phi(x) - \varphi(x) - x(1-\Phi(x)) + \varphi(x)) && \textit{(d'après les égalités } (R_1)\textit{)} \\
&\qquad + (1+x^2) \times 1 && \textit{(d'après la question } \textbf{\textit{4.a)}}\textit{)} \\
&= x(-\cancel{x\Phi(x)} - \cancel{\varphi(x)} - x + \cancel{x\Phi(x)} + \cancel{\varphi(x)}) + 1 + x^2 \\
&= -x^2 + 1 + x^2 \\
&= 1
\end{aligned}$$

$$\boxed{\text{On a bien : } \forall x \in \mathbb{R},\ \theta''(-x) \int_{-\infty}^{x} \Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} (1-\Phi(t))\ dt = 1.}$$

□

b) Établir pour tout x réel l'égalité :

$$f_h''(x) = -h'(x) + \theta''(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt$$

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

- D'une part, d'après la question ***3.b)***,

$$f_h'(x) = x f_h(x) + c_h - h(x)$$

donc, en dérivant (ce qui est possible car f_h est de classe $\mathcal{C}^2$ sur $\mathbb{R}$) :

$$\begin{aligned}
f_h''(x) &= f_h(x) + x f_h'(x) - h'(x) \\
&= f_h(x) + x(x f_h(x) + c_h - h(x)) - h'(x) \\
&= -h'(x) + (1+x^2) f_h(x) + x(c_h - h(x))
\end{aligned}$$

- D'autre part, en reprenant la trame du calcul de la question précédente :

$$\begin{aligned}
&\theta''(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \\
&= (-x + (1+(-x)^2)\theta(-x)) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt \\
&\quad + (x + (1+x^2)\theta(x)) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt && \textit{(d'après la question } \textbf{\textit{3.a)}}\textit{)} \\
&= x \left(- \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right) \\
&\quad + (1+x^2) \left(\theta(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt \right) \\
&= x(c_h - h(x)) + (1+x^2) f_h(x) && \textit{(d'après la question } \textbf{\textit{2.c)}}\textit{)} \\
&= f_h''(x) + h'(x)
\end{aligned}$$

$$\boxed{\text{D'où : } \forall x \in \mathbb{R},\ f_h''(x) = -h'(x) + \theta''(-x) \int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta''(x) \int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt.}$$

□

c) Étudier les variations sur $\mathbb{R}$ de la fonction $x \mapsto \Phi(x) + \dfrac{x}{1+x^2}\varphi(x)$. En déduire son signe et le signe de θ''.
En conclure que, pour tout x réel : $|f_h''(x)| \leqslant 2$.

Démonstration.
On note $g : x \mapsto \Phi(x) + \dfrac{x}{1+x^2}\varphi(x)$. La fonction g est dérivable sur $\mathbb{R}$ comme somme et produit de fonctions dérivables sur $\mathbb{R}$. Soit $x \in \mathbb{R}$.

$$\begin{aligned}
g'(x) &= \varphi(x) + \frac{1+x^2-2x^2}{(1+x^2)^2}\varphi(x) + \frac{x}{1+x^2}\varphi'(x) \\
&= \varphi(x) + \frac{1-x^2}{(1+x^2)^2}\varphi(x) - \frac{x}{1+x^2} x\varphi(x) \\
&= \frac{\varphi(x)}{(1+x^2)^2} \Big(\left(1+x^2\right)^2 + 1 - x^2 - x^2\left(1+x^2\right) \Big) \\
&= \frac{\varphi(x)}{(1+x^2)^2} \Big(1 + 2x^2 + x^4 + 1 - x^2 - x^2 - x^4 \Big) \\
&= \frac{2\varphi(x)}{(1+x^2)^2} > 0 && \textit{(car } \varphi(x) > 0\textit{)}
\end{aligned}$$

$$\boxed{\text{Donc } g \text{ est strictement croissante sur } \mathbb{R}.}$$

De plus,

× $\lim\limits_{x \to -\infty} \Phi(x) = 0$ car Φ est une fonction de répartition

× $\displaystyle\lim_{x \to -\infty} \frac{x}{1+x^2} = 0$

× $\displaystyle\lim_{x \to -\infty} \varphi(x) = 0$

donc

$$\lim_{x \to -\infty} g(x) = 0 + 0 \times 0 = 0$$

$$\boxed{\text{On en déduit que, pour tout } x \in \mathbb{R},\ g(x) > 0.}$$

Commentaire

Le tableau de variations de g n'est pas explicitement demandé, mais il est possible de l'inclure dans la réponse :

x	$-\infty$ $\qquad$ $+\infty$
Signe de $g'(x)$	$+$
Variations de g	$0 \nearrow 1$

D'autre part,

$$\begin{aligned}
\theta''(x) &= x + \left(1+x^2\right)\theta(x) && \textit{(d'après la question } \textbf{\textit{3.a)}}\textit{)} \\
&= x + \left(1+x^2\right)\frac{\Phi(x)}{\varphi(x)} \\
&= \frac{1+x^2}{\varphi(x)}\left(\frac{x}{1+x^2}\varphi(x) + \Phi(x)\right) \\
&= \frac{1+x^2}{\varphi(x)} g(x)
\end{aligned}$$

$$\boxed{\text{On en déduit que, pour tout } x \in \mathbb{R},\ \theta''(x) > 0.}$$

Ensuite, en majorant de manière analogue aux calculs effectués en question ***4.b)*** :

$$\begin{aligned}
&\left|f_h''(x)\right| \\
&= \left|-h'(x) + \theta''(-x)\int_{-\infty}^{x} h'(t)\Phi(t)\ dt + \theta''(x)\int_{x}^{+\infty} h'(t)(1-\Phi(t))\ dt\right| \\
&\leqslant \left|h'(x)\right| + \left|\theta''(-x)\right|\int_{-\infty}^{x} \left|h'(t)\Phi(t)\right|\ dt + \left|\theta''(x)\right|\int_{x}^{+\infty} \left|h'(t)(1-\Phi(t))\right|\ dt && \textit{(inégalité triangulaire)} \\
&\leqslant 1 + \theta''(-x)\int_{-\infty}^{x} \left|h'(t)\right|\Phi(t)\ dt + \theta''(x)\int_{x}^{+\infty} \left|h'(t)\right|(1-\Phi(t))\ dt && \textit{(car } h \in W,\ \theta''(-x) > 0 \textit{ et } \theta''(x) > 0\textit{)} \\
&\leqslant 1 + \theta''(-x)\int_{-\infty}^{x} \Phi(t)\ dt + \theta''(x)\int_{x}^{+\infty} (1-\Phi(t))\ dt && \textit{(} h \in W \textit{ et croissance de l'intégrale avec les bornes dans l'ordre croissant)} \\
&\leqslant 1 + 1 = 2 && \textit{(d'après la question } \textbf{\textit{5.a)}}\textit{)}
\end{aligned}$$

$$\boxed{\text{On a bien : } \forall x \in \mathbb{R},\ |f_h''(x)| \leqslant 2.}$$

□

Partie 2 - Majoration uniforme de la distance de Kolmogorov

Dans la suite du problème, si X est une variable aléatoire de fonction de répartition F_X, on définit, pour tout x réel, $d_X(x)$ la distance de Kolmogorov au point x entre la loi de X et la loi normale centrée réduite par :

$$d_X(x) = |F_X(x) - \Phi(x)|$$

On définit, pour tout x réel, la fonction h_x sur $\mathbb{R}$ par $h_x(t) = \begin{cases} 1 & \text{si } t \leqslant x \\ 0 & \text{si } t > x \end{cases}$.

On définit aussi la fonction γ sur $\mathbb{R}$ par :

$$\gamma(t) = \begin{cases} 1 & \text{si } t < 0 \\ 0 & \text{si } t > 1 \\ 1 - 3t^2 + 2t^3 & \text{si } t \in [0,1] \end{cases}$$

Soit X une variable aléatoire.

6. Pour tout x réel, quelle est la loi de la variable aléatoire $h_x(X)$? En déduire que $\mathbb{E}(h_x(X))$ existe et vaut $F_X(x)$.

Démonstration.
Soit $x \in \mathbb{R}$. La fonction h_x est à valeurs dans $\{0,1\}$ donc $h_x(X)(\Omega) \subset \{0,1\}$. On en déduit que $h_x(X)$ suit une loi de Bernoulli. Notons r son paramètre. Par définition de r :

$$r = \mathbb{P}([h_x(X) = 1])$$

Montrons que $[h_x(X) = 1] = [X \leqslant x]$.
Soit $\omega \in \Omega$.

$$\begin{aligned} \omega \in [h_x(X) = 1] &\iff h_x(X(\omega)) = 1 \\ &\iff X(\omega) \leqslant x \qquad \textit{(par définition de } h_x(t), \textit{ avec } t = X(\omega)) \\ &\iff \omega \in [X \leqslant x] \end{aligned}$$

Ainsi, on a bien $[h_x(X) = 1] = [X \leqslant x]$, ce qui donne

$$\begin{aligned} r &= \mathbb{P}([h_x(X) = 1]) \\ &= \mathbb{P}([X \leqslant x]) \\ &= F_X(x) \end{aligned}$$

$$\boxed{\text{Finalement, } h_x(X) \hookrightarrow \mathcal{B}(F_X(x)), \text{ donc } h_x(X) \text{ admet une espérance et } \mathbb{E}(h_x(X)) = F_X(x).}$$

□

7. a) Écrire une fonction **Python** `gamma(t)` qui calcule et renvoie la valeur de $\gamma(t)$, t étant donné.

Démonstration.

```
def gamma(t):
    if t < 0:
        return 1
    elif t > 1:
        return 0
    else:
        return 1 - 3 * t**2 + 2 * t**3
```

Commentaire

- La définition de $\gamma(t)$ par cas impose une structure conditionnelle en **Python**.
- En toute rigueur, `gamma(t)` n'est pas une fonction **Python**. Il vaudrait mieux parler de la fonction `gamma` qui prend en paramètre un réel `t`.

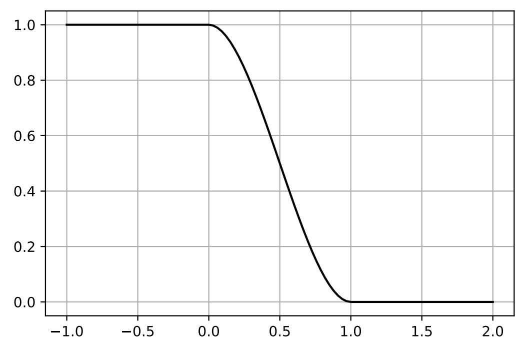
□

b) Utiliser la fonction précédente pour écrire un script qui affiche le graphe de γ sur le segment $[-1, 2]$ dans un repère.

Démonstration.

```
import numpy as np
import matplotlib.pyplot as plt
abscisse = np.linspace(-1,2,100)
ordonnees = [gamma(t) for t in abscisse]
plt.plot(abscisse, ordonnees)
plt.grid()
```

En exécutant ce script, on obtient le graphe :



Commentaire

Rappelons que le graphe $\mathcal{C}_\gamma$ de γ est un ensemble de points du plan dont la définition précise est :

$$\mathcal{C}_\gamma = \{(t, \gamma(t)) \mid t \in \mathbb{R}\}$$

Le tracé d'un graphe en **Python** se fait en discrétisant l'espace de départ (ici le segment $[-1, 2]$) et en reliant un nombre **fini** de points par des segments de droite. Plus précisément, si

$$-1 = x_0 < x_1 < \cdots < x_N = 2$$

est la discrétisation choisie du segment $[-1, 2]$, on relie alors, pour tout $k \in [\![0, N-1]\!]$, les points $(x_k, \gamma(x_k))$ et $(x_{k+1}, \gamma(x_{k+1}))$. On obtient alors une représentation approchée du graphe de γ.
La commande `np.linspace(-1,2,100)` produit le tableau

$$\texttt{[}x_0\texttt{, ..., }x_{99}\texttt{]}$$

contenant 100 nombres rangés dans l'ordre croissant et régulièrement espacés tels que le plus petit nombre est -1 et le plus grand nombre est 2. On a donc, pour tout $k \in [\![0, 98]\!]$,

$$x_{k+1} - x_k = \frac{2-(-1)}{99} = \frac{3}{99} = \frac{1}{33}$$

La commande `[gamma(t) for t in abscisse]` construit alors la liste

$$\texttt{[}\gamma(x_0)\texttt{, ..., }\gamma(x_{99})\texttt{]}$$

La commande `plt.plot(abscisse, ordonnees)` permet le tracé du graphe à partir des points de coordonnées $(x_k, \gamma(x_k))$ précédemment construits.
Pour finir, la commande `plt.grid()` trace des lignes horizontales et verticales permettant de mieux visualiser les valeurs prises par la fonction γ, ce qui donne notre repère.

□

8. a) Montrer que γ est continue sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé de 0 et 1.

Démonstration.

- La fonction γ est de classe $\mathcal{C}^1$:
  - × sur l'intervalle **ouvert** $]-\infty, 0[$ car constante sur cet intervalle
  - × sur l'intervalle **ouvert** $]1, +\infty[$ car constante sur cet intervalle
  - × sur l'intervalle **ouvert** $]0, 1[$ car polynomiale sur cet intervalle

  Ainsi,

  $$\boxed{\gamma \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathbb{R} \text{ privé de 0 et 1.}}$$

Il reste à montrer que γ est continue en 0 et en 1.

- Montrons que la fonction γ est continue en 0 :
  - × $\gamma(0) = 1$
  - × $\lim\limits_{t \to 0^-} \gamma(t) = 1$
  - × $\lim\limits_{t \to 0^+} \gamma(t) = 1$

  $$\boxed{\text{donc } \gamma \text{ est continue en 0.}}$$

- Montrons que la fonction γ est continue en 1 :
  - × $\gamma(1) = 1 - 3 + 2 = 0$
  - × $\lim\limits_{t \to 1^-} \gamma(t) = 1 - 3 + 2 = 0$

× $\lim_{t\to 1^+} \gamma(t) = 0$

$$\boxed{\text{donc } \gamma \text{ est continue en 1.}}$$

□

b) Étudier les variations de γ sur $[0,1]$ et montrer que pour tout $t \in \mathbb{R}$, $\gamma(t) \in [0,1]$.

Démonstration.
D'après la question précédente, γ est dérivable sur $]0,1[$. Soit $t \in]0,1[$.

$$\begin{aligned} \gamma'(t) &= -6t + 6t^2 \\ &= 6t(t-1) \end{aligned}$$

Comme $6t > 0$ et $t - 1 < 0$, alors $\gamma'(t) < 0$.
On en déduit que γ est strictement décroissante sur $]0,1[$.
De plus, γ est continue sur $[0,1]$,

$$\boxed{\text{donc } \gamma \text{ est strictement décroissante sur } [0,1].}$$

Le tableau de variations de γ est :

t	0		1
Signe de $\gamma'(t)$	$\Vert$	$-$	$\Vert$
Variations de γ	1	$\searrow$	0

D'après ce tableau de variations, pour tout $t \in [0,1]$, $\gamma(t) \in [0,1]$.
De plus, pour tout $t < 0$, $\gamma(t) = 1 \in [0,1]$ et pour tout $t > 1$, $\gamma(t) = 0 \in [0,1]$.

$$\boxed{\text{D'où, pour tout } t \in \mathbb{R},\ \gamma(t) \in [0,1].}$$

□

c) Établir que γ est dérivable en 1 et que $\gamma'(1) = 0$.
On montrerait de même que γ est dérivable en 0 et que $\gamma'(0) = 0$. On l'admet.

Démonstration.
Soit $t \neq 1$. Puisque $\gamma(1) = 0$, on a $\dfrac{\gamma(t) - \gamma(1)}{t-1} = \dfrac{\gamma(t)}{t-1}$.
Deux cas se présentent.

- Premier cas : $t > 1$

$$\frac{\gamma(t) - \gamma(1)}{t-1} = \frac{0}{t-1} = 0$$

donc $\lim_{t\to 1^+} \dfrac{\gamma(t) - \gamma(1)}{t-1} = 0$.

- Deuxième cas : $0 < t < 1$

$$\begin{aligned}\frac{\gamma(t)-\gamma(1)}{t-1} &= \frac{1-3t^2+2t^3}{t-1}\\ &= \frac{(t-1)(2t^2-t-1)}{t-1}\\ &= 2t^2-t-1\end{aligned}$$

donc $\displaystyle\lim_{t\to 1^-} \frac{\gamma(t)-\gamma(1)}{t-1} = 2-1-1 = 0.$

Finalement : $\displaystyle\lim_{t\to 1} \frac{\gamma(t)-\gamma(1)}{t-1} = 0$

$$\boxed{\text{donc } \gamma \text{ est dérivable en 1 et } \gamma'(1) = 0.}$$

□

d) Justifier que γ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et que pour tout t réel $|\gamma'(t)| \leqslant \dfrac{3}{2}$.

Démonstration.
On a vu dans les questions précédentes que

× γ est dérivable sur $\mathbb{R}$

× γ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé de 0 et 1.

Il reste à montrer que γ' est continue en 0 et en 1.

- Montrons que la fonction γ' est continue en 0 :
  × pour tout $t < 0$, $\gamma'(t) = 0$ donc $\displaystyle\lim_{t\to 0^-} \gamma'(t) = 0 = \gamma'(0)$
  × pour tout $0 < t < 1$, $\gamma'(t) = 6t(t-1)$ donc $\displaystyle\lim_{t\to 0^+} \gamma'(t) = 0 = \gamma'(0)$
- Montrons que la fonction γ' est continue en 1 :
  × pour tout $0 < t < 1$, $\gamma'(t) = 6t(t-1)$ donc $\displaystyle\lim_{t\to 1^-} \gamma'(t) = 0 = \gamma'(1)$
  × pour tout $t > 1$, $\gamma'(t) = 0$ donc $\displaystyle\lim_{t\to 1^+} \gamma'(t) = 0 = \gamma'(1)$

$$\boxed{\text{Donc } \gamma \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathbb{R}.}$$

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

× Si $t \leqslant 0$, alors $\gamma'(t) = 0$ et donc $|\gamma'(t)| \leqslant \dfrac{3}{2}$.

× Si $t \geqslant 1$, alors $\gamma'(t) = 0$ et donc $|\gamma'(t)| \leqslant \dfrac{3}{2}$.

× Si $0 < t < 1$, alors $\gamma'(t) = 6t(t-1)$.
  Or, l'inégalité $p(1-p) \leqslant \dfrac{1}{4}$ valable pour tout $p \in [0,1]$ est classique et on en déduit :

$$\begin{aligned}\left|\gamma'(t)\right| &= |6t(t-1)| = 6\,|t|\,|1-t|\\ &= 6t(1-t) && \textit{(car } 0 < t < 1\textit{)}\\ &\leqslant 6 \times \frac{1}{4}\\ &= \frac{3}{2}\end{aligned}$$

$$\boxed{\text{Finalement, on a bien, pour tout } t \text{ réel, } |\gamma'(t)| \leqslant \frac{3}{2}.}$$

Commentaire

La question ***8*** est découpée en quatre sous-questions qui permettent de bien guider les candidat·es. Elle ne présente aucune difficulté particulière pour un·e élève sérieux·se qui a travaillé ses définitions de première année. Il faut apprendre à reconnaître de telles questions qui peuvent être « cachées » au milieu de questions beaucoup plus difficiles. Puisque le raisonnement est guidé et que les résultats sont dans l'énoncé, on peut s'attendre à ce que la majeure partie du barème soit allouée à la précision de l'argumentation. Il est impensable d'arriver aux concours sans savoir étudier avec rigueur la régularité d'une fonction définie par morceaux.

□

On suppose dans la suite de cette partie que X admet une espérance et on considère un réel M_X qui vérifie, pour tout $h \in W$, $|\mathbb{E}(h(X)) - \mathbb{E}(h(N))| \leqslant M_X$.

Commentaire

Détaillons un peu l'argument de la preuve de l'existence de M_X. On suppose seulement que X admet une espérance. Soit $h \in W$. D'après la question ***2.a)***,

$$|h(X) - h(N)| \leqslant |X - N| \leqslant |X| + |N|$$

On remarque ensuite :

× si Y est une variable aléatoire, alors Y admet une espérance si et seulement si $|Y|$ admet une espérance.

(Le programme d'ECG maths appliquées permet de faire la preuve dans le cas discret ou le cas à densité, mais pas dans le cas général. Il s'agit essentiellement d'une traduction de la définition qui demande une convergence absolue.)

× en cas d'existence, par inégalité triangulaire :

$$|\mathbb{E}(Y)| \leqslant \mathbb{E}(|Y|)$$

(Ici l'ingrédient essentiel est la croissance de l'espérance, qui est valable dans le cas général)

Puisque X et N admettent chacune une espérance, toutes les variables aléatoires en jeu admettent une espérance (en particulier, on utilise le résultat admis en début d'énoncé pour $h(X)$ et $h(N)$) et on peut écrire, par linéarité et par croissance de l'espérance :

$$\begin{aligned} |\mathbb{E}(h(X)) - \mathbb{E}(h(N))| &= |\mathbb{E}(h(X) - h(N))| \\ &\leqslant \mathbb{E}(|h(X) - h(N)|) \\ &\leqslant \mathbb{E}(|X| + |N|) \\ &\leqslant \mathbb{E}(|X|) + \mathbb{E}(|N|) \end{aligned}$$

On pose alors $M_X = \mathbb{E}(|X|) + \mathbb{E}(|N|)$, qui ne dépend pas de h et qui vérifie bien la propriété voulue. Le concepteur a sans doute fait le choix de ne pas tester les candidat·es sur ce raisonnement abstrait, préférant poser d'autres questions. Il aurait peut être été préférable d'écrire, pour éviter toute ambiguïté :

On suppose dans la suite de cette partie que X admet une espérance. On admet alors qu'il existe un réel M_X qui vérifie, pour tout $h \in W$, $|\mathbb{E}(h(X)) - \mathbb{E}(h(N))| \leqslant M_X$.

9. Soit $t > 0$ et x un réel. Pour tout $y \in \mathbb{R}$, on pose $k_x(y) = \gamma\left(\dfrac{y-x}{t}\right)$.

a) Montrer que pour tout y réel, $h_x(y) \leqslant k_x(y)$.

Démonstration.
Soit $y \in \mathbb{R}$. Deux cas se présentent.

- Premier cas : $y \leqslant x$. Alors
  - $\times$ $h_x(y) = 1$.
  - $\times$ $y - x \leqslant 0$ et $t > 0$ donc $\dfrac{y-x}{t} \leqslant 0$ donc $k_x(y) = \gamma\left(\dfrac{y-x}{t}\right) = 1$ par définition de γ.

  On a bien $h_x(y) \leqslant k_x(y)$.
- Deuxième cas : $y > x$. Alors
  - $\times$ $h_x(y) = 0$.
  - $\times$ pour tout $u \in \mathbb{R}$, $\gamma(u) \in [0,1]$ (d'après la question ***8.b)***) donc $k_x(y) = \gamma\left(\dfrac{y-x}{t}\right) \geqslant 0$.

  On a bien $h_x(y) \leqslant k_x(y)$.

$$\boxed{\text{Pour tout } y \text{ réel, } h_x(y) \leqslant k_x(y).}$$

□

b) *On admet l'existence de* $\mathbb{E}(k_x(X))$ *et de* $\mathbb{E}(k_x(N))$. Justifier l'inégalité suivante :

$$\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leqslant \mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N)) + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N))$$

Commentaire

On peut démontrer l'existence de $\mathbb{E}(k_x(X))$ et de $\mathbb{E}(k_x(N))$ en utilisant le résultat admis dans le préambule du sujet. On démontre un peu plus loin que $g : u \mapsto \dfrac{2t}{3} k_x(u)$ est une fonction qui appartient à W. Ce résultat pourrait être démontré dès maintenant. On en déduit alors que $g(X)$ et $g(N)$ existent. Or, $k_x(X)$ (resp. $k_x(N)$) est une transformée affine de $g(X)$ (resp. $g(N)$) donc $k_x(X)$ et $k_x(N)$ admettent également une espérance.

Démonstration.
Soit $\omega \in \Omega$. D'après la question précédente :

$$\forall y \in \mathbb{R},\ h_x(y) \leqslant k_x(y)$$

En évaluant en $y = X(\omega)$, on obtient :

$$h_x(X(\omega)) \leqslant k_x(X(\omega))$$

Ceci étant vrai pour tout $\omega \in \Omega$:

$$[h_x(X) \leqslant k_x(X)] = \Omega$$

et, a fortiori :

$$\mathbb{P}([h_x(X) \leqslant k_x(X)]) = 1$$

Ainsi, par croissance de l'espérance :

$$\mathbb{E}(h_x(X)) \leqslant \mathbb{E}(k_x(X))$$

On en déduit que

$$\begin{aligned}\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) &\leqslant \mathbb{E}(k_x(X)) - \mathbb{E}(h_x(N)) \\ &= \mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N)) + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N))\end{aligned}$$

Commentaire

Il peut paraître assez surprenant de rajouter un terme « qui ne sert à rien » de manière artificielle à la toute fin, alors qu'on aurait pu s'arrêter à la majoration

$$\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leqslant \mathbb{E}(k_x(X)) - \mathbb{E}(h_x(N))$$

Cette opération est en fait assez courante lorsque l'on souhaite majorer une différence de deux termes qui diffèrent de deux paramètres (premier paramètre : la fonction k_x ou la fonction h_x, deuxième paramètre : la variable aléatoire X ou la variable aléatoire N). On diminue la difficulté en ne faisant varier qu'un seul paramètre à la fois, mais on a alors deux termes à majorer. C'est ce que fait le sujet en nous faisant majorer le terme $\mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N))$ dès la question suivante.
De manière plus abstraite, on écrit souvent

$$f(x,y) - f(a,b) = (f(x,y) - f(x,b)) + (f(x,b) - f(a,b))$$

et on majore alors le terme $f(x,y) - f(x,b)$ puis le terme $f(x,b) - f(a,b)$ afin d'obtenir une majoration de $f(x,y) - f(a,b)$.

$\square$

c) Montrer que $\mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) = \displaystyle\int_x^{x+t} k_x(u)\varphi(u)du$.

Démonstration.
Calculons tout d'abord les deux termes $\mathbb{E}(k_x(N))$ et $\mathbb{E}(h_x(N))$ séparément.

- La fonction k_x est continue sur $\mathbb{R}$ comme composée de γ et d'une fonction affine toutes les deux continues sur $\mathbb{R}$ (d'après la question ***8.a)***). De plus, il est admis que $k_x(N)$ admet une espérance. Par théorème de transfert, il vient que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} k_x(u)\varphi(u)du$ est absolument convergente et vérifie :
$$\mathbb{E}(k_x(N)) = \int_{-\infty}^{+\infty} k_x(u)\varphi(u)du$$
Pour tout $u \in \mathbb{R}$:
$$\begin{aligned} k_x(u) = 0 &\iff \gamma\left(\frac{u-x}{t}\right) = 0 \\ &\iff \frac{u-x}{t} \geqslant 1 && \textit{(d'après la question } \textbf{\textit{8.}}\textit{)} \\ &\iff u - x \geqslant t && \textit{(car } t > 0\textit{)} \\ &\iff u \geqslant x + t \end{aligned}$$
donc
$$\mathbb{E}(k_x(N)) = \int_{-\infty}^{x+t} k_x(u)\varphi(u)du$$
- D'autre part :
$$\begin{aligned} \mathbb{E}(h_x(N)) &= F_N(x) && \textit{(d'après la question } \textbf{\textit{6.}}\textit{)} \\ &= \Phi(x) \\ &= \int_{-\infty}^{x} \varphi(u)du \end{aligned}$$

Pour tout $u \in \mathbb{R}$:

$$\begin{aligned}
k_x(u) = 1 &\iff \gamma\left(\frac{u-x}{t}\right) = 1 \\
&\iff \frac{u-x}{t} \leqslant 0 && \textit{(d'après la question 8.)} \\
&\iff u - x \leqslant 0 && \textit{(car } t > 0\textit{)} \\
&\iff u \leqslant x
\end{aligned}$$

donc

$$\mathbb{E}(h_x(N)) = \int_{-\infty}^{x} 1 \times \varphi(u)du = \int_{-\infty}^{x} k_x(u)\varphi(u)du$$

Finalement,

$$\boxed{\mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) = \int_{-\infty}^{x+t} k_x(u)\varphi(u)du - \int_{-\infty}^{x} k_x(u)\varphi(u)du = \int_{x}^{x+t} k_x(u)\varphi(u)du}$$

Commentaire

Cette question est difficile car elle demande une prise d'initiative : faire apparaître le terme $k_x(u)$ là où il n'est pas naturellement.
La première idée doit être de penser au théorème de transfert lorsqu'on doit montrer une égalité entre une espérance et une intégrale. A partir de là, il faut s'efforcer de faire apparaître les termes de la formule de fin. Une bonne approche est de conjecturer des égalités au brouillon qui permettraient de faire le lien entre la formule que l'on a démontré et celle que l'on cherche à obtenir, puis de vérifier que ces conjectures sont vraies.

□

d) Établir que la fonction g, définie sur $\mathbb{R}$ par $g : u \mapsto \frac{2t}{3}k_x(u)$, appartient à W. En déduire que :

$$\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leqslant \frac{3}{2t}M_X + \frac{t}{\sqrt{2\pi}} \leqslant \frac{3}{2t}M_X + \frac{t}{2}$$

Démonstration.
La fonction k_x est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme composée de γ et d'une fonction affine toutes les deux de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (d'après la question ***8.d)***). Ainsi, g est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme transformée affine de k_x. Soit $u \in \mathbb{R}$.

$$\begin{aligned}
\left|g'(u)\right| &= \left|\frac{2t}{3}k_x'(u)\right| \\
&= \left|\frac{2t}{3}\frac{1}{t}\gamma'\left(\frac{u-x}{t}\right)\right| \\
&= \frac{2}{3}\left|\gamma'\left(\frac{u-x}{t}\right)\right| \\
&\leqslant \frac{2}{3} \times \frac{3}{2} && \textit{(d'après la question 8.d)} \\
&= 1
\end{aligned}$$

$$\boxed{\text{donc } g \in W.}$$

D'après l'énoncé, pour tout $h \in W$, $|\mathbb{E}(h(X)) - \mathbb{E}(h(N))| \leqslant M_X$.
On applique ce résultat à la fonction g :

$$|\mathbb{E}(g(X)) - \mathbb{E}(g(N))| \leqslant M_X$$

De plus,

$$\begin{aligned} |\mathbb{E}(g(X)) - \mathbb{E}(g(N))| &= \left|\mathbb{E}\left(\frac{2t}{3}k_x(X)\right) - \mathbb{E}\left(\frac{2t}{3}k_x(N)\right)\right| \\ &= \frac{2t}{3}\left|\mathbb{E}\left(k_x(X)\right) - \mathbb{E}\left(k_x(N)\right)\right| \qquad \textit{(par linéarité de l'espérance et car } t > 0\textit{)} \end{aligned}$$

On en déduit :

$$\left|\mathbb{E}\left(k_x(X)\right) - \mathbb{E}\left(k_x(N)\right)\right| \leqslant \frac{3}{2t}M_X$$

D'après les questions ***9.b)***, ***9.c)*** et ce qui précède :

$$\begin{aligned} \mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) &\leqslant \mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N)) + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) \\ &\leqslant |\mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N))| + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) \\ &\leqslant \frac{3}{2t}M_X + \int_x^{x+t} k_x(u)\varphi(u)du \end{aligned}$$

Il reste à montrer : $\displaystyle\int_x^{x+t} k_x(u)\varphi(u)du \leqslant \frac{t}{\sqrt{2\pi}}$.
On le fait en appliquant à deux reprises la croissance de l'intégrale, les bornes étant rangées dans l'ordre croissant ($t > 0$ donc $x \leqslant x + t$) :

$$\begin{aligned} \int_x^{x+t} k_x(u)\varphi(u)du &\leqslant \left|\int_x^{x+t} k_x(u)\varphi(u)du\right| \\ &\leqslant \int_x^{x+t} |k_x(u)\varphi(u)|\,du \qquad \textit{(par inégalité triangulaire et car } x \leqslant x+t\textit{)} \\ &\leqslant \int_x^{x+t} |k_x(u)|\,|\varphi(u)|\,du \\ &\leqslant \int_x^{x+t} |k_x(u)|\,\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u^2}{2}}du \\ &\leqslant \int_x^{x+t} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u^2}{2}}du \qquad \textit{(}|k_x(u)| \leqslant 1 \textit{ d'après la question } \boldsymbol{8.b)}\textit{)} \\ &\leqslant \int_x^{x+t} \frac{1}{\sqrt{2\pi}}du \qquad \textit{(}-\frac{u^2}{2} \leqslant 0 \textit{ donc } e^{-\frac{u^2}{2}} \leqslant 1\textit{)} \\ &= \frac{1}{\sqrt{2\pi}}((x+t) - x) \\ &= \frac{t}{\sqrt{2\pi}} \end{aligned}$$

$$\boxed{\text{D'où : } \mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leqslant \frac{3}{2t}M_X + \frac{t}{\sqrt{2\pi}}.}$$

Ensuite, $\pi \geqslant 3$ donc $2\pi \geqslant 6 \geqslant 4$ donc $\sqrt{2\pi} \geqslant \sqrt{4} = 2$ donc $\dfrac{t}{\sqrt{2\pi}} \leqslant \dfrac{t}{2}$ (car $t > 0$).

$$\boxed{\text{On a bien : } \frac{3}{2t}M_X + \frac{t}{\sqrt{2\pi}} \leqslant \frac{3}{2t}M_X + \frac{t}{2}.}$$

□

On admet de même, qu'en utilisant la fonction k_{x-t}, on a :

$$\mathbb{E}(h_x(N)) - \mathbb{E}(h_x(X)) \leqslant \frac{3}{2t}M_X + \frac{t}{2}$$

10. En étudiant la fonction $t \mapsto \dfrac{3}{2t}M_X + \dfrac{t}{2}$ sur $]0, +\infty[$, en déduire que, pour tout x réel,

$$|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leqslant \sqrt{3M_X}, \text{ puis que } d_X(x) \leqslant \sqrt{3M_X} \qquad (R_2)$$

Démonstration.
Notons $\psi : t \mapsto \dfrac{3}{2}M_X\dfrac{1}{t} + \dfrac{1}{2}t$. La fonction ψ est dérivable sur $]0, +\infty[$. Soit $t \in]0, +\infty[$.

$$\begin{aligned}\psi'(t) &= -\frac{3}{2t^2}M_X + \frac{1}{2} \\ &= \frac{t^2 - 3M_X}{2t^2} \\ &= \frac{(t - \sqrt{3M_X})(t + \sqrt{3M_X})}{2t^2} && \textit{(car } M_X \geqslant 0\textit{)}\end{aligned}$$

Ainsi,

$$\begin{aligned}\psi'(t) \geqslant 0 &\iff \frac{(t - \sqrt{3M_X})(t + \sqrt{3M_X})}{2t^2} \geqslant 0 \\ &\iff t - \sqrt{3M_X} \geqslant 0 && \textit{(car } 2t^2 > 0 \textit{ et } t + \sqrt{3M_X} \geqslant 0\textit{)} \\ &\iff t \geqslant \sqrt{3M_X}\end{aligned}$$

et, de plus,

$$\psi(\sqrt{3M_X}) = \frac{3M_X}{2\sqrt{3M_X}} + \frac{\sqrt{3M_X}}{2} = \frac{\sqrt{3M_X}}{2} + \frac{\sqrt{3M_X}}{2} = \sqrt{3M_X}$$

Le tableau de variations de ψ est donc :

t	0		$\sqrt{3M_X}$		$+\infty$
Signe de $\psi'(t)$	‖	$-$	0	$+$	
Variations de ψ	‖ $+\infty$	↘	$\sqrt{3M_X}$	↗	$+\infty$

Soit $x \in \mathbb{R}$. D'après la question ***9.d)***, l'encadrement

$$-\psi(t) \leqslant \mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leqslant \psi(t)$$

est valable pour tout réel $t > 0$.

Autrement dit, l'inégalité

$$|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leqslant \psi(t)$$

est valable pour tout réel $t > 0$.
On peut donc choisir $t = \sqrt{3M_X}$ (avec $M_X > 0$) et obtenir :

$$\boxed{|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leqslant \psi(\sqrt{3M_X}) = \sqrt{3M_X}}$$

Il reste à interpréter le terme de gauche. D'après la question ***6.***,

$$\mathbb{E}(h_x(X)) = F_X(x) \qquad \text{et} \qquad \mathbb{E}(h_x(N)) = F_N(x) = \Phi(x)$$

$$\boxed{\text{d'où, pour tout } x \text{ réel, } d_X(x) = |F_X(x) - \Phi(x)| = |\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leqslant \sqrt{3M_X}.}$$

Commentaire

Insistons sur ce qui nous apparaît être une subtilité dans cette question. On pourrait croire à la première lecture de l'énoncé que l'étude de la fonction ψ va nous donner :

$$\forall t \in \mathbb{R}, \psi(t) \leqslant \sqrt{3M_X}$$

Après tout, un tel résultat permettrait de conclure à l'inégalité voulue, et beaucoup de questions de majorations se résolvent ainsi. C'est pourtant le contraire que l'on trouve :

$$\forall t \in \mathbb{R}, \psi(t) \geqslant \sqrt{3M_X}$$

Lorsque l'on trouve le contraire de ce que l'on escomptait, il faut se relire pour débusquer une éventuelle erreur de signe. Ici, il ne s'agit pourtant pas d'une erreur. Il faut alors réfléchir aux rôles que jouent les différentes variables en présence. La quantité $|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))|$ dépend de x, de X et de N, mais pas de t. Le réel t a été fixé de manière arbitraire dans $]0, +\infty[$ au début de la question ***9***. Ainsi, on peut reformuler le résultat de la question ***9.d)*** en : pour tout $t > 0$, $|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leqslant \psi(t)$. On **choisit** alors le réel t de sorte à obtenir le plus petit majorant possible. D'après l'étude de la fonction ψ, il s'agit du réel $t = \sqrt{3M_X}$ (c'est le point en lequel ψ atteint son minimum).

□

Partie 3 - Estimation d'une densité

On considère X une variable aléatoire à densité de fonction de répartition F et de densité de probabilité f qui dépendent d'un paramètre inconnu θ, où $\theta \in \Theta$, Θ un intervalle de $\mathbb{R}$.

Soit a un point de continuité de f, fixé. On souhaite estimer $f(a)$.

Par exemple, si X suit la loi exponentielle de paramètre θ et $a > 0$, on souhaite estimer $\theta e^{-\theta a}$.

On dispose pour tout $\theta \in \Theta$, d'une suite de variables aléatoires $(X_i)_{i \geqslant 1}$ indépendantes de même loi que X.

On choisit une suite $(h_n)_{n \geqslant 1}$ de réels strictement positifs tels que :

$$\lim_{n \to +\infty} h_n = 0 \text{ et } \lim_{n \to +\infty} n h_n = +\infty$$

Pour tout $n \in \mathbb{N}^*$, et $\omega \in \Omega$, on définit :

$$C_n(\omega) \text{ comme le nombre d'indices } i \in [\![1, n]\!] \text{ tels que } X_i(\omega) \in\,]a - h_n, a + h_n]$$

et $f_n(\omega) = \dfrac{1}{2nh_n} C_n(\omega)$.

Commentaire

Deux difficultés se présentent en ce début de partie 3.

- L'objet C_n n'est pas défini comme une variable aléatoire mais comme une application définie sur Ω. Il aurait sans doute été judicieux d'admettre explicitement qu'on définit bien ainsi une variable aléatoire.
- La notation f_n peut faire penser à une fonction d'une variable réelle (d'ailleurs la notation f est utilisée pour parler d'une densité donc d'une fonction d'une variable réelle), mais c'est bien à nouveau une variable aléatoire qui est définie ici. De plus, si doute il y avait, la question ***13.*** parle de son espérance et de sa variance. Lorsqu'une notation n'est pas usuelle, il peut être pertinent d'écrire explicitement de quel type d'objet il s'agit à la première lecture du sujet, pour éviter de futures confusions.

11. On suppose que l'on dispose d'un fichier `stats.csv` qui comporte une colonne nommée `salaire`. On considère que les valeurs de cette colonne constituent la réalisation d'un échantillon de la loi de X dont la taille dépasse 10000.

a) Après avoir exécuté `import pandas as pd`, quelle(s) instruction(s) permet(tent) de lire dans le fichier `stats.csv` les valeurs de la colonne `salaire` et d'affecter cette série `pandas` obtenue à une variable `échantillon` ?

On supposera que le fichier `stats.csv` se trouve dans le répertoire de travail.

Démonstration.

```
donnees = pd.read_csv('stats.csv')
échantillon = donnees['salaire']
```

□

b) On souhaite calculer et afficher $f_n(\omega)$ pour a donné, lorsque la réalisation d'un échantillon $(X_1(\omega), \dots, X_n(\omega))$ de la loi de X est représentée en **Python** par `échantillon` et, pour tout $n \in \mathbb{N}^*$, $h_n = \dfrac{1}{\sqrt{n}}$.

Compléter le script suivant pour qu'il réalise cette tâche.

```
a = float(input('a='))
n = échantillon.count()
h = 1 / np.sqrt(n)
C = 0
for i in range(n):
    if ... and ...:
        ... += 1
print(C / ...)
```

Démonstration.

```
a = float(input('a='))
n = échantillon.count()
h = 1 / np.sqrt(n)
C = 0
for i in range(n):
    if échantillon[i] > a - h and échantillon[i] <= a + h:
        C += 1
print(C / (2 * n * h))
```

Commentaire

Il est indiqué dans le programme officiel que les commandes de la librairie `pandas` ne sont pas exigibles. On peut s'attendre à ce que la majorité des points soient donnés si la syntaxe est incorrecte mais que la compréhension du fonctionnement du script est bonne.

□

12. Montrer que C_n suit une loi binomiale de paramètres (n, p_n) en précisant l'expression de p_n en fonction de a et h_n.
En déduire que $\mathbb{E}(f_n)$ existe et vaut : $\dfrac{F(a+h_n) - F(a-h_n)}{2h_n}$.

Démonstration.
Notons, pour tout $i \in [\![1, n]\!]$, Z_i la variable aléatoire définie par :

$$\forall \omega \in \Omega,\ Z_i(\omega) = \begin{cases} 1 & \text{si } X_i(\omega) \in \,]a-h_n, a+h_n] \\ 0 & \text{sinon} \end{cases}$$

× Les variables aléatoires Z_i suivent toutes la même loi de Bernoulli, de paramètre

$$\begin{aligned} p_n &= \mathbb{P}\left([Z_i = 1]\right) \\ &= \mathbb{P}\left([X_i \in \,]a-h_n, a+h_n]]\right) \\ &= \mathbb{P}\left([X \in \,]a-h_n, a+h_n]]\right) && \textit{(car X_i suit la même loi que X)} \\ &= \mathbb{P}([a-h_n < X \leqslant a+h_n]) \\ &= F(a+h_n) - F(a-h_n) \end{aligned}$$

× Les variables aléatoires $(X_i)_{i\geqslant 1}$ sont mutuellement indépendantes, donc, par lemme des coalitions, les variables aléatoires $(Z_i)_{1\leqslant i\leqslant n}$ sont également mutuellement indépendantes.

On remarque maintenant que $C_n = \sum\limits_{i=1}^{n} Z_i$ et donc, d'après les deux points précédents :

$$\boxed{C_n \hookrightarrow \mathcal{B}(n, p_n) \text{ où } p_n = F(a+h_n) - F(a-h_n).}$$

La variable aléatoire C_n admet une espérance. On en déduit que f_n admet une espérance en tant que transformée affine de C_n et

$$\begin{aligned}\mathbb{E}(f_n) &= \mathbb{E}\left(\frac{1}{2nh_n}C_n\right)\\ &= \frac{1}{2nh_n}\mathbb{E}(C_n) && \textit{(par linéarité)}\\ &= \frac{1}{2\cancel{n}h_n}\cancel{n}p_n\\ &= \frac{F(a+h_n)-F(a-h_n)}{2h_n}\end{aligned}$$

Commentaire

On peut faire dans cette question un peu de rétro-ingénierie et s'aider du résultat donné dans l'énoncé pour trouver p_n si on bloque sur le calcul.
L'énoncé nous donne : $\mathbb{E}(f_n) = \dfrac{F(a+h_n)-F(a-h_n)}{2h_n}$.
On cherche p_n et on sait que $\mathbb{E}(C_n) = np_n$. Or, $f_n = \dfrac{1}{2nh_n}C_n$ donc $C_n = 2nh_nf_n$.
D'où, en passant à l'espérance :

$$np_n = 2nh_n\frac{F(a+h_n)-F(a-h_n)}{2h_n} = n(F(a+h_n)-F(a-h_n))$$

ce qui donne :

$$p_n = F(a+h_n)-F(a-h_n)$$

□

13. a) En utilisant la dérivabilité de F en a, montrer que $\lim\limits_{n\to+\infty} \mathbb{E}(f_n) = f(a)$.

Démonstration.
La fonction f est continue en a donc F est dérivable en a. La fonction F admet alors un développement limité à l'ordre 1 au point a qui s'écrit :

$$F(a+h) = F(a) + hF'(a) + \underset{h\to 0}{o}(h) = F(a) + hf(a) + \underset{h\to 0}{o}(h)$$

Puisque $h_n \underset{n\to+\infty}{\longrightarrow} 0$, on peut écrire :

$$F(a+h_n) = F(a) + h_nf(a) + \underset{n\to+\infty}{o}(h_n)$$

$$\text{et} \qquad F(a-h_n) = F(a) - h_nf(a) + \underset{n\to+\infty}{o}(h_n)$$

Par différence :

$$F(a+h_n)-F(a-h_n) = 2h_nf(a) + \underset{n\to+\infty}{o}(h_n)$$

Ainsi, en multipliant par $\dfrac{1}{2h_n}$:

$$\begin{aligned}\mathbb{E}(f_n) &= \frac{F(a+h_n)-F(a-h_n)}{2h_n}\\ &= \frac{2h_nf(a) + \underset{n\to+\infty}{o}(h_n)}{2h_n}\\ &= f(a) + \underset{n\to+\infty}{o}(1)\end{aligned}$$

$$\boxed{\text{donc on a bien : } \mathbb{E}(f_n) \underset{n\to+\infty}{\longrightarrow} f(a).}$$

□

b) Montrer que $\mathbb{V}(f_n)$ existe et que $\lim\limits_{n\to+\infty} \mathbb{V}(f_n) = 0$.

Démonstration.
La variable aléatoire C_n admet une variance. On en déduit que f_n admet une variance en tant que transformée affine de C_n. De plus :

$$\begin{aligned}
\mathbb{V}(f_n) &= \mathbb{V}\left(\frac{1}{2nh_n} C_n\right) \\
&= \frac{1}{(2nh_n)^2} \mathbb{V}(C_n) && \textit{(par propriété de la variance)} \\
&= \frac{1}{(2nh_n)^2} np_n(1-p_n) && \textit{(car } C_n \hookrightarrow \mathcal{B}(n, p_n)\textit{)} \\
&= \frac{1-p_n}{2nh_n} \frac{p_n}{2h_n} \\
&= \frac{1-p_n}{2nh_n} \mathbb{E}(f_n)
\end{aligned}$$

Or, on a vu à la question ***13.a)*** que $p_n \underset{n\to+\infty}{\longrightarrow} 0$ et $\mathbb{E}(f_n) \underset{n\to+\infty}{\longrightarrow} f(a)$.
De plus, par hypothèse, $nh_n \underset{n\to+\infty}{\longrightarrow} +\infty$.

$$\boxed{\text{Il vient alors : } \mathbb{V}(f_n) \underset{n\to+\infty}{\longrightarrow} 0 \times f(a) = 0.}$$

□

On suppose désormais, que $f(a) > 0$, que pour tout $n \in \mathbb{N}^*$, $p_n \in \,]0,1[$, que F est de classe $\mathcal{C}^2$ au voisinage de a et que $\lim\limits_{n\to+\infty} nh_n^3 = 0$.

On note pour tout $n \geqslant 1$, $\sigma_n = \sqrt{np_n(1-p_n)}$ et $\theta_n = \sqrt{2h_n f(a)}$.

On définit les variables aléatoires : $D_n = \dfrac{C_n - np_n}{\sigma_n}$ et $\hat{f}_n = \dfrac{\theta_n \sqrt{n}}{f(a)}(f_n - f(a))$.

14. a) En utilisant le développement limité de F à l'ordre 2 au point a, montrer que :

$$p_n \underset{n\to+\infty}{=} 2h_n f(a) + o(h_n^2) \underset{n\to+\infty}{=} \theta_n^2 + o(h_n^2)$$

Démonstration.
F est de classe $\mathcal{C}^2$ au voisinage de a donc admet un développement limité d'ordre 2 au point a qui s'écrit :

$$\begin{aligned}
F(a+h) &= F(a) + hF'(a) + \frac{1}{2}h^2 F''(a) + \underset{h\to 0}{o}(h^2) \\
&= F(a) + hf(a) + \frac{1}{2}h^2 F''(a) + \underset{h\to 0}{o}(h^2)
\end{aligned}$$

Puisque $h_n \underset{n\to+\infty}{\longrightarrow} 0$, on peut écrire :

$$F(a+h_n) = F(a) + h_n f(a) + \frac{1}{2}h_n^2 F''(a) + \underset{n\to+\infty}{o}(h_n^2)$$

$$\text{et} \qquad F(a-h_n) = F(a) - h_n f(a) + \frac{1}{2}h_n^2 F''(a) + \underset{n\to+\infty}{o}(h_n^2)$$

Par différence :

$$F(a+h_n) - F(a-h_n) = 2h_n f(a) + \underset{n\to+\infty}{o}(h_n^2)$$

D'après la question ***12.*** :

$$\begin{aligned} p_n &= F(a+h_n) - F(a-h_n) \\ &= 2h_n f(a) + \underset{n\to+\infty}{o}(h_n^2) \\ &= \theta_n^2 + \underset{n\to+\infty}{o}(h_n^2) && \textit{(par définition de } \theta_n\textit{)} \end{aligned}$$

$$\boxed{\text{Finalement, on a bien : } p_n \underset{n\to+\infty}{=} 2h_n f(a) + o(h_n^2) \underset{n\to+\infty}{=} \theta_n^2 + o(h_n^2)}$$

□

b) En déduire que : $p_n \underset{n\to+\infty}{\sim} 2h_n f(a)$, puis que $\lim\limits_{n\to+\infty} np_n = +\infty$.

Démonstration.
On a, d'après la question précédente :

$$p_n = 2h_n f(a) + \underset{n\to+\infty}{o}(h_n^2)$$

Or, $f(a) > 0$ et pour tout $n \in \mathbb{N}^*$, $h_n > 0$, donc :

$$\frac{p_n}{2h_n f(a)} = \frac{2h_n f(a) + \underset{n\to+\infty}{o}(h_n^2)}{2h_n f(a)} = 1 + \underset{n\to+\infty}{o}(h_n)$$

De plus, $h_n \underset{n\to+\infty}{\longrightarrow} 0$ donc $1 + \underset{n\to+\infty}{o}(h_n) \underset{n\to+\infty}{\longrightarrow} 1$ donc $\dfrac{p_n}{2h_n f(a)} \underset{n\to+\infty}{\longrightarrow} 1$.

$$\boxed{\text{D'où : } p_n \underset{n\to+\infty}{\sim} 2h_n f(a).}$$

On en déduit :

$$np_n \underset{n\to+\infty}{\sim} 2nh_n f(a)$$

Or, $2f(a) > 0$ et $nh_n \underset{n\to+\infty}{\longrightarrow} +\infty$ donc, par produit, $2nh_n f(a) \underset{n\to+\infty}{\longrightarrow} +\infty$.

$$\boxed{\text{Finalement : } np_n \underset{n\to+\infty}{\longrightarrow} +\infty.}$$

□

c) Montrer que $\hat{f}_n = \dfrac{\sigma_n}{\theta_n\sqrt{n}} D_n + \sqrt{n}\left(\dfrac{p_n}{\theta_n} - \theta_n\right)$ et que l'on a :

$$\lim_{n\to+\infty} \frac{\sigma_n}{\theta_n\sqrt{n}} = 1 \quad ; \quad \lim_{n\to+\infty} \sqrt{n}\left(\frac{p_n}{\theta_n} - \theta_n\right) = 0$$

Démonstration.

- D'une part :

$$\begin{aligned}
\hat{f}_n &= \frac{\theta_n\sqrt{n}}{f(a)}\left(f_n - f(a)\right) \\
&= \frac{\theta_n\sqrt{n}}{f(a)} f_n - \theta_n\sqrt{n} \\
&= \frac{\theta_n\sqrt{n}C_n}{2nh_nf(a)} - \theta_n\sqrt{n} && \textit{(par définition de } f_n\textit{)} \\
&= \frac{\theta_n\sqrt{n}C_n}{\theta_n^2 n} - \theta_n\sqrt{n} && \textit{(par définition de } \theta_n\textit{)} \\
&= \frac{1}{\theta_n\sqrt{n}}C_n - \theta_n\sqrt{n}
\end{aligned}$$

- D'autre part :

$$\begin{aligned}
\frac{\sigma_n}{\theta_n\sqrt{n}}D_n + \sqrt{n}\left(\frac{p_n}{\theta_n} - \theta_n\right) &= \frac{C_n - np_n}{\theta_n\sqrt{n}} + \sqrt{n}\frac{p_n - \theta_n^2}{\theta_n} && \textit{(par définition de } D_n\textit{)} \\
&= \frac{C_n - np_n}{\theta_n\sqrt{n}} + n\frac{p_n - \theta_n^2}{\theta_n\sqrt{n}} \\
&= \frac{C_n - \cancel{np_n} + \cancel{np_n} - n\theta_n^2}{\theta_n\sqrt{n}} \\
&= \frac{1}{\theta_n\sqrt{n}}C_n - \frac{n\theta_n^2}{\theta_n\sqrt{n}} \\
&= \frac{1}{\theta_n\sqrt{n}}C_n - \theta_n\sqrt{n}
\end{aligned}$$

$$\boxed{\text{D'où : } \hat{f}_n = \frac{\sigma_n}{\theta_n\sqrt{n}}D_n + \sqrt{n}\left(\frac{p_n}{\theta_n} - \theta_n\right).}$$

- De plus,

$$\begin{aligned}
\frac{\sigma_n}{\theta_n\sqrt{n}} &= \frac{\sqrt{np_n(1-p_n)}}{\theta_n\sqrt{n}} && \textit{(par définition de } \sigma_n\textit{)} \\
&= \frac{\sqrt{p_n(1-p_n)}}{\theta_n} \\
&= \frac{\sqrt{p_n(1-p_n)}}{\sqrt{2h_nf(a)}} && \textit{(par définition de } \theta_n\textit{)} \\
&\underset{n\to+\infty}{\sim} \frac{\sqrt{p_n(1-p_n)}}{\sqrt{p_n}} && \textit{(d'après la question } \textbf{\textit{14.b)}}\textit{)} \\
&= \sqrt{1-p_n}
\end{aligned}$$

$$\boxed{\text{Puisque } p_n \underset{n\to+\infty}{\longrightarrow} 0, \text{ il vient : } \frac{\sigma_n}{\theta_n\sqrt{n}} \underset{n\to+\infty}{\longrightarrow} 1.}$$

- Enfin,

$$
\begin{aligned}
\sqrt{n}\left(\frac{p_n}{\theta_n}-\theta_n\right) &= \sqrt{n}\frac{p_n-\theta_n^2}{\theta_n} \\
&= \sqrt{n}\frac{\underset{n\to+\infty}{o}(h_n^2)}{\sqrt{2h_nf(a)}} && \textit{(d'après la question } \textbf{\textit{14.a)}}\textit{)} \\
&= \underset{n\to+\infty}{o}\left(\sqrt{n}h_n^{\frac{3}{2}}\right) && \textit{(car } 2f(a) \textit{ est une constante)} \\
&= \underset{n\to+\infty}{o}\left(\sqrt{nh_n^3}\right)
\end{aligned}
$$

Or, par hypothèse, $nh_n^3 \underset{n\to+\infty}{\longrightarrow} 0$ donc $\sqrt{nh_n^3} \underset{n\to+\infty}{\longrightarrow} 0$.

$$\boxed{\text{Finalement : } \sqrt{n}\left(\frac{p_n}{\theta_n}-\theta_n\right) \underset{n\to+\infty}{\longrightarrow} 0.}$$

Commentaire

Le nombre de notations intervenant dans la formule $\hat{f}_n = \dfrac{\sigma_n}{\theta_n\sqrt{n}}D_n + \sqrt{n}\left(\dfrac{p_n}{\theta_n}-\theta_n\right)$ et sa preuve est assez impressionnant. Il est difficile dans ces conditions de trouver du premier coup comment faire un calcul direct en partant d'un terme et en arrivant à l'autre.
La meilleure méthode dans ces cas là est souvent de calculer et de simplifier les deux termes en parallèle au brouillon, en essayant de tout exprimer à l'aide d'un nombre minimal de notations (ici nous avons choisi de tout exprimer à l'aide de C_n et θ_n).
Il est normal pour un calcul difficile comme celui-ci de devoir tester plusieurs pistes. Le résultat étant donné dans l'énoncé, il ne faut pas hésiter à l'admettre si l'intuition ne vient pas après quelques échecs.

$\square$

On admet, dans la suite de cette partie, que $(\hat{f}_n)_{n\geqslant 1}$ converge en loi vers N ce qui implique que pour tout $(x,y)\in\mathbb{R}^2$, avec $x\leqslant y$, on a :

$$\lim_{n\to+\infty} \mathbb{P}(x\leqslant \hat{f}_n\leqslant y) = \Phi(y)-\Phi(x)$$

15. Soit $\alpha\in\,]0,1[$. On pose $\eta_\alpha = t_\alpha^2$ où t_α est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi normale $(0,1)$.

a) Montrer que $\displaystyle\lim_{n\to+\infty} \mathbb{P}\left((f(a))^2-\left(2f_n+\frac{\eta_\alpha}{2nh_n}\right)f(a)+f_n^2\leqslant 0\right) = 1-\alpha$.

Démonstration.
Par définition $\Phi(t_\alpha) = 1-\dfrac{\alpha}{2}$. De plus, $\alpha\in\,]0,1[$ donc $\Phi(t_\alpha) > \dfrac{1}{2}$. D'où $t_\alpha > 0$.

De plus,

$$\begin{aligned}
&\left[(f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right) f(a) + f_n^2 \leqslant 0\right] \\
&= \left[(f(a))^2 - 2f_n f(a) - \frac{\eta_\alpha}{2nh_n} f(a) + f_n^2 \leqslant 0\right] \\
&= \left[(f_n - f(a))^2 - \frac{\eta_\alpha}{2nh_n} f(a) \leqslant 0\right] \\
&= \left[(f_n - f(a))^2 \leqslant \frac{\eta_\alpha}{2nh_n} f(a)\right] \\
&= \left[(f_n - f(a))^2 \leqslant \frac{t_\alpha^2}{2nh_n} f(a)\right] && \textit{(par définition de } \eta_\alpha\textit{)} \\
&= \left[\left(\frac{\theta_n \sqrt{n}}{f(a)}(f_n - f(a))\right)^2 \leqslant \frac{n\theta_n^2}{f(a)^2} \frac{t_\alpha^2}{2nh_n} f(a)\right] && \textit{(en multipliant par } \frac{n\theta_n^2}{f(a)^2} \geqslant 0\textit{)} \\
&= \left[\hat{f}_n^2 \leqslant \frac{2nh_n f(a)}{f(a)^2} \frac{t_\alpha^2}{2nh_n} f(a)\right] && \textit{(par définition de } \hat{f}_n \textit{ et } \theta_n\textit{)} \\
&= \left[\hat{f}_n^2 \leqslant t_\alpha^2\right] \\
&= \left[-t_\alpha \leqslant \hat{f}_n \leqslant t_\alpha\right] \qquad \textit{(car } t_\alpha \geqslant 0\textit{)}
\end{aligned}$$

Par convergence en loi de $(\hat{f}_n)_{n \geqslant 1}$ vers N, on a :

$$\mathbb{P}\left(\left[-t_\alpha \leqslant \hat{f}_n \leqslant t_\alpha\right]\right) \underset{n \to +\infty}{\longrightarrow} \Phi(t_\alpha) - \Phi(-t_\alpha) = 2\Phi(t_\alpha) - 1 = 1 - \alpha$$

$$\boxed{\text{D'où : } \lim_{n \to +\infty} \mathbb{P}\left((f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right) f(a) + f_n^2 \leqslant 0\right) = 1 - \alpha.}$$

Commentaire

On commence à construire dans cette question un intervalle de confiance asymptotique pour $f(a)$ au niveau de confiance $1 - \alpha$.
Il faut donc nécessairement faire apparaître $\hat{f}_n$ et t_α pour utiliser la convergence en loi qui est admise.
On se souvient alors que

$$\hat{f}_n = \frac{\theta_n \sqrt{n}}{f(a)}(f_n - f(a))$$

donc il faut faire apparaître le terme $f_n - f(a)$.
On fait le lien avec les termes $(f(a))^2$ et f_n^2 qui font penser au développement de $(f_n - f(a))^2$.

□

b) On note, pour $n \geqslant 1$, $\Delta_n = \sqrt{\left(f_n + \dfrac{\eta_\alpha}{4nh_n}\right)^2 - f_n^2}$.

Montrer que :

$$\lim_{n \to +\infty} \mathbb{P}\left(f(a) \in \left[f_n + \frac{\eta_\alpha}{4nh_n} - \Delta_n, f_n + \frac{\eta_\alpha}{4nh_n} + \Delta_n\right]\right) = 1 - \alpha$$

Démonstration.

$$\begin{aligned}
&\left[(f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right) f(a) + f_n^2 \leqslant 0\right] \\
&= \left[(f(a))^2 - 2f(a)\left(f_n + \frac{\eta_\alpha}{4nh_n}\right) + f_n^2 \leqslant 0\right] \\
&= \left[\left(f(a) - \left(f_n + \frac{\eta_\alpha}{4nh_n}\right)\right)^2 - \left(f_n + \frac{\eta_\alpha}{4nh_n}\right)^2 + f_n^2 \leqslant 0\right] \\
&= \left[\left(f(a) - \left(f_n + \frac{\eta_\alpha}{4nh_n}\right)\right)^2 \leqslant \left(f_n + \frac{\eta_\alpha}{4nh_n}\right)^2 - f_n^2\right] \\
&= \left[\left(f(a) - \left(f_n + \frac{\eta_\alpha}{4nh_n}\right)\right)^2 \leqslant \Delta_n^2\right] \\
&= \left[-\Delta_n \leqslant f(a) - \left(f_n + \frac{\eta_\alpha}{4nh_n}\right) \leqslant \Delta_n\right] \\
&= \left[f_n + \frac{\eta_\alpha}{4nh_n} - \Delta_n \leqslant f(a) \leqslant f_n + \frac{\eta_\alpha}{4nh_n} + \Delta_n\right] \\
&= \left[f(a) \in \left[f_n + \frac{\eta_\alpha}{4nh_n} - \Delta_n, f_n + \frac{\eta_\alpha}{4nh_n} + \Delta_n\right]\right]
\end{aligned}$$

D'après la question ***15.a)***, on a bien

$$\boxed{\lim_{n\to+\infty} \mathbb{P}\left(\left[f(a) \in \left[f_n + \frac{\eta_\alpha}{4nh_n} - \Delta_n, f_n + \frac{\eta_\alpha}{4nh_n} + \Delta_n\right]\right]\right) = 1 - \alpha.}$$

Commentaire

Puisque l'on cherche à construire un intervalle de confiance asymptotique pour $f(a)$ au niveau de confiance $1-\alpha$, il faut réussir à isoler $f(a)$ dans l'événement dont on contrôle la probabilité (celui de la question précédente).
On se rend alors compte que l'expression

$$(f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right) f(a) + f_n^2$$

n'est rien d'autre qu'un trinôme du second degré en $f(a)$. On met donc en place dans cette question le calcul classique de mise sous forme canonique d'un trinôme du second degré.
D'autre part, soulignons que Δ_n est bien définie. En effet,

$$\begin{aligned}
\left(f_n + \frac{\eta_\alpha}{4nh_n}\right)^2 - f_n^2 &= f_n^2 + 2f_n\frac{\eta_\alpha}{4nh_n} + \left(\frac{\eta_\alpha}{4nh_n}\right)^2 - f_n^2 \\
&= 2f_n\frac{\eta_\alpha}{4nh_n} + \left(\frac{\eta_\alpha}{4nh_n}\right)^2 \geqslant 0
\end{aligned}$$

$\square$

Partie 4 - Convergence « uniforme » en loi vers la loi normale

Soit $n \in \mathbb{N}^*$. On considère n variables aléatoires $X_1, \dots, X_n$ indépendantes centrées qui possèdent un moment d'ordre 3. On admet alors que ces variables aléatoires possèdent une variance.

On pose, pour tout $k \in [\![1, n]\!]$, $\mathbb{E}(X_k^2) = v_k$, $S_n = \sum\limits_{k=1}^{n} X_k$, $Y_k = S_n - X_k$ et on suppose que $\sum\limits_{k=1}^{n} v_k = 1$.

Soit f une fonction de classe $\mathcal{C}^2$ sur $\mathbb{R}$ telle que pour tout x réel, $|f(x)| \leqslant 1$ et $|f''(x)| \leqslant 2$.

On admet que si X et Y sont des variables aléatoires possédant une espérance alors $\mathbb{E}(Y f(X))$ et $\mathbb{E}(f'(X))$ existent.

Commentaire

Insistons à nouveau sur le fait que l'existence d'une espérance n'est jamais immédiate a priori sauf pour les variables aléatoires finies. Il faut donc avoir le réflexe de démontrer l'existence avant de se lancer dans le calcul. Cette remarque est valable pour les trois questions qui suivent.

16. a) Montrer que $\sum\limits_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum\limits_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k)) = \mathbb{E}(f'(S_n))$.

Démonstration.

- Les variables aléatoires X_k sont toutes centrées donc possèdent chacune une espérance.
- On en déduit que S_n et Y_k admettent également une espérance comme sommes de variables aléatoires admettant chacune une espérance.
- D'après le résultat admis, $f'(S_n)$ et $f'(Y_k)$ admettent donc une espérance.
  Par suite, $f'(S_n) - f'(Y_k)$ admet une espérance.
- Pour finir, $Y_k = S_n - X_k = X_1 + \cdots + X_{k-1} + X_{k+1} + \cdots + X_n$ donc, par lemme des coalitions, Y_k et X_k sont indépendantes. Toujours par lemme des coalitions, X_k^2 et $f'(Y_k)$ sont aussi indépendantes.
- Comme les variables aléatoires X_k^2 et $f'(Y_k)$ sont indépendantes et admettent chacune une espérance, on peut conclure que $X_k^2 f'(Y_k)$ admet une espérance.
- On peut maintenant faire le calcul :

$$\sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k))$$

$$= \sum_{k=1}^{n} v_k (\mathbb{E}(f'(S_n)) - \mathbb{E}(f'(Y_k))) + \sum_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k)) \qquad \textit{(par linéarité)}$$

$$= \sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n)) - \sum_{k=1}^{n} v_k \mathbb{E}(f'(Y_k)) + \sum_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k))$$

$$= \mathbb{E}(f'(S_n)) \sum_{k=1}^{n} v_k - \cancel{\sum_{k=1}^{n} \mathbb{E}(X_k^2)\mathbb{E}(f'(Y_k))} + \cancel{\sum_{k=1}^{n} \mathbb{E}(X_k^2)\mathbb{E}(f'(Y_k))}$$

(car $\mathbb{E}(X_k^2 f'(Y_k)) = \mathbb{E}(X_k^2)\mathbb{E}(f'(Y_k))$ par indépendance)

$$= \mathbb{E}(f'(S_n)) \qquad \textit{(car } \sum_{k=1}^{n} v_k = 1\textit{)}$$

□

b) Montrer que $\sum\limits_{k=1}^{n} \mathbb{E}(X_k\,[f(S_n) - f(Y_k)]) = \sum\limits_{k=1}^{n} \mathbb{E}(X_k f(S_n)) = \mathbb{E}(S_n f(S_n))$.

Démonstration.

- Les variables aléatoires X_k et S_n admettent une espérance donc $X_k f(S_n)$ admet une espérance (cf résultat admis). Par somme $S_n f(S_n) = \sum\limits_{k=1}^{n} (X_k f(S_n))$ admet aussi une espérance.
- Les variables aléatoires X_k et Y_k admettent une espérance donc $X_k f(Y_k)$ admet une espérance (cf résultat admis).
- Par somme, $X_k f(S_n) - X_k f(Y_k) = X_k(f(S_n) - f(Y_k))$ admet une espérance.
- Les variables Y_k et X_k sont indépendantes donc par lemme des coalitions, $f(Y_k)$ et X_k sont aussi indépendantes.
- On peut maintenant faire le calcul :

$$\begin{aligned}
&\sum_{k=1}^{n} \mathbb{E}\Big(X_k\big(f(S_n) - f(Y_k)\big)\Big) \\
&= \sum_{k=1}^{n} \mathbb{E}(X_k f(S_n) - X_k f(Y_k)) \\
&= \sum_{k=1}^{n} (\mathbb{E}(X_k f(S_n)) - \mathbb{E}(X_k f(Y_k))) && \textit{(par linéarité)} \\
&= \sum_{k=1}^{n} \mathbb{E}(X_k f(S_n)) - \sum_{k=1}^{n} \mathbb{E}(X_k f(Y_k)) \\
&= \sum_{k=1}^{n} \mathbb{E}(X_k f(S_n)) - \sum_{k=1}^{n} \mathbb{E}(X_k)\mathbb{E}(f(Y_k)) && \textit{(par indépendance)} \\
&= \sum_{k=1}^{n} \mathbb{E}(X_k f(S_n)) && \textit{(car } \forall k \in [\![1,n]\!],\ \mathbb{E}(X_k) = 0\textit{)} \\
&= \mathbb{E}\left(\sum_{k=1}^{n} X_k f(S_n)\right) && \textit{(par linéarité)} \\
&= \mathbb{E}\left(f(S_n)\sum_{k=1}^{n} X_k\right) \\
&= \mathbb{E}\left(f(S_n)S_n\right)
\end{aligned}$$

□

c) En déduire que :

$$\mathbb{E}(f'(S_n) - S_n f(S_n)) = \sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^{n} \mathbb{E}\left(X_k\left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\right)$$

Démonstration.

- Les variables aléatoires $f'(S_n)$ et $S_n f(S_n)$ admettent une espérance donc $f'(S_n) - S_n f(S_n)$ admet une espérance par somme.
- Les variables aléatoires $X_k^2 f'(Y_k)$, $X_k f(S_n)$ et $X_k f(Y_k)$ admettent une espérance donc la variable aléatoire

$$X_k\Big(X_k f'(Y_k) - (f(S_n) - f(Y_k))\Big) = X_k^2 f'(Y_k) - X_k f(S_n) + X_k f(Y_k)$$

admet une espérance par somme.

- On peut maintenant faire le calcul :

$$\begin{aligned}
& \mathbb{E}(f'(S_n) - S_n f(S_n)) \\
= \ & \mathbb{E}(f'(S_n)) - \mathbb{E}(S_n f(S_n)) && \textit{(par linéarité)} \\
= \ & \sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k)) && \textit{(d'après la question } \textbf{16.a)}\textit{)} \\
& \quad - \sum_{k=1}^{n} \mathbb{E}(X_k \left[f(S_n) - f(Y_k)\right]) && \textit{(d'après la question } \textbf{16.b)}\textit{)} \\
= \ & \sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) \\
& \quad + \sum_{k=1}^{n} \mathbb{E}(X_k^2 f'(Y_k) - X_k \left[f(S_n) - f(Y_k)\right]) && \textit{(par linéarité)} \\
= \ & \sum_{k=1}^{n} v_k \mathbb{E}(f'(S_n) - f'(Y_k)) \\
& \quad + \sum_{k=1}^{n} \mathbb{E}\big(X_k \left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\big)
\end{aligned}$$

□

17. Soit a et b deux réels.

a) Montrer que :

$$b f'(a) - (f(a+b) - f(a)) = \int_0^1 b(f'(a) - f'(a+tb)) \, dt$$

Démonstration.
La fonction f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ donc la fonction $t \mapsto f'(a+tb)$ est continue sur $\mathbb{R}$ comme composée de deux fonctions continues sur $\mathbb{R}$.
Ainsi, l'intégrale $\displaystyle\int_0^1 b(f'(a) - f'(a+tb)) \, dt$ est bien définie.
Tout d'abord, par linéarité de l'intégrale :

$$\int_0^1 b(f'(a) - f'(a+tb)) \, dt = \int_0^1 b f'(a) \, dt - \int_0^1 b f'(a+tb) \, dt$$

Ensuite, on remarque que $t \mapsto b f'(a+tb)$ est la dérivée de $t \mapsto f(a+tb)$, d'où :

$$\int_0^1 b f'(a+tb) \, dt = \left[\, f(a+tb) \, \right]_0^1 = f(a+b) - f(a)$$

Pour finir :

$$\int_0^1 b f'(a) \, dt = b f'(a)$$

$$\boxed{\text{Finalement : } \int_0^1 b(f'(a) - f'(a+tb)) \, dt = b f'(a) - (f(a+b) - f(a)).}$$

Commentaire

Il ne faut jamais se décourager au cours de l'épreuve car une question facile et indépendante du reste de l'énoncé peut toujours arriver après plusieurs questions (trop) difficiles. C'est le cas de cette question ***17.a)*** qui est un résultat d'analyse. Il faut donc prendre le temps de lire le sujet en début d'épreuve pour repérer les questions qui semblent abordables et leur réserver du temps en fin d'épreuve même si l'on est resté bloqué sur beaucoup de questions précédentes.

$\square$

b) En déduire que :

$$\left|bf'(a) - (f(a+b) - f(a))\right| \leqslant b^2$$

Démonstration.
D'après la question précédente :

$$\begin{aligned}\left|bf'(a) - (f(a+b) - f(a))\right| &= \left|\int_0^1 b(f'(a) - f'(a+tb))\, dt\right| \\ &\leqslant \int_0^1 \left|b(f'(a) - f'(a+tb))\right|\, dt \quad \textit{(par inégalité triangulaire, car } 0 \leqslant 1\textit{)}\end{aligned}$$

Soit $t \in [0,1]$.

$$\left|b(f'(a) - f'(a+tb))\right| = |b|\left|f'(a) - f'(a+tb)\right|$$

Or, pour tout $x \in \mathbb{R}$, $|f''(x)| \leqslant 2$, donc par inégalité des accroissements finis :

$$\left|f'(a) - f'(a+tb)\right| \leqslant 2\,|a - (a+tb)| = 2\,|tb| = 2\,|b|\,t \qquad \textit{(car } t \geqslant 0\textit{)}$$

Par croissance de l'intégrale, les bornes étant rangées dans l'ordre croissant ($0 \leqslant 1$), on a

$$\begin{aligned}\left|bf'(a) - (f(a+b) - f(a))\right| &\leqslant \int_0^1 \left|b(f'(a) - f'(a+tb))\right|\, dt \\ &\leqslant \int_0^1 2\,|b|^2\, t\, dt \\ &= 2b^2 \int_0^1 t\, dt \\ &= 2b^2 \times \frac{1}{2} \\ &= b^2\end{aligned}$$

$$\boxed{\text{On a bien : } |bf'(a) - (f(a+b) - f(a))| \leqslant b^2.}$$

$\square$

c) En conclure que :

$$\left|\mathbb{E}(f'(S_n) - S_n f(S_n))\right| \leqslant 2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3)$$

puis, grâce à l'inégalité (R_2), que, pour tout x réel :

$$d_{S_n}(x) \leqslant \sqrt{3\left(2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3)\right)} \tag{R_3}$$

Démonstration.
D'après la question ***16.c)*** et par inégalité triangulaire :

$$\begin{aligned}
&\left|\mathbb{E}(f'(S_n) - S_n f(S_n))\right| \\
&\leqslant \sum_{k=1}^{n} |v_k| \left|\mathbb{E}(f'(S_n) - f'(Y_k))\right| + \sum_{k=1}^{n} \left|\mathbb{E}\left(X_k \left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\right)\right| \\
&\leqslant \sum_{k=1}^{n} v_k \left|\mathbb{E}(f'(S_n) - f'(Y_k))\right| + \sum_{k=1}^{n} \left|\mathbb{E}\left(X_k \left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\right)\right|
\end{aligned}$$

En effet, $X_k^2 \geqslant 0$ donc $v_k \geqslant 0$ par positivité de l'espérance.
Soit $k \in [\![1, n]\!]$. Majorons maintenant chaque terme séparément.

- D'une part :

$$\begin{aligned}
\left|\mathbb{E}(f'(S_n) - f'(Y_k))\right| &\leqslant \mathbb{E}(\left|f'(S_n) - f'(Y_k)\right|) && \textit{(par inégalité triangulaire)} \\
&\leqslant \mathbb{E}(2\,|S_n - Y_k|) && \textit{(par inégalité des accroissements finis et par croissance de l'espérance)} \\
&= 2\mathbb{E}(|X_k|) && \textit{(par linéarité)}
\end{aligned}$$

- D'autre part :

$$\begin{aligned}
&\left|\mathbb{E}\left(X_k \ \left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\right)\right| \\
&\leqslant \mathbb{E}\Big(\left|X_k \ \left[X_k f'(Y_k) - (f(S_n) - f(Y_k))\right]\right| \Big) && \textit{(par inégalité triangulaire)} \\
&= \mathbb{E}\Big(\ |X_k| \ \left|X_k f'(Y_k) - (f(S_n) - f(Y_k))\right| \Big) \\
&= \mathbb{E}\Big(\ |X_k| \ \left|X_k f'(Y_k) - (f(Y_k + X_k) - f(Y_k))\right| \Big) \\
&\leqslant \mathbb{E}\Big(|X_k|\, X_k^2 \Big) && \textit{(d'après la question \textbf{17.b)} et par croissance de l'espérance)} \\
&= \mathbb{E}\left(|X_k|\, |X_k|^2 \right) \\
&= \mathbb{E}\left(|X_k|^3 \right)
\end{aligned}$$

On en déduit, en sommant toutes ces inégalités, que :

$$\boxed{|\mathbb{E}(f'(S_n) - S_n f(S_n))| \leqslant 2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3).}$$

Notons maintenant

$$M_X = 2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3)$$

Soit $h \in W$. D'après la partie 1, la fonction f_h vérifie :

× f_h est de classe $\mathcal{C}^2$ sur $\mathbb{R}$,

× pour tout $x \in \mathbb{R}$, $|f_h(x)| \leqslant 1$,

× pour tout $x \in \mathbb{R}$, $|f_h''(x)| \leqslant 2$.

On peut donc appliquer la majoration précédente avec $f = f_h$:

$$\left|\mathbb{E}(f_h'(S_n) - S_n f_h(S_n))\right| \leqslant M_X$$

Or, d'après la question ***3.c)*** (la variable S_n admet bien une espérance),

$$|\mathbb{E}(h(S_n)) - \mathbb{E}(h(N))| = \left|\mathbb{E}(f_h'(S_n) - S_n f_h(S_n))\right|$$

On a donc démontré que, pour tout $h \in W$,

$$|\mathbb{E}(h(S_n)) - \mathbb{E}(h(N))| \leqslant M_X$$

On peut donc appliquer la question ***10.*** afin obtenir, pour tout x réel :

$$d_{S_n}(x) \leqslant \sqrt{3M_X}$$

$$\boxed{\text{Autrement dit : pour tout } x \text{ réel, } d_{S_n}(x) \leqslant \sqrt{3\left(2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3)\right)}.}$$

Commentaire

- Cette question est à considérer comme très difficile parce qu'elle fait le bilan des deux premières parties ainsi que du début de la quatrième partie. Ainsi, il faut être capable en un temps court de prendre du recul sur les résultats précédemment démontrés au cours de l'épreuve pour les utiliser conjointement.
- La preuve donnée nous semble être dans l'esprit du sujet, mais nous n'avons pas démontré que toutes les espérances écrites existent bien. En particulier, pourquoi $\mathbb{E}(|X_k|^3)$ existe ?
  A nouveau, cela repose sur le résultat suivant :
  « Soit Y une variable aléatoire. Alors Y admet une espérance si et seulement si $|Y|$ admet une espérance. »
  (seul le sens direct est utile ici)
  Nous avons les outils pour faire la preuve dans le cas discret ou le cas à densité, mais pas dans le cas général.
  Ce commentaire est analogue à celui fait au moment de la question ***9***.

$\square$

Une définition - *Dans la suite du sujet, si $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles et $(\delta_n)_{n\geqslant 1}$ une suite réelle de limite nulle qui vérifient,*

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, d_{X_n}(x) \leqslant \delta_n$$

on dira alors que $(X_n)_{n\geqslant 1}$ converge uniformément en loi vers N.
On remarque, et on l'admet pour la suite, que si $(X_n)_{n\geqslant 1}$ converge uniformément en loi vers N alors $(X_n)_{n\geqslant 1}$ converge en loi vers N.

Commentaire

La définition de la convergence (simple) en loi vers N, qui est au programme, s'écrit dans ce contexte (puisque Φ est continue sur $\mathbb{R}$) :

$$\forall x \in \mathbb{R}, F_{X_n}(x) \underset{n \to +\infty}{\longrightarrow} \Phi(x)$$

c'est-à-dire

$$\forall x \in \mathbb{R}, d_{X_n}(x) \underset{n \to +\infty}{\longrightarrow} 0$$

Pour chaque x fixé, la suite $(d_{X_n}(x))_{n \geqslant 1}$ des distances au point x converge vers 0. La *vitesse de convergence* peut alors dépendre de x. Cela peut poser des problèmes dans certaines preuves, où l'on a besoin que cette vitesse ne dépende pas du point x considéré. C'est à cet effet que l'on introduit la notion de convergence « uniforme ».
Écrire

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, d_{X_n}(x) \leqslant \delta_n$$

c'est affirmer que **toutes** les suites $(d_{X_n}(x))_{n \geqslant 1}$ convergent vers 0 au moins aussi vite que la suite $(\delta_n)_{n \geqslant 1}$ (quelque soit le point x considéré). On a alors une vitesse de convergence uniforme.

18. *Une première application.* On suppose dans cette question que $(Z_k)_{k \geqslant 1}$ est une suite de variables aléatoires indépendantes, suivant la même loi et admettant des moments d'ordre 1 à 3.

On note pour $i \in [\![1, 3]\!]$, $s_i = \mathbb{E}(|Z_k - \mathbb{E}(Z_k)|^i)$, $\sigma = \sqrt{s_2}$ et , $X_k = \dfrac{Z_k - \mathbb{E}(Z_k)}{\sigma\sqrt{n}}$ pour tout $k \in \mathbb{N}^*$.

On suppose que $\sigma \neq 0$.

On utilise les notations de la question précédente.

a) Montrer que l'on peut appliquer l'inégalité (R_3) qui donne ici :

$$d_{S_n}(x) \leqslant \sqrt{3\frac{2\sigma^2 s_1 + s_3}{\sigma^3\sqrt{n}}}$$

Démonstration.

- Les variables aléatoires Z_k sont indépendantes donc les variables aléatoires $X_k = \dfrac{Z_k - \mathbb{E}(Z_k)}{\sigma\sqrt{n}}$ sont indépendantes par lemme des coalitions.
- Les variables aléatoires Z_k admettent chacune des moments d'ordre 1 à 3 donc les variables aléatoires X_k admettent également chacune des moments d'ordre 1 à 3 (par transformation affine).
- Soit $k \geqslant 1$. Par linéarité :

$$\mathbb{E}(X_k) = \mathbb{E}\left(\frac{Z_k - \mathbb{E}(Z_k)}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma\sqrt{n}}\mathbb{E}\left(Z_k - \mathbb{E}(Z_k)\right) = \frac{1}{\sigma\sqrt{n}}(\mathbb{E}(Z_k) - \mathbb{E}(Z_k)) = 0$$

donc les variables aléatoires X_k sont centrées.

- Enfin :

$$\begin{aligned}
\sum_{k=1}^{n} v_k &= \sum_{k=1}^{n} \mathbb{E}(X_k^2)\\
&= \sum_{k=1}^{n} \mathbb{E}\left(\left(\frac{Z_k-\mathbb{E}(Z_k)}{\sigma\sqrt{n}}\right)^2\right)\\
&= \sum_{k=1}^{n} \frac{1}{n\sigma^2}\mathbb{E}\left((Z_k-\mathbb{E}(Z_k))^2\right)\\
&= \sum_{k=1}^{n} \frac{1}{ns_2}s_2\\
&= 1
\end{aligned}$$

Nous sommes donc dans les conditions d'application de la majoration (R_3) qui s'écrit, pour tout x réel :

$$d_{S_n}(x) \leqslant \sqrt{3M_X}$$

où

$$M_X = 2\sum_{k=1}^{n} v_k\mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3)$$

Or,

$$\begin{aligned}
\mathbb{E}(|X_k|) &= \mathbb{E}\left(\left|\frac{Z_k-\mathbb{E}(Z_k)}{\sigma\sqrt{n}}\right|\right)\\
&= \frac{1}{\sigma\sqrt{n}}\,\mathbb{E}(|Z_k-\mathbb{E}(Z_k)|) \qquad \textit{(par linéarité de l'espérance et car } \frac{1}{\sigma\sqrt{n}} \geqslant 0\textit{)}\\
&= \frac{s_1}{\sigma\sqrt{n}}
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{E}(|X_k|^3) &= \mathbb{E}\left(\left|\frac{Z_k-\mathbb{E}(Z_k)}{\sigma\sqrt{n}}\right|^3\right)\\
&= \frac{1}{(\sigma\sqrt{n})^3}\,\mathbb{E}(|Z_k-\mathbb{E}(Z_k)|^3) \qquad \textit{(par linéarité de l'espérance et car } \frac{1}{(\sigma\sqrt{n})^3} \geqslant 0\textit{)}\\
&= \frac{s_3}{(\sigma\sqrt{n})^3}
\end{aligned}$$

donc

$$\begin{aligned}
M_X &= 2\sum_{k=1}^{n} v_k\frac{s_1}{\sigma\sqrt{n}} + \sum_{k=1}^{n} \frac{s_3}{(\sigma\sqrt{n})^3}\\
&= \frac{2s_1}{\sigma\sqrt{n}}\sum_{k=1}^{n} v_k + n\frac{s_3}{(\sigma\sqrt{n})^3}\\
&= \frac{2s_1}{\sigma\sqrt{n}} + n\frac{s_3}{(\sigma\sqrt{n})^3} \qquad \textit{(car } \sum_{k=1}^{n} v_k = 1\textit{)}\\
&= \frac{2s_1}{\sigma\sqrt{n}} + \frac{s_3}{\sigma^3\sqrt{n}}\\
&= \frac{2\sigma^2 s_1 + s_3}{\sigma^3\sqrt{n}}
\end{aligned}$$

$$\boxed{\text{D'où : } d_{S_n}(x) \leqslant \sqrt{3\dfrac{2\sigma^2 s_1 + s_3}{\sigma^3\sqrt{n}}}.}$$

Commentaire

L'énoncé nous demande explicitement de vérifier que l'on a le droit d'appliquer l'inégalité (R_3). C'est en fait un très bon réflexe qu'il faut s'efforcer d'avoir même lorsque l'énoncé ne nous dit pas de le faire.

□

b) En déduire que $(S_n)_{n\geqslant 1}$ converge uniformément en loi vers N, donc converge en loi vers N. Quel résultat du cours nous aurait permis d'obtenir cette dernière convergence directement ?

Démonstration.
Les constantes σ, s_1 et s_3 ne dépendent pas de n donc :

$$\delta_n = \sqrt{3\frac{2\sigma^2 s_1 + s_3}{\sigma^3\sqrt{n}}} \underset{n\to+\infty}{\longrightarrow} 0$$

D'après la question précédente :

$$\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ d_{S_n}(x) \leqslant \delta_n$$

donc

$$\boxed{(S_n)_{n\geqslant 1} \text{ converge uniformément en loi vers } N}$$

D'après la remarque, on a donc également :

$$\boxed{(S_n)_{n\geqslant 1} \text{ converge en loi vers } N}$$

On aurait pu obtenir cette dernière convergence par une application directe du théorème central limite.
En effet, notons $W_n = \sum\limits_{k=1}^{n} Z_k$ et $m = \mathbb{E}(Z_k)$. Alors :

× les variables aléatoires Z_k sont indépendantes,

× les variables aléatoires Z_k suivent la même loi,

× les variables aléatoires Z_k admettent une espérance et une variance non nulle.

Donc $W_n^* \xrightarrow[n\to+\infty]{\mathscr{L}} N$.
Il reste à montrer que $S_n = W_n^*$. Or :

$$\begin{aligned} S_n &= \sum_{k=1}^{n} X_k \\ &= \sum_{k=1}^{n} \frac{Z_k - m}{\sigma\sqrt{n}} \\ &= \frac{W_n - nm}{\sigma\sqrt{n}} \\ &= W_n^* \qquad \textit{(calcul classique)} \end{aligned}$$

□

19. *Une deuxième application.* On suppose dans cette question que $Z_1, \dots, Z_n$ sont des variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre $p_n \in \,]0,1[$.

On pose $\sigma_n = \sqrt{np_n(1-p_n)}$ et $X_k = \dfrac{Z_k - p_n}{\sigma_n}$ pour tout $k \in [\![1,n]\!]$.

a) Montrer que $\mathbb{E}(|X_k|) = \dfrac{2\sigma_n}{n}$ et $\mathbb{E}(|X_k|^3) \leqslant \dfrac{2}{n\sigma_n}$.

Démonstration.
Soit $k \in [\![1,n]\!]$. Donnons les lois des différentes variables aléatoires en jeu.

- Loi de Z_k :
  - $\times$ $Z_k(\Omega) = \{0,1\}$
  - $\times$ $\mathbb{P}([Z_k = 0]) = 1 - p_n$ et $\mathbb{P}([Z_k = 1]) = p_n$
- Loi de X_k :
  - $\times$ $X_k(\Omega) = \left\{ -\dfrac{p_n}{\sigma_n}, \dfrac{1-p_n}{\sigma_n} \right\}$
  - $\times$ $\mathbb{P}\left(\left[X_k = -\dfrac{p_n}{\sigma_n}\right]\right) = 1 - p_n$ et $\mathbb{P}\left(\left[X_k = \dfrac{1-p_n}{\sigma_n}\right]\right) = p_n$
- Loi de $|X_k|$:
  - $\times$ $|X_k|(\Omega) = \left\{ \dfrac{p_n}{\sigma_n}, \dfrac{1-p_n}{\sigma_n} \right\}$
  - $\times$ $\mathbb{P}\left(\left[\ |X_k| = \dfrac{p_n}{\sigma_n}\ \right]\right) = 1 - p_n$ et $\mathbb{P}\left(\left[\ |X_k| = \dfrac{1-p_n}{\sigma_n}\ \right]\right) = p_n$
- Loi de $|X_k|^3$:
  - $\times$ $|X_k|^3(\Omega) = \left\{ \left(\dfrac{p_n}{\sigma_n}\right)^3, \left(\dfrac{1-p_n}{\sigma_n}\right)^3 \right\}$
  - $\times$ $\mathbb{P}\left(\left[\ |X_k|^3 = \left(\dfrac{p_n}{\sigma_n}\right)^3\ \right]\right) = 1 - p_n$ et $\mathbb{P}\left(\left[\ |X_k|^3 = \left(\dfrac{1-p_n}{\sigma_n}\right)^3\ \right]\right) = p_n$

  Remarquons que $\dfrac{p_n}{\sigma_n} = \dfrac{1-p_n}{\sigma_n}$ si et seulement si $p_n = \dfrac{1}{2}$. De plus, la formule utilisée pour calculer les espérances ci-dessous reste vraie même dans ce cas.

Les variables aléatoires $|X_k|$ et $|X_k|^3$ sont finies donc admettent chacune une espérance.

$$\begin{aligned}
\mathbb{E}(|X_k|) &= \frac{p_n}{\sigma_n}(1-p_n) + \frac{1-p_n}{\sigma_n}p_n \\
&= \frac{2p_n(1-p_n)}{\sigma_n} \\
&= \frac{2np_n(1-p_n)}{n\sigma_n} \\
&= \frac{2\sigma_n^2}{n\sigma_n} && \textit{(par définition de } \sigma_n\textit{)} \\
&= \frac{2\sigma_n}{n}
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{E}(|X_k|^3) &= \left(\frac{p_n}{\sigma_n}\right)^3 (1-p_n) + \left(\frac{1-p_n}{\sigma_n}\right)^3 p_n \\
&= \frac{p_n(1-p_n)}{\sigma_n^3} \left(p_n^2 + (1-p_n)^2\right) \\
&= \frac{2}{n\sigma_n} \frac{np_n(1-p_n)}{2\sigma_n^2} \left(p_n^2 + (1-p_n)^2\right) \\
&= \frac{\cancel{2}}{n\sigma_n} \frac{\cancel{\sigma_n^2}}{\cancel{2\sigma_n^2}} \left(p_n^2 + (1-p_n)^2\right) && \textit{(par définition de } \sigma_n\textit{)} \\
&= \frac{1}{n\sigma_n} \left(p_n^2 + (1-p_n)^2\right) \\
&\leqslant \frac{1}{n\sigma_n} (1+1) && \textit{(car } p_n^2 \leqslant 1 \textit{ et } (1-p_n)^2 \leqslant 1\textit{)} \\
&= \frac{2}{n\sigma_n}
\end{aligned}$$

$$\boxed{\text{D'où : } \mathbb{E}(|X_k|) = \frac{2\sigma_n}{n} \text{ et } \mathbb{E}(|X_k|^3) \leqslant \frac{2}{n\sigma_n}.}$$

□

b) En déduire que, pour tout x réel :

$$d_{S_n}(x) \leqslant 2\sqrt{3\left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n}\right)}$$

Démonstration.
On démontre comme à la question ***18.a)*** que l'on peut appliquer l'inégalité (R_3).
Pour tout $x \in \mathbb{R}$:

$$d_{S_n}(x) \leqslant \sqrt{3M_X}$$

où

$$\begin{aligned}
M_X &= 2\sum_{k=1}^{n} v_k \mathbb{E}(|X_k|) + \sum_{k=1}^{n} \mathbb{E}(|X_k|^3) \\
&\leqslant 2\sum_{k=1}^{n} v_k \frac{2\sigma_n}{n} + \sum_{k=1}^{n} \frac{2}{n\sigma_n} && \textit{(d'après la question } \textbf{\textit{19.a)}}\textit{)} \\
&= \frac{4\sigma_n}{n} \sum_{k=1}^{n} v_k + \frac{2}{\sigma_n} \\
&= 4\left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n}\right) && \textit{(car } \sum_{k=1}^{n} v_k = 1\textit{)}
\end{aligned}$$

$$\boxed{\text{Ainsi : pour tout } x \text{ réel, } d_{S_n}(x) \leqslant \sqrt{3 \times 4\left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n}\right)} = 2\sqrt{3\left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n}\right)}.}$$

□

c) Justifier le résultat suivant :
si pour tout $n \in \mathbb{N}^*$, T_n est une variable aléatoire qui suit la loi binomiale (n, p_n) avec $\lim\limits_{n \to +\infty} p_n = 0$ et $\lim\limits_{n \to +\infty} np_n = +\infty$ alors, $\left(\dfrac{T_n - np_n}{\sqrt{np_n(1-p_n)}} \right)_{n \geqslant 1}$ converge uniformément en loi vers N.

Démonstration.
On note $T_n = \sum\limits_{k=1}^{n} Z_k$ où les Z_k sont indépendantes et suivent toutes la même loi de Bernoulli de paramètre p_n.
On note comme précédemment $\sigma_n = \sqrt{np_n(1-p_n)}$ et $X_k = \dfrac{Z_k - p_n}{\sigma_n}$.
Alors

$$\begin{aligned} S_n &= \sum_{k=1}^{n} X_k \\ &= \sum_{k=1}^{n} \frac{Z_k - p_n}{\sigma_n} \\ &= \frac{T_n - np_n}{\sigma_n} \\ &= \frac{T_n - np_n}{\sqrt{np_n(1-p_n)}} \end{aligned}$$

et donc il faut montrer que $(S_n)_{n \geqslant 1}$ converge uniformément en loi vers N.
Notons $\delta_n = 2\sqrt{3\left(\dfrac{\sigma_n}{n} + \dfrac{1}{2\sigma_n}\right)}$. D'après la question précédente, pour tout $x \in \mathbb{R}$,

$$d_{S_n}(x) \leqslant \delta_n$$

Il reste donc à montrer que $\delta_n \underset{n \to +\infty}{\longrightarrow} 0$.

- Par hypothèse : $p_n \underset{n \to +\infty}{\longrightarrow} 0$. Ainsi : $1 - p_n \underset{n \to +\infty}{\longrightarrow} 1$ et donc $1 - p_n \underset{n \to +\infty}{\sim} 1$. On en déduit que :

  $$\sigma_n \underset{n \to +\infty}{\sim} \sqrt{np_n}$$

  et donc $\sigma_n \underset{n \to +\infty}{\longrightarrow} +\infty$ car $np_n \underset{n \to +\infty}{\longrightarrow} +\infty$ par hypothèse.
- Par suite :

  $$\frac{\sigma_n}{n} \underset{n \to +\infty}{\sim} \sqrt{\frac{p_n}{n}}$$

  et donc $\dfrac{\sigma_n}{n} \underset{n \to +\infty}{\longrightarrow} 0$.

On en déduit que $\delta_n \underset{n \to +\infty}{\longrightarrow} 0$.

$$\boxed{\text{Donc } \left(\frac{T_n - np_n}{\sqrt{np_n(1-p_n)}} \right)_{n \geqslant 1} \text{ converge uniformément en loi vers } N.}$$

□

20. *Un petit lemme.* Soit $(V_n)_{n \geqslant 1}$ une suite de variables aléatoires réelles qui converge uniformément en loi vers N. Soit $(a_n)_{n \geqslant 1}$ une suite de réels strictement positifs qui converge vers a, tel que $a > 0$, et $(b_n)_{n \geqslant 1}$ une suite de réels qui converge vers b.

a) Soit X une variable aléatoire et (α, β) un couple de réels avec $\alpha > 0$. On note $F_{\alpha X+\beta}$ et $F_{\alpha N+\beta}$ les fonctions de répartition respectives de $\alpha X + \beta$ et $\alpha N + \beta$.
Montrer que, pour tout x réel,

$$|F_{\alpha X+\beta}(x) - F_{\alpha N+\beta}(x)| = d_X\left(\frac{x-\beta}{\alpha}\right)$$

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

$$\begin{aligned}
|F_{\alpha X+\beta}(x) - F_{\alpha N+\beta}(x)| &= |\mathbb{P}([\alpha X + \beta \leqslant x]) - \mathbb{P}([\alpha N + \beta \leqslant x])| \\
&= |\mathbb{P}([\alpha X \leqslant x - \beta]) - \mathbb{P}([\alpha N \leqslant x - \beta])| \\
&= \left|\mathbb{P}\left(\left[X \leqslant \frac{x-\beta}{\alpha}\right]\right) - \mathbb{P}\left(\left[N \leqslant \frac{x-\beta}{\alpha}\right]\right)\right| \qquad \textit{(car } \alpha > 0\textit{)} \\
&= \left|F_X\left(\frac{x-\beta}{\alpha}\right) - \Phi\left(\frac{x-\beta}{\alpha}\right)\right| \\
&= d_X\left(\frac{x-\beta}{\alpha}\right)
\end{aligned}$$

□

b) Montrer que pour tout x réel,

$$\lim_{n\to+\infty} \left(\mathbb{P}(a_nV_n + b_n \leqslant x) - \mathbb{P}(a_nN + b_n \leqslant x)\right) = 0$$

Démonstration.
Soit $(\delta_n)_{n\geqslant 1}$ une suite de limite nulle telle que

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, d_{V_n}(x) \leqslant \delta_n$$

Soit $x \in \mathbb{R}$. D'après la question précédente :

$$\begin{aligned}
|\mathbb{P}(a_nV_n + b_n \leqslant x) - \mathbb{P}(a_nN + b_n \leqslant x)| &= |F_{a_nV_n+b_n}(x) - F_{a_nN+b_n}(x)| \\
&= d_{V_n}\left(\frac{x-b_n}{a_n}\right) \qquad \textit{(}\forall n \in \mathbb{N}^*, a_n > 0\textit{)} \\
&\leqslant \delta_n
\end{aligned}$$

Or, $\delta_n \underset{n\to+\infty}{\longrightarrow} 0$ donc, par théorème d'encadrement,

$$\boxed{\lim_{n\to+\infty} \left(\mathbb{P}(a_nV_n + b_n \leqslant x) - \mathbb{P}(a_nN + b_n \leqslant x)\right) = 0}$$

□

c) Établir que $\lim\limits_{n\to+\infty} \mathbb{P}(a_nN + b_n \leqslant x) = \mathbb{P}(aN + b \leqslant x)$ puis en déduire que $(a_nV_n + b_n)_{n\geqslant 1}$ converge en loi vers $aN + b$. Quelle est la loi de la variable aléatoire $aN + b$?

Démonstration.

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

$$\mathbb{P}(a_n N + b_n \leqslant x) - \mathbb{P}(aN + b \leqslant x) = \mathbb{P}\left(N \leqslant \frac{x - b_n}{a_n}\right) - \mathbb{P}\left(N \leqslant \frac{x-b}{a}\right) \quad (a_n > 0 \text{ } et \text{ } a > 0)$$
$$= \Phi\left(\frac{x - b_n}{a_n}\right) - \Phi\left(\frac{x-b}{a}\right)$$

Or, $a_n \underset{n \to +\infty}{\longrightarrow} a$ et $b_n \underset{n \to +\infty}{\longrightarrow} b$ donc $\dfrac{x - b_n}{a_n} \underset{n \to +\infty}{\longrightarrow} \dfrac{x-b}{a}$. De plus, Φ est continue en $\dfrac{x-b}{a}$.
Donc
$$\Phi\left(\frac{x - b_n}{a_n}\right) - \Phi\left(\frac{x-b}{a}\right) \underset{n \to +\infty}{\longrightarrow} 0$$
D'où
$$\boxed{\lim_{n \to +\infty} \mathbb{P}(a_n N + b_n \leqslant x) = \mathbb{P}(aN + b \leqslant x).}$$
De plus,
$$\mathbb{P}(a_n V_n + b_n \leqslant x) - \mathbb{P}(aN + b \leqslant x)$$
$$= (\mathbb{P}(a_n V_n + b_n \leqslant x) - \mathbb{P}(a_n N + b_n \leqslant x)) + (\mathbb{P}(a_n N + b_n \leqslant x) - \mathbb{P}(aN + b \leqslant x))$$
et

× $\lim\limits_{n \to +\infty} (\mathbb{P}(a_n V_n + b_n \leqslant x) - \mathbb{P}(a_n N + b_n \leqslant x)) = 0$

× $\lim\limits_{n \to +\infty} (\mathbb{P}(a_n N + b_n \leqslant x) - \mathbb{P}(aN + b \leqslant x)) = 0$

d'où
$$\mathbb{P}(a_n V_n + b_n \leqslant x) - \mathbb{P}(aN + b \leqslant x) \underset{n \to +\infty}{\longrightarrow} 0$$
Ceci étant vrai pour tout $x \in \mathbb{R}$:
$$\boxed{a_n V_n + b_n \underset{n \to +\infty}{\overset{\mathscr{L}}{\longrightarrow}} aN + b.}$$

- Ensuite, $N \hookrightarrow \mathcal{N}(0,1)$ et $a > 0$ donc, par transformation affine, $aN + b$ suit une loi normale.
De plus,

× $\mathbb{E}(aN + b) = a\mathbb{E}(N) + b = b$

× $\mathbb{V}(aN + b) = a^2 \mathbb{V}(N) = a^2$

$$\boxed{\text{donc } aN + b \hookrightarrow \mathcal{N}\left(b, a^2\right)}$$

□

21. On reprend les notations de la partie 3.

a) Justifier que $(D_n)_{n \geqslant 1}$ converge uniformément en loi vers N.

Démonstration.
D'après les résultats de la partie 3, on a, pour tout $n \in \mathbb{N}^*$,
$$D_n = \frac{C_n - np_n}{\sigma_n} = \frac{C_n - np_n}{\sqrt{np_n(1-p_n)}}$$
où

× $C_n \hookrightarrow \mathcal{B}(n, p_n)$

× $\lim\limits_{n \to +\infty} p_n = 0$ (car $h_n \underset{n \to +\infty}{\longrightarrow} 0$)

× $\lim\limits_{n\to+\infty} np_n = +\infty$ (car $nh_n \underset{n\to+\infty}{\longrightarrow} +\infty$)

D'après la question ***19.c)***, $\left(\dfrac{C_n - np_n}{\sqrt{np_n(1-p_n)}}\right)_{n\geqslant 1}$ converge uniformément en loi vers N.

$$\boxed{\text{Autrement dit, } (D_n)_{n\geqslant 1} \text{ converge uniformément en loi vers } N.}$$

□

b) En utilisant les résultats des questions ***14.*** et ***20.***, en déduire que la suite $(\hat{f}_n)_{n\geqslant 1}$ converge en loi vers N.

Démonstration.
Récapitulons les résultats de la question ***14.*** :

× pour tout $n \in \mathbb{N}^*$, $\hat{f}_n = a_n D_n + b_n$ où $a_n = \dfrac{\sigma_n}{\theta_n\sqrt{n}}$ et $b_n = \sqrt{n}\left(\dfrac{p_n}{\theta_n} - \theta_n\right)$

× pour tout $n \in \mathbb{N}^*$, $a_n > 0$ et $\lim\limits_{n\to+\infty} a_n = 1 > 0$

× $\lim\limits_{n\to+\infty} b_n = 0$

D'après la question précédente, $(D_n)_{n\geqslant 1}$ converge uniformément en loi vers N.
Ainsi, d'après la question ***20.c)***, $(a_n D_n + b_n)_{n\geqslant 1}$ converge en loi vers $1 \times N + 0 = N$.

$$\boxed{\text{Autrement dit, } (\hat{f}_n)_{n\geqslant 1} \text{ converge en loi vers } N.}$$

□