

Partie 1 - Transformation de Stein

Soit $h \in W$. On définit sur $\mathbb{R}$, la fonction $\theta : x \mapsto \frac{\Phi(x)}{\varphi(x)}$ 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.

1. a) Pour tout $x \geq 0$ et pour tout $t \in [x; +\infty[$, on a $0 \leq x \leq t$, et comme $\varphi(t)$ est toujours positif car φ est une densité, la multiplication de tous les membres de la double inégalité par $\varphi(t)$ donne :

$$0 \leq x\varphi(t) \leq t\varphi(t).$$

Les fonctions concernées par cette double inégalité sont continues sur $\mathbb{R}$ selon la variable t .

De plus, l'intégrale $\int_x^{+\infty} \varphi(t) dt$ converge et vaut $\mathbf{P}(X > x)$; comme l'énoncé le fait remarquer,

on a par ailleurs : $\varphi'(t) = \frac{1}{\sqrt{2\pi}} \cdot (-t) \cdot e^{-\frac{t^2}{2}} = -t\varphi(t)$, donc $\int_x^{+\infty} t\varphi(t) dt = -\int_x^{+\infty} \varphi'(t) dt$ converge et vaut $-(\lim_{t \rightarrow +\infty} \varphi(t) - \varphi(x)) = \varphi(x)$.

Les propriétés de positivité et de croissance de l'intégrale peuvent donc ici s'appliquer, qui donnent pour tout $x \geq 0$:

$$0 \leq \int_x^{+\infty} x\varphi(t) dt \leq \int_x^{+\infty} t\varphi(t) dt \iff 0 \leq x\mathbf{P}(N > x) \leq x\varphi(x) \iff 0 \leq x(1 - \Phi(x)) \leq x\varphi(x).$$

- b) Pour tout réel x négatif, et pour tout $t \in]-\infty; x]$, on a $t \leq x \leq 0 \implies t\varphi(t) \leq x\varphi(t) \leq 0$, soit $-\varphi(t) \leq t\varphi(x) \leq 0$.

Les fonctions concernées sont continues et leurs intégrales sur $] -\infty; x]$ convergent, donc par croissance de l'intégrale :

$$-\int_{-\infty}^x \varphi'(t) dt \leq x \int_{-\infty}^x \varphi(t) dt \leq 0 \iff -(\varphi(x) - \underbrace{\lim_{x \rightarrow -\infty} \varphi(x)}_{=0}) \leq x\mathbf{P}(N \leq x) \leq 0 \iff -\varphi(x) \leq x\Phi(x) \leq 0.$$

- c) Soit x un réel quelconque, et A un réel négatif et inférieur à x : dans l'intégrale $\int_A^x \Phi(t) dt$, on pose :

$$u(t) = \Phi(t) \longrightarrow u'(t) = \varphi(t)$$

$$v'(t) = 1 \longrightarrow v(t) = t$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc par intégration par parties :

$$\int_A^x \Phi(t) dt = \left[t\Phi(t) \right]_A^x - \int_A^x \underbrace{t\varphi(t)}_{=-\varphi'(t)} dt = x\Phi(x) - A\Phi(A) + \int_A^x \varphi'(t) dt = x\Phi(x) - A\Phi(A) + \varphi(x) - \varphi(A).$$

D'après b) : $-\varphi(A) \leq A\Phi(A) \leq 0$ avec $\lim_{A \rightarrow -\infty} \varphi(A) = 0$, donc par encadrement, $\lim_{A \rightarrow -\infty} A\Phi(A) = 0$.

On peut donc passer à la limite dans l'égalité ci-dessus lorsque A tend vers $-\infty$, ce qui donne la relation vraie pour tout réel x :

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

De même, pour x réel quelconque, soit B un réel positif et supérieur à x .

Dans l'intégrale $\int_x^B (1 - \Phi(t)) dt$, on pose :

$$\begin{aligned} u(t) &= 1 - \Phi(t) & \longrightarrow & u'(t) = -\varphi(t) \\ v'(t) &= 1 & \longrightarrow & v(t) = t \end{aligned}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc par intégration par parties :

$$\int_x^B (1 - \Phi(t)) dt = \left[t(1 - \Phi(t)) \right]_x^B + \int_x^B \underbrace{t\varphi(t)}_{=-\varphi'(t)} dt = B(1 - \Phi(B)) - x(1 - \Phi(x)) - \varphi(B) + \varphi(x).$$

L'encadrement : $0 \leq B(1 - \Phi(B)) \leq \varphi(B)$ et le fait que $\lim_{B \rightarrow +\infty} \varphi(B) = 0$ assurent d'après le même théorème que précédemment, que $\lim_{B \rightarrow +\infty} B(1 - \Phi(B)) = 0$, et donc qu'on peut passer à la limite quand B tend vers $+\infty$ dans l'égalité précédente. On obtient ainsi, pour tout réel x :

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

ce qui achève de démontrer les deux égalités demandées, appelées ensemble (R_1) par l'énoncé.

2. a) D'après les hypothèses faites dans l'énoncé : la fonction $h \in W$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et vérifie $\forall x \in \mathbb{R}, |h'(x)| \leq 1$.

On peut donc appliquer l'Inégalité des Accroissements Finis à la fonction h avec le majorant $K = 1$ pour $|h'|$; ce résultat de cours assure directement que :

$$\forall (x, y) \in \mathbb{R}^2, |h(x) - h(y)| \leq K|x - y| \iff |f(x) - f(y)| \leq |x - y|.$$

Si on écrit cette inégalité pour $y = 0$, on obtient : $\forall x \in \mathbb{R}, |h(x) - h(0)| \leq |x|$.

Or d'après l'inégalité triangulaire pour la valeur absolue :

$$|h(x)| = |(h(x) - h(0)) + h(0)| \leq |h(x) - h(0)| + |h(0)|,$$

donc on a bien :

$$\forall x \in \mathbb{R}, |h(x)| \leq |x| + |h(0)|.$$

- b) Pour tout réel x : $0 \leq |h'(x)\Phi(x)| = |h'(x)|\Phi(x) \leq \Phi(x)$ car $|h'(x)| \leq 1$ pour tout réel, tandis que Φ est positive sur $\mathbb{R}$ (c'est une fonction de répartition).

Les fonctions concernées étant continues et positives sur $\mathbb{R}$, la convergence de l'intégrale $\int_{-\infty}^x \Phi(t) dt$ assure alors la convergence absolue de $\int_{-\infty}^x h'(t)\Phi(t) dt$, donc sa convergence, d'après le théorème de comparaison des intégrales de fonctions continues, positives.

On réalise alors une nouvelle intégration par parties dans l'intégrale $\int_A^x h'(t)\Phi(t) dt$ ($x \in \mathbb{R}$ et $A \leq x$), en posant :

$$\begin{aligned} u(t) &= \Phi(t) & \longrightarrow & u'(t) = \varphi(t) \\ v'(t) &= h'(t) & \longrightarrow & v(t) = h(t) \end{aligned}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc :

$$\int_A^x h'(t)\Phi(t)dt = \left[h(t)\Phi(t) \right]_A^x - \int_A^x h(t)\varphi(t)dt = h(x)\Phi(x) - h(A)\Phi(A) - \int_A^x h(t)\varphi(t)dt.$$

Or pour tout réel A , d'après l'inégalité obtenue en 2.a) :

$$|h(A)\Phi(A)| = |h(A)|\Phi(A) \leq |A|\Phi(A) + |h(0)|\Phi(A) = |A\Phi(A)| + |h(0)|\Phi(A),$$

où on a déjà vu que $\lim_{A \rightarrow -\infty} A\Phi(A) = 0$ et $\lim_{A \rightarrow -\infty} \Phi(A) = 0$ (car Φ est une fonction de répartition) : par encadrement, $\lim_{A \rightarrow -\infty} |A\Phi(A)| = 0 = \lim_{A \rightarrow -\infty} A\Phi(A)$.

La convergence de l'intégrale $\int_{-\infty}^x h'(t)\Phi(t)dt$ et ce résultat assurent alors la convergence de $\int_{-\infty}^x h(t)\varphi(t)dt$, et on peut passer à la limite dans l'égalité issue de la dernière intégration par parties quand $A \rightarrow -\infty$, pour obtenir :

$$\forall x \in \mathbb{R}, \quad \int_{-\infty}^x h'(t)\Phi(t)dt = h(x)\Phi(x) - \int_{-\infty}^x h(t)\varphi(t)dt.$$

L'énoncé admettait alors que, de même, $\int_x^{+\infty} h'(t)(1 - \Phi(t))dt$ converge et que,

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

c) On en déduit donc que, pour tout x réel :

$$\begin{aligned} & - \int_{-\infty}^x h'(t)\Phi(t)dt + \int_x^{+\infty} h'(t)(1 - \Phi(t))dt \\ &= -h(x)\Phi(x) + \int_{-\infty}^x h(t)\varphi(t)dt - h(x) + h(x)\Phi(x) - \int_x^{+\infty} h(t)\varphi(t)dt \\ &= \int_{-\infty}^{+\infty} h(t)\varphi(t)dt - h(x) \\ &= E(h(N)) - h(x) = c_h - h(x) \quad \text{d'après le théorème de transfert, puisque } \varphi \text{ est une densité de } N \end{aligned}$$

3. a) La fonction $\theta : x \mapsto \frac{\Phi(x)}{\varphi(x)}$ est indéfiniment dérivable sur $\mathbb{R}$ comme quotient de fonctions qui le sont, et pour tout réel x : $\Phi'(x) = \varphi(x)$ et $\varphi'(x) = -x\varphi(x)$, donc

$$\theta'(x) = \frac{\Phi'(x) \cdot \varphi(x) - \Phi(x) \varphi'(x)}{(\varphi'(x))^2} = \frac{(\varphi(x))^2 + x\Phi(x)\varphi(x)}{(\varphi(x))^2} = 1 + x \frac{\Phi(x)}{\varphi(x)} = 1 + x\theta(x)$$

$$\theta''(x) = 0 + 1 \cdot \theta(x) + x\theta'(x) = \theta(x) + x(1 + x\theta(x)) = x + (1 + x^2)\theta(x),$$

$$\theta(-x)\Phi(x) = \frac{\Phi(-x)}{\varphi(-x)}\Phi(x) = \frac{1 - \Phi(x)}{\varphi(x)}\Phi(x) = \frac{\Phi(x)}{\varphi(x)}(1 - \Phi(x)) = \theta(x)(1 - \Phi(x)).$$

En effet pour la troisième relation, la fonction φ est paire sur $\mathbb{R}$ donc $\varphi(-x) = \varphi(x)$ pour tout réel x , et $\Phi(-x) = 1 - \Phi(x)$ pour tout réel x (propriété du cours de Φ).

b) Puisque les fonctions h et Φ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, et puisque les fonctions $a : x \mapsto \int_{-\infty}^x h'(t)\Phi(t)dt$

et $b : x \mapsto \int_x^{+\infty} h'(t)(1 - \Phi(t))dt$ sont bien définies sur $\mathbb{R}$, alors ces deux dernières fonctions sont aussi de classe $\mathcal{C}^1$ sur $\mathbb{R}$, avec :

$$\forall x \in \mathbb{R}, \quad a'(x) = h'(x)\Phi(x) \quad \text{et} \quad b'(x) = -h'(x)(1 - \Phi(x)).$$

La fonction $f_h : x \mapsto \theta(-x).a(x) + \theta(x).b(x)$ est donc de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme somme et produit de fonctions qui le sont, avec pour tout réel x :

$$\begin{aligned} f'_h(x) &= -\theta'(-x).a(x) + \theta(-x).a'(x) + \theta'(x).b(x) + \theta(x).b'(x) \\ &= -(1 - x\theta(-x)).a(x) + \theta(-x).h'(x)\Phi(x) + (1 + x\theta(x)).b(x) - \theta(x).h'(x)(1 - \Phi(x)) \\ &= -a(x) + b(x) + x(\theta(-x).a(x) + \theta(x).b(x)) + \theta(-x).h'(x)\Phi(x) - \theta(x).h'(x) + \theta(x).h'(x)\Phi(x) \end{aligned}$$

Or 2.c) dit que : $\forall x \in \mathbb{R}, -a(x) + b(x) = c_h - h(x)$, $\theta(-x).a(x) + \theta(x).b(x) = f_h(x)$ pour tout réel x ,

et $\theta(-x).h'(x)\Phi(x) + \theta(x).h'(x)\Phi(x) = \theta(x)(1 - \Phi(x)).h'(x) + \theta(x).h'(x)\Phi(x) = \theta(x).h'(x)$ d'après la troisième relation de 3.a), donc :

$$\forall x \in \mathbb{R}, \quad f'_h(x) = c_h - h(x) + x.f_h(x) + \theta(x).h'(x) - \theta(x).h'(x) \iff \boxed{\forall x \in \mathbb{R}, \quad f'_h(x) - x.f_h(x) = c_h - h(x)}$$

Cette relation s'écrit aussi : $\forall x \in \mathbb{R}, f'_h(x) = x f_h(x) + c_h - h(x)$. Sous cette forme, f'_h apparaît comme une somme (et un produit) de fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}$: f'_h est donc de classe $\mathcal{C}^1$ sur $\mathbb{R}$, ce qui implique bien que la fonction f_h est alors de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

- c) Si X est une variable aléatoire qui admet une espérance, alors la relation précédente donne l'égalité de variables aléatoires $f'_h(X) - X f_h(X) = c_h - h(X)$.

L'énoncé admettant que dans ce cas, $\mathbb{E}(h(X))$ existe, alors $c_h - h(X) = f'_h(X) - X f_h(X)$ admet une espérance, et on peut écrire :

$$|\mathbb{E}(c_h - h(X))| = |\mathbb{E}(f'_h(X) - X f_h(X))| \iff |\mathbb{E}(h(X)) - c_h| = |\mathbb{E}(f'_h(X) - X f_h(X))|$$

par linéarité de l'espérance et parité de la valeur absolue, ce qui est bien :

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

4. Majoration de $|f''_h|$.

- a) D'après les égalités (R_1) , pour tout réel x :

$$\begin{aligned} \theta(-x) \int_{-\infty}^x \Phi(t) dt + \theta(x) \int_x^{+\infty} (1 - \Phi(t)) dt &= x\theta(-x)\Phi(x) + \theta(-x)\varphi(x) - x\theta(x)(1 - \Phi(x)) + \theta(x)\varphi(x) \\ &\stackrel{3.a)}{=} x\theta(x)(1 - \Phi(x)) + (\theta(-x) + \theta(x))\varphi(x) - x\theta(x)(1 - \Phi(x)) \\ \text{or } \begin{cases} \Phi(-x) = 1 - \Phi(x) \\ \text{et } \varphi(-x) = \varphi(x) \end{cases} &= \left(\frac{\Phi(-x)}{\varphi(-x)} + \frac{\Phi(x)}{\varphi(x)} \right) \varphi(x) = \frac{1 - \Phi(x) + \Phi(x)}{\varphi(x)} \varphi(x) \\ &= 1. \end{aligned}$$

- b) Pour tout réel x , d'après l'inégalité triangulaire (pour la somme et les intégrales) :

$$\begin{aligned} |f_h(x)| &= \left| \theta(-x) \int_{-\infty}^x h'(t)\Phi(t) dt + \theta(x) \int_x^{+\infty} h'(t)(1 - \Phi(t)) dt \right| \\ &\leq \theta(-x) \left| \int_{-\infty}^x h'(t)\Phi(t) dt \right| + \theta(x) \left| \int_x^{+\infty} h'(t)(1 - \Phi(t)) dt \right| \quad \text{car } \theta(-x) \geq 0 \text{ et } \theta(x) \geq 0 \\ &\leq \theta(-x) \int_{-\infty}^x |h'(t)|\Phi(t) dt + \theta(x) \int_x^{+\infty} |h'(t)|(1 - \Phi(t)) dt \end{aligned}$$

Or : $\forall t \in \mathbb{R}, |h'(t)| \leq 1$ donc $|h'(t)|\Phi(t) \leq \Phi(t)$ et $|h'(t)|(1 - \Phi(t)) \leq 1 - \Phi(t)$ puisque $0 \leq \Phi(t) \leq 1$.

Par croissance de l'intégrale (toutes les intégrales convergent), on a donc :

$$\forall x \in \mathbb{R}, \quad \int_{-\infty}^x |h'(t)|\Phi(t)dt \leq \int_{-\infty}^x \Phi(t)dt \quad \text{et} \quad \int_x^{+\infty} |h'(t)|(1-\Phi(t))dt \leq \int_x^{+\infty} (1-\Phi(t))dt,$$

donc puisque $\theta(-x)$ et $\theta(x)$ sont positifs, par transitivité de l'inégalité, on a pour tout réel x :

$$|f_h(x)| \leq \theta(-x) \int_{-\infty}^x \Phi(t)dt + \theta(x) \int_x^{+\infty} (1-\Phi(t))dt \iff |f_h(x)| \leq 1 \quad \text{d'après 4.b).}$$

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

a) Il suffit ici de reprendre la relation : $\forall x \in \mathbb{R}, \theta''(x) = x + (1+x^2)\theta(x)$ et de repartir de l'égalité obtenue en 4.a), pour pouvoir écrire, pour tout réel x :

$$\begin{aligned} & \theta''(-x) \int_{-\infty}^x \Phi(t)dt + \theta''(x) \int_x^{+\infty} (1-\Phi(t))dt \\ &= (-x + (1+(-x)^2)\theta(-x)) \int_{-\infty}^x \Phi(t)dt + (x + (1+x^2)\theta(x)) \int_x^{+\infty} (1-\Phi(t))dt \\ &= x \left(- \int_{-\infty}^x \Phi(t)dt + \int_x^{+\infty} (1-\Phi(t))dt \right) + (1+x^2) \left(\theta(-x) \int_{-\infty}^x \Phi(t)dt + \theta(x) \int_x^{+\infty} (1-\Phi(t))dt \right) \\ &\stackrel{(R_1)}{=} x(-x\Phi(x) - \varphi(x) - x(1-\Phi(x)) + \varphi(x)) + (1+x^2) \times 1 \quad \leftarrow \text{d'après 4.a)} \\ &= -x^2 + 1 + x^2 = 1. \end{aligned}$$

b) En reprenant les notations de la question 3.b) où on avait déjà dérivé f_h sur $\mathbb{R}$, et en dérivant une fois de plus (f_h est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ d'après cette question précédente) : pour tout $x \in \mathbb{R}$,

$$f_h'(x) = -\theta'(-x).a(x) + \theta(-x).a'(x) + \theta'(x).b(x) + \theta(x).b'(x)$$

$$\text{où } \theta(-x).a'(x) + \theta(x).b'(x) = h'(x)\theta(-x)\Phi(x) - h'(x)\theta(x)(1-\Phi(x)) = 0$$

$$\text{car } \theta(-x)\Phi(x) = \theta(x)(1-\Phi(x)) \text{ d'après 3.a),}$$

$$\text{donc } f_h''(x) = +\theta''(-x).a(x) - \theta'(-x).a'(x) + \theta''(x).b(x) + \theta'(x).b'(x),$$

où :

$$\begin{aligned} -\theta'(-x).a'(x) + \theta'(x).b'(x) &= -(1-x\theta(-x)).h'(x)\Phi(x) - (1+x\theta(x)).h'(x)(1-\Phi(x)) \\ &= h'(x) \left(-\Phi(x) + x\theta(x)(1-\Phi(x)) - 1 + \Phi(x) - x\theta(x) + x\theta(x)\Phi(x) \right) \\ &= -h'(x), \end{aligned}$$

de sorte qu'on a bien :

$$\forall x \in \mathbb{R}, \quad f_h''(x) = -h'(x) + \theta''(-x) \int_{-\infty}^x h'(t)\Phi(t)dy + \theta''(x) \int_x^{+\infty} h'(t)(1-\Phi(t))dt.$$

c) La fonction $g : x \mapsto \Phi(x) + \frac{x}{1+x^2}\varphi(x)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme somme et produit de fonctions qui le sont, et :

$$\begin{aligned} \forall x \in \mathbb{R}, \quad g'(x) &= \Phi'(x) + \frac{1 \cdot (1+x^2) - x \cdot 2x}{(1+x^2)^2} \varphi(x) + \frac{x}{1+x^2} \varphi'(x) = \varphi(x) + \frac{1-x^2}{(1+x^2)^2} \varphi(x) - \frac{x^2}{1+x^2} \varphi(x) \\ &= \frac{(1+x^2)^2 + 1 - x^2 - x^2(1+x^2)}{(1+x^2)^2} \varphi(x) = \frac{1+2x^2+x^4+1-2x^2-x^4}{(1+x^2)^2} \varphi(x) \\ &= \frac{2}{(1+x^2)^2} \cdot \varphi(x). \end{aligned}$$

Pour tout réel x , $\frac{2}{(1+x^2)^2} > 0$ et $\varphi(x) > 0$, donc $g'(x) > 0$ et g est strictement croissante sur $\mathbb{R}$.

En remarquant que puisque $\lim_{x \rightarrow -\infty} \Phi(x) = 0$, $\frac{x}{1+x^2} \underset{x \rightarrow -\infty}{\sim} \frac{1}{x}$ donc $\lim_{x \rightarrow -\infty} \frac{x}{1+x^2} = 0$ et puisque $\lim_{x \rightarrow -\infty} \varphi(x) = 0$ aussi, alors $\lim_{x \rightarrow -\infty} g(x) = 0$: la fonction g strictement croissante sur $\mathbb{R}$ et de limite nulle en $-\infty$, est alors strictement positive sur $\mathbb{R}$.

Mais alors, pour tout réel x , en reprenant la formule obtenue en 3.a) :

$$\theta''(x) = x + (1+x^2)\theta(x) = x + (1+x^2)\frac{\Phi(x)}{\varphi(x)} = \frac{1+x^2}{\varphi(x)} \left(\frac{x}{1+x^2}\varphi(x) + \Phi(x) \right) = \frac{1+x^2}{\varphi(x)} \cdot g(x) > 0,$$

comme produit et quotient de termes tous strictement positifs sur $\mathbb{R}$.

On a alors, par une démarche analogue à celle qui a été suivie en 4.b) :

$$\begin{aligned} |f_h''(x)| &\leq |h'(x)| + \theta''(-x) \left| \int_{-\infty}^x h'(t)\Phi(t)dt \right| + \theta''(x) \left| \int_x^{+\infty} h'(t)(1-\Phi(t))dt \right| \\ &\leq 1 + \theta''(-x) \int_x^{+\infty} |h'(t)|\Phi(t)dt + \theta''(x) \int_x^{+\infty} |h'(t)|(1-\Phi(t))dt \quad \text{car } |h'(x)| \leq 1 \\ &\leq 1 + \theta''(-x) \underbrace{\int_x^{+\infty} \Phi(t)dt + \theta''(x) \int_x^{+\infty} (1-\Phi(t))dt}_{=1 \text{ d'après 5.a)}, \end{aligned}$$

ce qui donne bien : $\forall x \in \mathbb{R}, \quad |f_h''(x)| \leq 2.$

Partie 2 - Majoration uniforme de la distance de Kolmogorov

6. Par définition de la fonction h_x , la variable aléatoire $h_x(X)$ ne peut prendre que les valeurs 0 ou 1 : c'est donc une variable de Bernoulli, avec :

$$\mathbf{P}(h_x(X) = 1) = \mathbf{P}(X \leq x) = F_X(x), \quad \text{et } \mathbf{P}(h_x(X) = 0) = \mathbf{P}(X > x) = 1 - F_X(x).$$

La variable aléatoire finie $h_x(X)$ admet donc une espérance, qui vaut :

$$\mathbb{E}(h_x(X)) = \mathbf{P}(h_x(X) = 1) = F_X(x).$$

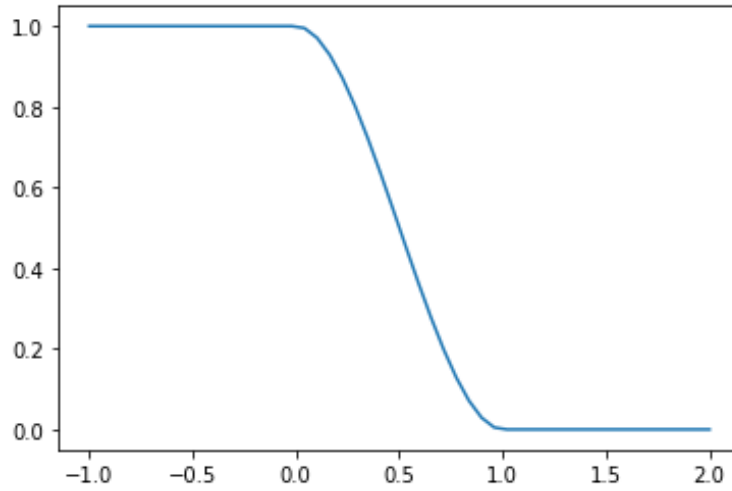
7. a) La fonction gamma demandée est très facile à programmer avec une structure de test :

```
def gamma(t):
    if t < 0 :
        return 1
    if t > 1 :
        return 0
    return 1-3*t**2+2*t**3 # si les deux tests précédents sont faux
```

b) On écrit ici les commandes qui, avec la bibliothèque `matplotlib.pyplot`, permettent d'afficher le graphe de la fonction γ sur l'intervalle $[-1; 2]$.

```
import matplotlib.pyplot as plt
import numpy as np

abscisses = np.linspace(-1,2)
images = [gamma(x) for x in abscisses]
plt.plot(abscisses,images)
plt.show()
```



Ci-dessus le résultat rendu par Python.

8. a) La fonction γ est de classe $\mathcal{C}^1$ donc continue sur chacun des intervalles $]-\infty; 0[$, $]1; +\infty[$ et $]0; 1[$: c'est une fonction constante sur chacun des deux premiers intervalles, et une fonction polynômiale sur le troisième intervalle.

De plus : $\lim_{t \rightarrow 0^+} \gamma(t) = \lim_{t \rightarrow 0^+} 1 - 3t^2 + 2t^3 = 1 = \gamma(1) = \lim_{t \rightarrow 0^-} \gamma(t)$, donc γ est continue en 0.

De même : $\lim_{t \rightarrow 1^-} \gamma(t) = \lim_{t \rightarrow 1^-} 1 - 3t^2 + 2t^3 = 1 - 3 + 2 = 0 = \gamma(0) = \lim_{t \rightarrow 1^+} \gamma(t)$, donc γ est aussi continue en 1, et finalement γ est continue sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé de 0 et 1.

- b) Pour tout $t \in]0; 1[$: $\gamma'(t) = -6t + 6t^2 = 6t(t - 1) < 0$ puisque $t > 0$ et $t < 1$, donc la fonction γ est strictement décroissante sur $]0; 1[$, de sorte que :

$$\forall t \in [0; 1], \quad \gamma(1) \leq \gamma(t) \leq \gamma(0) \iff 0 \leq \gamma(t) \leq 1.$$

- c) On étudie la dérivabilité de la fonction γ en 1 en calculant les limites à gauche et à droite en 1, du taux d'accroissement $\frac{\gamma(t) - \gamma(1)}{t - 1}$.

$$\text{Si } t > 1 : \quad \frac{\gamma(t) - \gamma(1)}{t - 1} = \frac{0 - 0}{t - 1} = 0 \xrightarrow{t \rightarrow 1^+} 0.$$

$$\text{Si } 0 < t < 1 : \quad \frac{\gamma(t) - \gamma(1)}{t - 1} = \frac{1 - 3t^2 + 2t^3}{t - 1}. \text{ Or la division euclidienne de } 1 - 3t^2 + 2t^3 \text{ par } t - 1 \text{ est exacte (en fait car 1 est racine évidente du polynôme de degré 3), et si on l'écrit on obtient : } 1 - 3t^2 + 2t^3 = (t - 1)(2t^2 - t - 1), \text{ donc pour } 0 < t < 1, \frac{1 - 3t^2 + 2t^3}{t - 1} = 2t^2 - t - 1 \xrightarrow{t \rightarrow 1^-} 2 - 1 - 1 = 0.$$

On en conclut donc, vu que les limites à gauche et à droite sont égales, que : $\lim_{t \rightarrow 1} \frac{\gamma(t) - \gamma(1)}{t - 1} = 0$, ce qui prouve que la fonction γ est bien dérivable en 1, avec $\gamma'(1) = 0$.

L'énoncé admettait de même que γ est dérivable en 0 avec $\gamma'(0) = 0$.

- d) La fonction γ est déjà de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé de 0 et de 1, mais dérivable sur tout $\mathbb{R}$, avec :

$$\forall t \in \mathbb{R}, \quad \gamma'(t) = \begin{cases} 0 & \text{si } t \leq 0 \text{ ou } t \geq 1 \\ -6t + 6t^2 & \text{si } 0 < t < 1. \end{cases}$$

En 0 : $\lim_{t \rightarrow 0^-} \gamma'(t) = \lim_{t \rightarrow 0^-} 0 = 0 = \gamma'(0) = \lim_{t \rightarrow 0^+} -6t + 6t^2$, donc γ' est continue en 0.

De même : $\lim_{t \rightarrow 1^+} \gamma'(t) = 0 = \gamma'(1) = \lim_{t \rightarrow 1^-} -6t + 6t^2$, donc γ' est aussi continue en 1.

Finalement, γ' est continue sur $\mathbb{R}$, donc γ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Pour tout $t \in]-\infty; 0] \cup [1; \infty[$, $\gamma'(t) = 0$ et pour tout $t \in]0; 1]$: $|\gamma'(t)| = 6t - 6t^2$, qui est un trinôme du second degré de racines évidentes 0 et 1.

Le sommet de la parabole (ici tournée vers le bas se situe donc au milieu des deux racines, en $t = \frac{1}{2}$, et donne la valeur maximale de $6t - 6t^2$ sur $[0; 1]$: $|\gamma'(\frac{1}{2})| = 3 - \frac{6}{4} = \frac{3}{2}$.

Ce maximum sur $[0; 1]$ de $|\gamma'|$, est en fait valable sur tout $\mathbb{R}$: $\forall t \in \mathbb{R}, |\gamma'(t)| \leq \frac{3}{2}$.

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

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

a) Pour tout réel y : si $y \leq x$, alors $h_x(y) = 1$ et $\frac{y-x}{t} \leq 0$, donc $\gamma(\frac{y-x}{t}) = 1$ aussi et $h_x(y) = k_x(y)$.

Sinon, $y > x$ et alors $h_x(y) = 0$.

Comme la fonction γ est à valeurs positives, on a donc $h_x(y) \leq \gamma(\frac{y-x}{t}) = k_x(y)$.

On a donc bien démontré que : $\forall y \in \mathbb{R}, h_x(y) \leq k_x(y)$.

b) L'énoncé admettait l'existence de $\mathbb{E}(k_x(X))$ et de $\mathbb{E}(k_x(N))$.

L'inégalité obtenue à la question précédente assure que $h_x(X) \leq k_x(X)$, et par croissance de l'espérance, on a alors :

$$\mathbb{E}(h_x(X)) \leq \mathbb{E}(k_x(X)) \iff \mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leq \mathbb{E}(k_x(X)) - \mathbb{E}(h_x(N)),$$

soit en effet : $\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) \leq \mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N)) + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N))$.

c) Si on examine de plus près les expressions comparées des fonctions h_x et k_x , on réalise que :

- Si $y \leq x$, alors $h_x(y) = 1$ et $\frac{y-x}{t} \leq 0$ donc $k_x(y) = \gamma(\frac{y-x}{t}) = 1$ aussi.
- Si $y > x + t$, alors $y > x$ puisque $t > 0$ et $h_x(y) = 0$, tandis que $y - x > t \iff \frac{y-x}{t} > 1$ et on a $k_x(y) = 0$ aussi.

Les espérances $\mathbb{E}(k_x(N))$ et $\mathbb{E}(h_x(N))$ sont données par le théorème de transfert, et grâce aussi à la relation de Chasles, on peut écrire :

$$\begin{aligned} \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) &= \int_{-\infty}^{+\infty} k_x(u) \varphi(u) du - \int_{-\infty}^{+\infty} h_x(u) \varphi(u) du \\ &= \int_{-\infty}^x \varphi(u) du + \int_x^{x+t} k_x(u) \varphi(u) dtu + \int_{x+t}^{+\infty} 0 \cdot \varphi(u) du - \int_{-\infty}^x \varphi(u) du + \int_x^{+\infty} 0 \cdot \varphi(u) du \\ &= \int_x^{x+t} k_x(u) \varphi(u) du. \end{aligned}$$

d) Soit la fonction $g : u \mapsto \frac{2t}{3} k_x(u)$, bien définie sur $\mathbb{R}$.

La fonction $k_x : u \mapsto \gamma(\frac{u-x}{t})$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme composée de deux fonctions qui le sont, à savoir la fonction affine $u \mapsto \frac{u-x}{t}$ et la fonction γ . La fonction g est donc de classe $\mathcal{C}^1$ sur $\mathbb{R}$ aussi, puisqu'on l'obtient en multipliant la fonction k_x par la constante $\frac{2t}{3}$.

Pour tout réel u , on a ensuite : $g'(u) = \frac{2t}{3} k'_x(u) = \frac{2t}{3} \times \frac{1}{t} \gamma'(\frac{u-x}{t}) = \frac{2}{3} \gamma'(\frac{u-x}{t})$, donc d'après la majoration de $|\gamma'|$ obtenue en 8.d) :

$$\forall u \in \mathbb{R}, |g'(u)| = \frac{2}{3} |\gamma'(\frac{u-x}{t})| \leq \frac{2}{3} \times \frac{3}{2}, \text{ soit : } \forall u \in \mathbb{R}, |g'(u)| \leq 1,$$

ce qui achève de démontrer que la fonction g appartient à W .

L'hypothèse faite par l'énoncé au début de cette question 9., permet alors d'écrire que :

$$|\mathbb{E}(g(X)) - \mathbb{E}(g(N))| \leq M_X \iff \frac{2t}{3} |\mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N))| \leq M_X \iff |\mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N))| \leq \frac{3}{2t} M_X,$$

par linéarité de l'espérance, et du fait que $\frac{2t}{3} > 0$.

D'après 9.c), on a aussi, puisque les fonctions γ donc k_x , et φ sont positives sur $\mathbb{R}$:

$$|\mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N))| = \int_x^{x+t} k_x(u) \varphi(u) du, \text{ où :}$$

$$\forall u \in \mathbb{R}, 0 \leq k_x(u) = \gamma\left(\frac{u-x}{t}\right) \leq 1 \text{ et } 0 < \varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \leq \frac{1}{\sqrt{2\pi}}, \text{ donc } 0 \leq k_x(u) \varphi(u) \leq \frac{1}{\sqrt{2\pi}}.$$

Les fonctions concernées sont continues sur $\mathbb{R}$, donc d'après l'inégalité de la moyenne :

$$0 \leq \int_x^{x+t} k_x(u) \varphi(u) du \leq (x+t-x) \cdot \frac{1}{\sqrt{2\pi}} = \frac{ty}{\sqrt{2\pi}}.$$

On met tout cela en lien avec le résultat de 9.b), et on conclut alors que :

$$\begin{aligned} \mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N)) &\leq \mathbb{E}(k_x(X)) - \mathbb{E}(k_x(N)) + \mathbb{E}(k_x(N)) - \mathbb{E}(h_x(N)) \\ &\leq \frac{3}{2t} M_X + \frac{t}{\sqrt{2\pi}} \leq \frac{3}{2t} M_X + \frac{t}{2}, \end{aligned}$$

$$\text{puisque } 6 < 2\pi < 8 \implies 2 < \sqrt{2\pi} < 3 \implies \frac{t}{\sqrt{2\pi}} < \frac{t}{2}.$$

L'énoncé admettait alors qu'en utilisant la fonction k_{x-t} , on a :

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

ce qui implique donc avec l'inégalité précédemment démontrée, que :

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

10. La fonction $\psi : t \mapsto \frac{3}{2t} M_X + \frac{t}{2}$ est dérivable sur $]0; +\infty[$, avec :

$$\forall t \in]0; +\infty[, \quad \psi'(t) = -\frac{3}{2t^2} M_X + \frac{1}{2} = \frac{t^2 - 3M_X}{2t^2}.$$

Pour tout $t > 0$: $\psi'(t) > 0 \iff t^2 - 3M_X > 0 \iff t^2 > 3M_X \iff t > \sqrt{3M_X}$, donc ψ est décroissante sur $]0; \sqrt{3M_X}]$, puis croissante sur $[\sqrt{3M_X}; +\infty[$.

Cette fonction ψ admet donc un minimum en $t = \sqrt{3M_X}$, qui vaut :

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

de sorte que : $\forall t \in]0; +\infty[, \quad \frac{3}{2t} M_X + \frac{t}{2} \leq \sqrt{3M_X}.$

L'inégalité obtenue à la question précédente assure alors que, pour tout x réel :

$$|\mathbb{E}(h_x(X)) - \mathbb{E}(h_x(N))| \leq \sqrt{3M_X}.$$

Comme on a vu à la question 6. : $\mathbb{E}(h_x(X)) = F_X(x)$, et comme ce résultat est valable pour tout variable aléatoire, on a aussi $\mathbb{E}(h_x(N)) = \Phi(x)$, donc l'inégalité qu'on vient d'obtenir se réécrit effectivement, pour tout réel x :

$$d_X(x) = |F_X(x) - \Phi(x)| \leq \sqrt{3M_X}.$$

Partie 3 - Estimation d'une densité

11. L'énoncé 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) Les instructions suivantes permettent d'accéder aux valeurs de la colonne `salaire` et d'affecter cette série `pandas` obtenue à une variable `échantillon` :

```
import pandas as pd

stats = pd.read_csv("stats.csv", delimiter = ",")
échantillon = stats["salaire"]
```

On sélectionne ainsi une sous-table à une seule colonne correspondant aux statistiques sur le `salaire`, dont les entrées restent indicées de la même façon que les entrées complètes du fichier originel.

- 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 $n \in \mathbb{N}^*$, $h_n = \frac{1}{\sqrt{n}}$.

Le script suivant réalise cette tâche :

```
a = float(input('a = '))
n = échantillon.count() # n est donc la taille de l'échantillon
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)) # calcul de f_n(omega)
```

- c) Pour tout entier $i \in \llbracket 1; n \rrbracket$, on peut associer à la variable aléatoire X_i , la variable aléatoire Y_i qui vaut 1 si $[a - h_n \leq X_i \leq a + h_n]$ est réalisé, et 0 sinon.

Comme les (X_i) sont mutuellement indépendantes et de même loi, alors les (Y_i) sont des variables aléatoires de Bernoulli elles-mêmes mutuellement indépendantes d'après le lemme des coalitions, et de même loi $\mathcal{B}(p_n)$, où $p_n = \mathbf{P}(Y_n = 1) = \mathbf{P}(a - h_n \leq X \leq a + h_n)$ ne dépend pas de i et correspond à la probabilité de succès pour Y_i .

De ce point de vue, la variable aléatoire C_n est donc celle qui compte le nombre de succès obtenus au cours de n épreuves identiques et mutuellement indépendantes : elle suit donc la loi binomiale de paramètres (n, p_n) (et en fait, $C_n = Y_1 + \dots + Y_n$).

On en déduit d'après le cours sur la loi binomiale, que C_n admet une espérance qui vaut $\mathbb{E}(C_n) = np_n$; par linéarité de l'espérance, f_n admet donc une espérance qui vaut :

$$\mathbb{E}(f_n) = \frac{1}{2nh_n} \mathbb{E}(C_n) = \frac{p_n}{2h_n} = \frac{\mathbf{P}(a - h_n \leq X \leq a + h_n)}{2h_n} = \frac{F(a + h_n) - F(a - h_n)}{2h_n}.$$

12. a) La fonction de répartition F est dérivable en a (puisque la densité f est continue en ce point), donc elle y admet un développement limité à l'ordre 1 qui s'écrit :

$$F(a + h) = F(a) + hF'(a) + o(h) = F(a) + hf(a) + o(h) \quad \text{lorsque } h \text{ est au voisinage de } 0.$$

Puisque $\lim_{n \rightarrow +\infty} h_n = 0$, on peut alors écrire, lorsque n est au voisinage de $+\infty$:

$$\frac{F(a + h_n) - F(a - h_n)}{2h_n} = \frac{F(a) + h_nf(a) + o(h_n) - (F(a) - h_nf(a) - o(h_n))}{2h_n} = \frac{2h_nf(a) + o(h_n)}{2h_n} = f(a) + o(1),$$

$$\text{ce qui prouve bien que : } \lim_{n \rightarrow +\infty} \frac{F(a + h_n) - F(a - h_n)}{2h_n} = f(a) = \lim_{n \rightarrow +\infty} \mathbb{E}(f_n).$$

- b) La variable aléatoire C_n admet aussi une variance qui vaut $\mathbb{V}(C_n) = np_n(1-p_n) = \mathbb{E}(C_n) \times (1-p_n)$, donc $f_n = \frac{1}{2nh_n}C_n$ admet une variance qui vaut :

$$\mathbb{V}(f_n) = \frac{1}{4n^2h_n^2} \cdot \mathbb{V}(C_n) = \frac{\mathbb{E}(C_n)}{2nh_n} \times \frac{1-p_n}{2nh_n} = \mathbb{E}(f_n) \times \frac{1-p_n}{2nh_n}.$$

On a $\lim_{n \rightarrow +\infty} \mathbb{E}(f_n) = f(a)$, $0 \leq 1-p_n \leq 1$ et $\lim_{n \rightarrow +\infty} nh_n = +\infty$ donc $\lim_{n \rightarrow +\infty} \frac{1-p_n}{2nh_n} = 0$.

Par produit, on a donc bien $\lim_{n \rightarrow +\infty} \mathbb{V}(f_n) = 0$.

L'énoncé suppose ici que désormais, $f(a) > 0$, 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 $\lim_{n \rightarrow +\infty} nh_n^3 = 0$.

13. a) Puisque F est supposée de classe $\mathcal{C}^2$ au voisinage de a , alors elle admet un développement limité à l'ordre 2 en a qui s'écrit :

$$F(a+h) = F(a) + hF'(a) + \frac{h^2}{2}F''(a) + o(h^2) = F(a) + hf(a) + \frac{h^2}{2}f'(a) + o(h^2) \quad \text{lorsque } h \text{ est au voisinage de } 0.$$

En reprenant l'expression trouvée à la question 12. : $p_n = F(a+h_n) - F(a-h_n)$ où $\lim_{n \rightarrow +\infty} h_n = 0$, donc on peut écrire, pour n au voisinage de $+\infty$:

$$\begin{aligned} p_n &= F(a) + h_nf(a) + \frac{h_n^2}{2}f'(a) + o(h_n^2) - \left(F(a) - h_nf(a) + \frac{(-h_n)^2}{2}f'(a) + o(h_n^2) \right) \\ &= 2h_nf(a) + o(h_n^2) = \theta_n^2 + o(h_n^2), \quad \text{puisque } \theta_n = \sqrt{2h_nf(a)}. \end{aligned}$$

- b) Puisque $\lim_{n \rightarrow +\infty} h_n = 0$, un $o(h_n^2)$ est aussi un $o(h_n)$, et le développement limité $p_n = 2h_nf(a) + o(h_n)$ donne de façon équivalente : $p_n \underset{n \rightarrow +\infty}{\sim} 2h_nf(a)$.

L'équivalence étant compatible avec le produit, on a alors : $np_n \underset{n \rightarrow +\infty}{\sim} 2nh_nf(a)$.

Comme par hypothèse, $\lim_{n \rightarrow +\infty} nh_n = +\infty$ et $f(a) > 0$, alors $\lim_{n \rightarrow +\infty} 2nh_nf(a) = +\infty = \lim_{n \rightarrow +\infty} np_n$, puisque deux suites équivalentes ont la même limite.

- c) On part ici du membre de droite de la relation demandée :

$$\begin{aligned} \frac{\sigma_n}{\theta_n \sqrt{n}} D_n + \sqrt{n} \left(\frac{p_n}{\theta_n} - \theta_n \right) &= \frac{C_n - np_n}{\theta_n \sqrt{n}} + \frac{np_n - n\theta_n^2}{\theta_n \sqrt{n}} = \frac{2nh_nf_n - 2nh_nf(a)}{\theta_n \sqrt{n}} = \frac{2nh_n}{\sqrt{2nh_nf(a)}} (f_n - f(a)) \\ &= \frac{\sqrt{2nh_nf(a)}}{f(a)} (f_n - f(a)) = \frac{\theta_n \sqrt{n}}{f(a)} (f_n - f(a)) = \hat{f}_n. \end{aligned}$$

Par ailleurs :

$$\frac{\sigma_n}{\theta_n \sqrt{n}} = \sqrt{\frac{np_n(1-p_n)}{2nh_nf(a)}} = \sqrt{\frac{p_n}{2h_nf(a)}(1-p_n)} \xrightarrow{n \rightarrow +\infty} 1,$$

puisque $p_n \underset{n \rightarrow +\infty}{\sim} 2h_nf(a)$ et $\lim_{n \rightarrow +\infty} 1-p_n = 1$ vu que $\lim_{n \rightarrow +\infty} p_n = \lim_{n \rightarrow +\infty} 2h_nf(a) = 0$,

$$\text{et } \sqrt{n} \left(\frac{p_n}{\theta_n} - \theta_n \right) = \sqrt{n} \left(\frac{\theta_n^2 + o(h_n^2)}{\theta_n} - \theta_n \right) = \frac{\sqrt{n}}{\theta_n} \cdot o(h_n^2) = o(h_n^2) \xrightarrow{n \rightarrow +\infty} 0 \text{ puisque } \lim_{n \rightarrow +\infty} h_n = 0.$$

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

$$\lim_{n \rightarrow +\infty} \mathbf{P}(x \leq \hat{f}_n \leq y) = \Phi(y) - \Phi(x).$$

14. 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 $\mathcal{N}(0, 1)$, c'est-à-dire l'unique réel qui vérifie : $\Phi(\eta_\alpha) = 1 - \frac{\alpha}{2}$.

a) Pour n assez grand, on peut écrire :

$$\begin{aligned} & \mathbf{P}\left((f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right)f(a) + f_n^2 \leq 0\right) = \mathbf{P}\left((f(a))^2 - 2f_nf(a) + f_n^2 \leq \frac{\eta_\alpha}{2nh_n}f(a)\right) \\ &= \mathbf{P}\left((f_n - f(a))^2 \leq \frac{\eta_\alpha}{2nh_n}f(a)\right) = \mathbf{P}\left(-\frac{\sqrt{\eta_\alpha f(a)}}{\sqrt{2nh_n}} \leq f_n - f(a) \leq \frac{\sqrt{\eta_\alpha f(a)}}{\sqrt{2nh_n}}\right) \\ &= \mathbf{P}\left(-\frac{t_\alpha f(a)}{\theta_n} \leq f_n - f(a) \leq \frac{t_\alpha f(a)}{\theta_n \sqrt{n}}\right) = \mathbf{P}\left(t_\alpha \leq \frac{\theta_n \sqrt{n}}{f(a)}(f_n - f(a)) \leq t_\alpha\right) \\ &= \mathbf{P}(t_\alpha \leq \hat{f}_n \leq t_\alpha), \end{aligned}$$

Donc :

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbf{P}\left((f(a))^2 - \left(2f_n + \frac{\eta_\alpha}{2nh_n}\right)f(a) + f_n^2 \leq 0\right) &= \lim_{n \rightarrow +\infty} \mathbf{P}(t_\alpha \leq \hat{f}_n \leq t_\alpha) = \Phi(t_\alpha) - \Phi(-t_\alpha) \\ &= 2\Phi(t_\alpha) - 1 = 2 - \alpha - 1 = 1 - \alpha. \end{aligned}$$

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

La probabilité initiale de la question précédente se réécrit aussi :

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

Le résultat de la question précédente s'applique donc quant à la limite de cette probabilité lorsque n tend vers $+\infty$:

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

ce qui signifie que $\left[f_n + \frac{\eta_\alpha}{4nh_n} - \Delta_n; f_n + \frac{\eta_\alpha}{4nh_n} + \Delta_n\right]$ est un intervalle de confiance asymptotique pour $f(a)$ au niveau de confiance $1 - \alpha$.

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

16. a) Remarquons d'emblée qu'avec les notations et hypothèses introduites au début de cette partie, pour tout $k \in \llbracket 1; n \rrbracket$, X_k et $Y_k = S_n - X_k = \sum_{\substack{1 \leq j \leq n \\ j \neq k}} X_j$ sont indépendantes d'après le lemme des

coalitions, puisque les $(X_k)_{1 \leq k \leq n}$ sont mutuellement indépendantes.

Il en est par conséquent de même, d'après le même lemme, de X_k^2 et $f'(Y_k)$. Toutes les espérances que l'on va écrire existent d'après l'énoncé, donc :

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

où l'énoncé suppose que $\sum_{k=1}^n v_k = 1$ et où $\mathbb{E}(X_k^2 f'(Y_k)) = \mathbb{E}(X_k^2) \times \mathbb{E}(f'(Y_k)) = v_k \mathbb{E}(f'(Y_k))$ pour tout $k \in \llbracket 1; n \rrbracket$, vu que X_k^2 et $f'(Y_k)$ sont indépendantes.

Les deux dernières sommes s'annulent donc, et on a bien :

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

b) Toujours par indépendance de X_k avec Y_k qui entraîne celle de X_k et $f(Y_k)$ pour tout $k \in \llbracket 1; n \rrbracket$, et par linéarité de l'espérance, on peut écrire :

$$\begin{aligned} \sum_{k=1}^n \mathbb{E}(X_k [f(S_n) - f(Y_k)]) &= \sum_{k=1}^n \mathbb{E}(X_k f(S_n) - X_k f(Y_k)) \\ &= \sum_{k=1}^n \mathbb{E}(X_k f(S_n)) - \sum_{k=1}^n \mathbb{E}(X_k) \times \mathbb{E}(f(Y_k)) \\ &\text{où par hypothèse, } X_k \text{ est centrée : } \mathbb{E}(X_k) = 0 \\ &= \sum_{k=1}^n \mathbb{E}(X_k f(S_n)) = \mathbb{E}\left(f(S_n) \sum_{k=1}^n X_k\right) = \mathbb{E}(S_n f(S_n)). \end{aligned}$$

c) En faisant la différence membre à membre des deux égalités précédentes, on en déduit :

$$\begin{aligned} \mathbb{E}(f'(S_n) - S_n f(S_n)) &= \mathbb{E}(f'(S_n)) - \mathbb{E}(S_n f(S_n)) \\ &= \sum_{k=1}^n v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^n \mathbb{E}(X_k^2 f'(Y_k)) - \sum_{k=1}^n \mathbb{E}(X_k [f(S_n) - f(Y_k)]) \\ &= \sum_{k=1}^n v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^n \mathbb{E}(X_k^2 f'(Y_k) - X_k [f(S_n) - f(Y_k)]) \\ &= \sum_{k=1}^n v_k \mathbb{E}(f'(S_n) - f'(Y_k)) + \sum_{k=1}^n \mathbb{E}(X_k [X_k f'(Y_k) - (f(S_n) - f(Y_k))]) \end{aligned}$$

17. Soient a et b deux réels.

a) On part du membre de droite et on calcule l'intégrale :

$$\int_0^1 b(f'(a) - f'(a+tb))dt = (1-0) \times b f'(a) - \int_0^1 b f'(a+tb)dt = b f'(a) - \left[f(a+tb) \right]_0^1 = b f'(a) - (f(a+b) - f(a)).$$

b) On a supposé au début de cette partie que pour tout réel x , $|f''(x)| \leq 2$: l'Inégalité des Accroissements Finis écrite pour f' , assure alors que :

$$\forall t \in [0; 1], \quad |f'(a) - f'(a+tb)| \leq |a - a - tb| = t|b| \implies |f'(a) - f'(a+tb)| \leq |b| \quad \text{puisque } 0 \leq t \leq 1$$

Mais alors :

$$\begin{aligned} |b f'(a) - (f(a+b) - f(a))| &= \left| \int_0^1 b(f'(a) - f'(a+tb))dt \right| \\ &\leq \int_0^1 |b| \times |f'(a) - f'(a+tb)|dt \quad \text{d'après l'inégalité triangulaire} \\ &\leq \int_0^1 |b| \times |b|dt = b^2 \quad \text{par croissance de l'intégrale } (0 < 1) \end{aligned}$$

c) L'inégalité triangulaire et le résultat de la question 11.c) donnent :

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

où :

- $|v_k| = v_k$ car $v_k = \mathbb{E}(X_k^2)$ est positive, comme espérance d'une variable aléatoire positive,
- $|\mathbb{E}(f'(S_n) - f'(Y_k))| \leq \mathbb{E}(|f'(S_n) - f'(Y_k)|)$, où puisque $|f''(x)| \leq 2$ pour tout réel x :

$$\forall (x, y) \in \mathbb{R}^2, \quad |f'(S_n) - f'(Y_k)| \leq 2|S_n - Y_k| = 2|X_k| \implies \mathbb{E}(|f'(S_n) - f'(Y_k)|) \leq 2\mathbb{E}(|X_k|)$$

par croissance de l'espérance. Ainsi :

$$\forall k \in \llbracket 1; n \rrbracket, \quad |\mathbb{E}(f'(S_n) - f'(Y_k))| \leq 2\mathbb{E}(|X_k|) \implies \sum_{k=1}^n |\mathbb{E}(f'(S_n) - f'(Y_k))| \leq 2 \sum_{k=1}^n \mathbb{E}(|X_k|).$$

- L'inégalité démontrée en 17.b) étant vraie pour tout réel, elle entraîne l'inégalité entre variables aléatoires :

$$\forall k \in \llbracket 1; n \rrbracket, \quad |X_k f'(Y_k) - (f(X_k + Y_k) - f(Y_k))| \leq |X_k|^2 \implies \left| X_k (X_k f'(Y_k) - (f(X_k + Y_k) - f(Y_k))) \right| \leq |X_k|^3,$$

Donc sachant que $X_k + Y_k = n$:

$$\left| \mathbb{E} \left(X_k [X_k f'(Y_k) - (f(S_n) - f(Y_k))] \right) \right| \leq \mathbb{E} \left(\left| X_k [X_k f'(Y_k) - (f(S_n) - f(Y_k))] \right| \right) \leq \mathbb{E}(|X_k|^3),$$

et par passage à la somme dans cette inégalité, on obtient bien :

$$\sum_{k=1}^n |v_k| \cdot |\mathbb{E}(f'(S_n) - f'(Y_k))| + \sum_{k=1}^n \left| \mathbb{E} \left(X_k [X_k f'(Y_k) - (f(S_n) - f(Y_k))] \right) \right| \leq 2 \sum_{k=1}^n v_k \mathbb{E}(|x_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3),$$

donc par transitivité de l'inégalité :

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

On termine cette question par une synthèse de plusieurs résultats précédents :

D'après l'ensemble de la partie 1, pour toute fonction h de W , la fonction f_h définie au début de cette partie est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et vérifie :

$$\forall x \in \mathbb{R}, \quad |f_h(x)| \leq 1 \text{ et } |f_h''(x)| \leq 2.$$

Le résultat qu'on vient d'établir s'applique donc avec $f = f_h$, qui assure que :

$$\forall h \in W, \quad |\mathbb{E}(f_h'(S_n) - S_n f_h(S_n))| \leq 2 \sum_{k=1}^n v_k \mathbb{E}(|x_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3),$$

ce qui s'écrit aussi, d'après la question 3.c) :

$$\forall h \in W, \quad |\mathbb{E}(h(S_n)) - \mathbb{E}(h(N))| \leq 2 \sum_{k=1}^n v_k \mathbb{E}(|x_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3).$$

On a donc trouvé un réel $M_{S_n} = 2 \sum_{k=1}^n v_k \mathbb{E}(|x_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3)$ qui vérifie les conditions établies au début de la question 9. : le résultat de la partie 2 s'ensuit pour ce réel, qui assure que :

$$\forall x \in \mathbb{R}, \quad d_{S_n}(x) \leq \sqrt{3M_{S_n}}, \text{ soit : } d_{S_n}(x) \leq \sqrt{3 \left(2 \sum_{k=1}^n v_k \mathbb{E}(|x_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3) \right)} \quad (R_3)$$

18. a) Avec les notations et conventions introduites au début de cette question 18 :

- Les variables aléatoires $X_k = \frac{Z_k - \mathbb{E}(Z_k)}{\sigma \sqrt{n}}$ sont bien centrées :

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

- On a aussi :

$$v_k = \mathbb{E}(X_k^2) = \frac{1}{n\sigma^2} \mathbb{E}(|Z_k - \mathbb{E}(Z_k)|^2) = \frac{s_2}{ns_2} = \frac{1}{n}, \text{ donc } \sum_{k=1}^n \mathbb{E}(X_k^2) = \sum_{k=1}^n v_k = n \times \frac{1}{n} = 1.$$

Ces deux conditions étant vérifiées, les résultats des questions 16. et 17. sont ici valables, et l'inégalité (R_3) s'applique.

Pour l'écrire explicitement avec les variables aléatoires actuelles, il reste à calculer :

$$2 \sum_{k=1}^n v_k \mathbb{E}(|X_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3) = 2 \sum_{k=1}^n \frac{1}{n} \times \frac{s_1}{\sigma \sqrt{n}} + \sum_{k=1}^n \frac{s_3}{\sigma^3 \cdot (\sqrt{n})^3} = \frac{2s_1}{\sigma \sqrt{n}} + \frac{ns_3}{\sigma^3 \cdot n \sqrt{n}} = \frac{2\sigma^2 s_1 + s_3}{\sigma^3 \sqrt{n}},$$

donc (R_3) s'écrit ici :

$$\forall x \in \mathbb{R}, \quad d_{S_n}(x) \leq \sqrt{3 \frac{2\sigma^2 s_1 + s_3}{\sigma^3 \sqrt{n}}}.$$

b) Il est alors clair que $\delta_n = \sqrt{3 \frac{2\sigma^2 s_1 + s_3}{\sigma^3 \sqrt{n}}}$ est bien le terme général d'une suite qui tend vers 0 :

d'après la définition donnée par l'énoncé, on peut donc conclure que la suite $(S_n)_{n \geq 1}$ converge uniformément en loi vers N , donc converge en loi vers N .

Ce dernier résultat (convergence en loi vers N) aurait pu directement être obtenu par le théorème de la limite centrée, vu que les $(Z_k)_{k \geq 1}$ sont mutuellement indépendantes et de même loi, admettent

une espérance et une variance non nulle σ^2 , et vu que $S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n (Z_k - \mathbb{E}(Z_k)) = \sqrt{n} \left(\frac{\overline{Z_n} - m}{\sigma} \right)$

avec $\overline{Z_n} = \frac{1}{n} \sum_{k=1}^n Z_k$ et m l'espérance commune aux Z_k .

19. a) Avec les notations introduites dans cette question, et d'après le théorème de transfert :

$$\mathbb{E}(|X_k|) = \left| \frac{0 - p_n}{\sigma_n} \right| \mathbf{P}(Z_k = 0) + \left| \frac{1 - p_n}{\sigma_n} \right| \mathbf{P}(Z_k = 1) = \frac{p_n(1 - p_n)}{\sigma_n} + \frac{(1 - p_n)p_n}{\sigma_n} = \frac{2p_n(1 - p_n)}{\sigma_n} = \frac{2\sigma_n^2}{n\sigma_n} = \frac{2\sigma_n}{n},$$

et de même :

$$\begin{aligned} \mathbb{E}(|X_k|^3) &= \left| \frac{0 - p_n}{\sigma_n} \right|^3 \mathbf{P}(Z_k = 0) + \left| \frac{1 - p_n}{\sigma_n} \right|^3 \mathbf{P}(Z_k = 1) = \frac{p_n(1 - p_n)}{\sigma_n^3} \cdot (1 - p_n)^2 + \frac{p_n(1 - p_n)}{\sigma_n^3} \cdot p_n^2 \\ &= \frac{p_n(1 - p_n)}{\sigma_n} \cdot (p_n^2 + (1 - p_n)^2) = \frac{\sigma_n^2}{n\sigma_n^3} \cdot (p_n^2 + (1 - p_n)^2) \leq \frac{2}{n\sigma_n} \end{aligned}$$

puisque $0 \leq p_n \leq 1$ et $0 \leq 1 - p_n \leq 1$, donc $p_n^2 + (1 - p_n)^2 \leq 2$.

b) Les variables X_k sont à nouveau centrées puisque $\mathbb{E}(Z_k) = p_n$ vu que Z_k est une variable de Bernoulli de paramètre p_n ,

$$\text{et } v_k = \mathbb{E}(X_k^2) = \mathbb{V}(X_k) = \frac{1}{\sigma_n^2} \mathbb{V}(Z_k) = \frac{p_n(1 - p_n)}{\sigma_n} = \frac{\sigma_n^2}{n\sigma_n^2} = \frac{1}{n}, \text{ donc } \sum_{k=1}^n v_k = 1.$$

L'inégalité (R_3) s'applique donc, et pour pouvoir l'écrire il faut à nouveau considérer :

$$2 \sum_{k=1}^n v_k \mathbb{E}(|X_k|) + \sum_{k=1}^n \mathbb{E}(|X_k|^3) \leq \frac{2}{n} \sum_{k=1}^n \frac{2\sigma_n}{n} + \sum_{k=1}^n \frac{2}{n\sigma_n} = \frac{4\sigma_n}{n} + \frac{2}{\sigma_n},$$

d'où pour tout réel x :

$$d_{S_n}(x) \leq \sqrt{3 \left(\frac{4\sigma_n}{n} + \frac{2}{\sigma_n} \right)} = 2\sqrt{3 \left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n} \right)}.$$

c) Avec les notations de l'énoncé :

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

Or avec $\delta'_n = 2\sqrt{3 \left(\frac{\sigma_n}{n} + \frac{1}{2\sigma_n} \right)}$:

- $\lim_{n \rightarrow +\infty} \frac{\sigma_n}{n} = \sqrt{\frac{p_n(1 - p_n)}{n}} = 0$ si $\lim_{n \rightarrow +\infty} p_n = 0$,
- $\lim_{n \rightarrow +\infty} \frac{1}{2\sigma_n} = \lim_{n \rightarrow +\infty} \frac{1}{2\sqrt{np_n(1 - p_n)}} = 0$ si $\lim_{n \rightarrow +\infty} np_n = 0$ et $\lim_{n \rightarrow +\infty} p_n = 0$,
car alors $\lim_{n \rightarrow +\infty} (1 - p_n) = 1$.

Si donc $\lim_{n \rightarrow +\infty} p_n = 0$ et $\lim_{n \rightarrow +\infty} np_n = +\infty$, alors $\lim_{n \rightarrow +\infty} \delta'_n = 0$ et on peut conclure que

la suite $\left(\frac{T_n - np_n}{\sqrt{np_n(1 - p_n)}} \right)_{n \geq 1}$ converge uniformément en loi vers N .

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

a) Soit X une variable aléatoire et (α, β) un couple de réels avec $\alpha > 0$.

Pour tout x réel :

$$|F_{\alpha X + \beta}(x) - F_{\alpha N + \beta}(x)| = |\mathbf{P}(\alpha X + \beta \leq x) - \mathbf{P}(\alpha N + \beta \leq x)| \stackrel{\alpha > 0}{=} \left| \mathbf{P}\left(X \leq \frac{x - \beta}{\alpha}\right) - \mathbf{P}\left(N \leq \frac{x - \beta}{\alpha}\right) \right| = d_X\left(\frac{x - \beta}{\alpha}\right).$$

- b) Puisque la suite $(V_n)_{n \geq 1}$ converge uniformément en loi vers N , alors il existe une suite $(\delta_n)_{n \geq 1}$ telle que :

$$\lim_{n \rightarrow +\infty} \delta_n = 0 \text{ et } \forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, d_{V_n}(x) \leq \delta_n.$$

On a alors, d'après la question précédente et pour tout $n \in \mathbb{N}^*$ et tout réel x :

$$|\mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(a_n N + b_n \leq x)| = d_{V_n}\left(\frac{x - b_n}{a_n}\right) \leq \delta_n,$$

donc par encadrement : $\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} |\mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(a_n N + b_n \leq x)| = 0$, ce qui donne bien :

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(a_n N + b_n \leq x) = 0.$$

- c) Puisque $\lim_{n \rightarrow +\infty} a_n = a > 0$ alors pour tout réel x , $\lim_{n \rightarrow +\infty} \frac{x - b_n}{a_n} = \frac{x - b}{a}$ et puisque Φ est continue sur $\mathbb{R}$, alors :

$$\lim_{n \rightarrow +\infty} \mathbf{P}(a_n N + b_n \leq x) = \Phi\left(\frac{x - b_n}{a_n}\right) = \Phi\left(\frac{x - b}{a}\right) = \mathbf{P}(aN + b \leq x).$$

On écrit alors, pour tout réel x et tout entier $n \in \mathbb{N}^*$:

$$\mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(aN + b \leq x) = (\mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(a_n N + b_n \leq x)) + (\mathbf{P}(a_n N + b_n \leq x) - \mathbf{P}(aN + b \leq x)).$$

La combinaison des deux limites précédentes assure alors que :

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbf{P}(a_n V_n + b_n \leq x) - \mathbf{P}(aN + b \leq x),$$

ce qui signifie bien par définition que la suite de variables aléatoires $(a_n V_n + b_n)_{n \geq 1}$ converge en loi vers $aN + b$.

Comme N suit la loi normale centrée réduite, le théorème de stabilité pour la loi normale assure que $aN + b$ suit la loi normale de paramètres (b, a^2) .

21. On reprend les notations de la partie 3.

- a) La variable aléatoire C_n suit, d'après la partie 3, la loi binomiale de paramètres n et $p_n \sim 2h_n f(a)$ qui vérifie : $\lim_{n \rightarrow +\infty} p_n = 0$ et $\lim_{n \rightarrow +\infty} np_n = +\infty$ (voir la question 14.b)), donc d'après 19.c), la suite de variables aléatoires $(D_n)_{n \geq 1} = \left(\frac{C_n - np_n}{\sigma_n}\right)_{n \geq 1}$ avec $\sigma_n = \sqrt{np_n(1 - p_n)}$, converge uniformément en loi vers N .

- b) Le résultat de la question 20.c) s'applique alors, avec $a_n = \frac{\sigma_n}{\theta_n \sqrt{n}}$ qui tend vers $a = 1 > 0$ et $b_n = \sqrt{n} \left(\frac{p_n}{\theta_n} - \theta_n\right)$ qui tend vers $b = 0$:

La suite de variables aléatoires $(\hat{f}_n)_{n \geq 1} = (a_n \cdot D_n + b_n)_{n \geq 1}$ converge uniformément en loi vers $aN + b = N$.

★★★ FIN DU SUJET ★★★