



PRÉPASIMPLE

prepasimple.fr

HEC / ESSEC

Mathématiques — Mathématiques appliquées

2023

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Corrigé

Partie 1 - Transformation de Stein

1. a) La densité φ est positive. Comme $0 \leq x \leq t$, on a $0 \leq x\varphi(t) \leq t\varphi(t)$. Pour $A \geq x$, l'intégration sur $[x, A]$ donne :

$$0 \leq x \int_x^A \varphi(t) dt \leq \int_x^A t\varphi(t) dt = \varphi(x) - \varphi(A).$$

Puisque $\varphi(A) \rightarrow 0$ et $\int_x^A \varphi(t) dt \rightarrow 1 - \Phi(x)$, on obtient $0 \leq x(1 - \Phi(x)) \leq \varphi(x)$, y compris pour $x = 0$.

1. b) Méthode 1 — Intégration sur la demi-droite négative

Si $t \leq x \leq 0$, alors $(-x)\varphi(t) \leq (-t)\varphi(t)$. Pour $B < x$:

$$0 \leq (-x) \int_B^x \varphi(t) dt \leq \int_B^x (-t)\varphi(t) dt = \varphi(x) - \varphi(B).$$

En faisant tendre B vers $-\infty$, on trouve $0 \leq -x\Phi(x) \leq \varphi(x)$, soit $-\varphi(x) \leq x\Phi(x) \leq 0$.

Méthode 2 — Symétrie de la loi normale

Appliquons 1. a) à $-x \geq 0$: $0 \leq (-x)(1 - \Phi(-x)) \leq \varphi(-x)$. La symétrie de la loi normale donne $\varphi(-x) = \varphi(x)$ et $1 - \Phi(-x) = \Phi(x)$. Ainsi $0 \leq -x\Phi(x) \leq \varphi(x)$, soit $-\varphi(x) \leq x\Phi(x) \leq 0$. Cette approche réutilise la borne positive ; la première établit directement la borne négative.

1. c) Les encadrements précédents impliquent $B\Phi(B) \rightarrow 0$ lorsque $B \rightarrow -\infty$ et $A(1 - \Phi(A)) \rightarrow 0$ lorsque $A \rightarrow +\infty$. L'intégration par parties sur $[B, x]$, avec $u = \Phi$ et $v' = 1$, donne :

$$\int_B^x \Phi(t) dt = x\Phi(x) - B\Phi(B) - \int_B^x t\varphi(t) dt = x\Phi(x) + \varphi(x) - B\Phi(B) - \varphi(B).$$

Le passage à la limite prouve la convergence et la première égalité de (R_1) . Sur $[x, A]$, de même :

$$\int_x^A (1 - \Phi(t)) dt = A(1 - \Phi(A)) - x(1 - \Phi(x)) + \varphi(x) - \varphi(A).$$

Le passage à la limite prouve la deuxième égalité de (R_1) .

2. a) L'inégalité des accroissements finis, avec $|h'| \leq 1$, donne $|h(x) - h(y)| \leq |x - y|$.
Pour $y = 0$, l'inégalité triangulaire fournit $|h(x)| \leq |x| + |h(0)|$.

2. b) Par 1. c), $\int_{-\infty}^x \Phi(t) dt$ converge. Comme $|h'(t)\Phi(t)| \leq \Phi(t)$, l'intégrale demandée converge absolument. De plus, $|h(t)\varphi(t)| \leq (|t| + |h(0)|)\varphi(t)$ est intégrable : N possède une espérance, et donc $E(|N|) < +\infty$.

Sur $[B, x]$, l'intégration par parties donne :

$$\int_B^x h'(t)\Phi(t) dt = h(x)\Phi(x) - h(B)\Phi(B) - \int_B^x h(t)\varphi(t) dt.$$

Pour $B \leq 0$, $|h(B)\Phi(B)| \leq |B|\Phi(B) + |h(0)|\Phi(B) \rightarrow 0$ d'après 1. b). Le passage à la limite donne l'égalité annoncée. L'égalité relative à la seconde intégrale est admise dans le sujet.

2. c) En soustrayant les égalités de 2. b), les termes contenant $h(x)$ ont pour somme $-h(x)\Phi(x) - h(x)(1 - \Phi(x)) = -h(x)$. Les intégrales restantes se réunissent en $\int_{-\infty}^{+\infty} h(t)\varphi(t) dt = c_h$ par transfert, avec convergence absolue justifiée en 2. b). On obtient donc $c_h - h(x)$.

3. a) Puisque $\Phi' = \varphi$ et $\varphi' = -x\varphi$, la dérivation du quotient donne :

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

En dérivant encore, $\theta''(x) = \theta(x) + x\theta'(x) = x + (1 + x^2)\theta(x)$. Enfin, la symétrie normale donne :

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

3. b) Posons $A(x) = \int_{-\infty}^x h'(t)\Phi(t) dt$ et $B(x) = \int_x^{+\infty} h'(t)(1 - \Phi(t)) dt$. Ces intégrales convergent par comparaison avec (R_1) . À l'aide de Chasles et du théorème fondamental sur un segment, A et B sont de classe C^1 , avec $A' = h'\Phi$ et $B' = -h'(1 - \Phi)$. Ainsi f_h est de classe C^1 et :

$$f'_h(x) = -\theta'(-x)A(x) + \theta'(x)B(x) + h'(x)[\theta(-x)\Phi(x) - \theta(x)(1 - \Phi(x))].$$

Le dernier terme est nul. Or $-\theta'(-x) = -1 + x\theta(-x)$ et $\theta'(x) = 1 + x\theta(x)$. D'où :

$$f'_h(x) = xf_h(x) - A(x) + B(x) = xf_h(x) + c_h - h(x).$$

Le second membre est de classe C^1 puisque f_h et h le sont. Donc f'_h est de classe C^1 et f_h est de classe C^2 .

3. c) L'égalité ponctuelle précédente donne $f'_h(X) - Xf_h(X) = c_h - h(X)$. Le membre droit est intégrable par l'admission du préambule. La différence du membre gauche possède donc bien une espérance, sans avoir à séparer ses termes. Sa moyenne est $c_h - E(h(X))$. En prenant les valeurs absolues et en utilisant $c_h = E(h(N))$, on obtient le résultat.

4. a) Avec $I(x) = x\Phi(x) + \varphi(x)$ et $J(x) = -x(1 - \Phi(x)) + \varphi(x)$ donnés par (R_1) , le membre gauche vaut :

$$\frac{1 - \Phi(x)}{\varphi(x)}[x\Phi(x) + \varphi(x)] + \frac{\Phi(x)}{\varphi(x)}[-x(1 - \Phi(x)) + \varphi(x)] = 1.$$

Les deux termes proportionnels à x s'annulent et les autres donnent $1 - \Phi(x) + \Phi(x) = 1$.

4. b) Les fonctions θ , Φ et $1 - \Phi$ sont positives. Comme $|h'| \leq 1$, $|A(x)| \leq I(x)$ et $|B(x)| \leq J(x)$. L'inégalité triangulaire et 4. a) donnent $|f_h(x)| \leq \theta(-x)I(x) + \theta(x)J(x) = 1$.

5. a) En utilisant 3. a) et 4. a), le membre gauche est :

$$[-x + (1 + x^2)\theta(-x)]I(x) + [x + (1 + x^2)\theta(x)]J(x) = -xI(x) + xJ(x) + 1 + x^2.$$

Or $J(x) - I(x) = -x$. Le résultat est donc $-x^2 + 1 + x^2 = 1$.

5. b) On dérive l'expression $f'_h = -\theta'(-x)A + \theta'(x)B$ obtenue en 3. b) :

$$f''_h(x) = \theta''(-x)A(x) + \theta''(x)B(x) - h'(x)[\theta'(-x)\Phi(x) + \theta'(x)(1 - \Phi(x))].$$

Par 3. a), l'expression entre crochets est :

$$[1 - x\theta(-x)]\Phi(x) + [1 + x\theta(x)](1 - \Phi(x)) = 1.$$

Les termes proportionnels à x s'annulent. On obtient l'égalité demandée.

5. c) Posons $q(x) = \Phi(x) + \frac{x}{1+x^2}\varphi(x)$. Une dérivation explicite fournit :

$$q'(x) = \varphi(x) \left(1 + \frac{1-x^2}{(1+x^2)^2} - \frac{x^2}{1+x^2} \right) = \frac{2\varphi(x)}{(1+x^2)^2} > 0.$$

Ainsi q est strictement croissante. Comme $\frac{x}{1+x^2}\varphi(x) \rightarrow 0$ aux deux infinis, ses limites sont 0 en $-\infty$ et 1 en $+\infty$. Donc $0 < q(x) < 1$. D'après 3. a),

$$\theta''(x) = \frac{1+x^2}{\varphi(x)}q(x) > 0.$$

Dans 5. b), les coefficients $\theta''(-x)$ et $\theta''(x)$ sont donc positifs. Avec $|h'| \leq 1$, puis 5. a), on obtient :

$$|f_h''(x)| \leq 1 + \theta''(-x)I(x) + \theta''(x)J(x) = 2.$$

Partie 2 - Majoration uniforme de la distance de Kolmogorov

6. La variable $h_x(X)$ est l'indicatrice de $[X \leq x]$. Elle suit une loi de Bernoulli de paramètre $F_X(x)$, les paramètres extrêmes 0 et 1 correspondant à des lois certaines. Elle est bornée, donc possède une espérance, égale à $P(X \leq x) = F_X(x)$.

7. a) Une traduction directe de la définition est :

Python

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

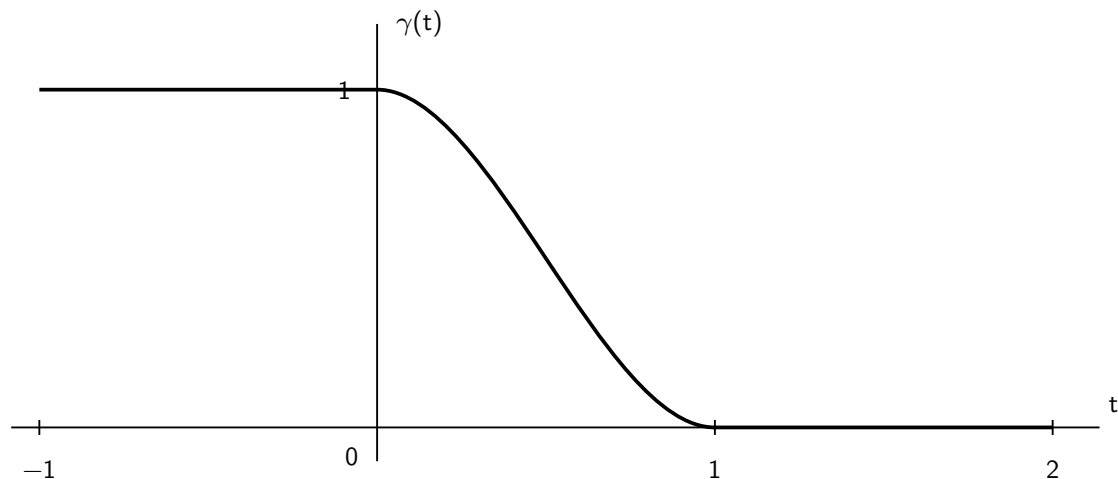
7. b) On évalue la fonction sur une grille, puis on relie les valeurs calculées :

Python

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

T = np.linspace(-1, 2, 301)
Y = [gamma(t) for t in T]
plt.plot(T, Y)
plt.show()
```

La représentation est constante égale à 1 jusqu'à 0, décroît de 1 à 0 entre 0 et 1, puis reste nulle. La portion cubique ci-dessous est tracée exactement à partir de sa formule.



8. a) Sur chacun des intervalles $] -\infty, 0[$, $]0, 1[$ et $]1, +\infty[$, la fonction est polynomiale ou constante et donc de classe C^1 . En 0, les limites valent $1 = \gamma(0)$; en 1, elles valent $1 - 3 + 2 = 0 = \gamma(1)$. Cela prouve sa continuité sur $\mathbb{R}$.

8. b) Sur $]0, 1[$, $\gamma'(t) = -6t + 6t^2 = 6t(t - 1) < 0$. La fonction décroît donc de $\gamma(0) = 1$ à $\gamma(1) = 0$. À l'extérieur, ses valeurs sont 1 ou 0; donc $0 \leq \gamma(t) \leq 1$ sur tout $\mathbb{R}$.

8. c) Pour $u > 0$, le taux d'accroissement en 1 est nul. Pour $u < 0$ assez proche de zéro, un développement polynomial exact donne $\gamma(1 + u) = 3u^2 + 2u^3$, donc :

$$\frac{\gamma(1 + u) - \gamma(1)}{u} = 3u + 2u^2 \rightarrow 0.$$

Les deux limites sont égales : γ est dérivable en 1 et $\gamma'(1) = 0$. Le résultat analogue en 0 est admis dans le sujet; on peut également le vérifier par le quotient $(-3u^2 + 2u^3)/u$ à droite et le quotient nul à gauche.

8. d) La dérivée vaut 0 hors $[0, 1]$ et $6t(t - 1)$ sur $[0, 1]$. Cette expression se raccorde continûment à zéro aux bornes, donc $\gamma \in C^1(\mathbb{R})$. Sur $[0, 1]$:

$$|\gamma'(t)| = 6t(1 - t) = \frac{3}{2} - 6\left(t - \frac{1}{2}\right)^2 \leq \frac{3}{2}.$$

Hors de ce segment, la dérivée est nulle.

9. a) Si $y \leq x$, alors $(y - x)/t \leq 0$, donc $k_x(y) = 1 = h_x(y)$. Si $y > x$, alors $h_x(y) = 0 \leq k_x(y)$ par 8. b). On a donc partout $h_x \leq k_x$.
9. b) La croissance de l'espérance donne $E(h_x(X)) \leq E(k_x(X))$. On soustrait $E(h_x(N))$, puis on ajoute et soustrait $E(k_x(N))$. On obtient exactement l'inégalité annoncée.
9. c) La différence $k_x(u) - h_x(u)$ est nulle si $u \leq x$ ou $u \geq x + t$, sauf éventuellement en des points isolés qui ne changent pas l'intégrale. Pour $x < u < x + t$, elle vaut $k_x(u)$. Le transfert, applicable à ces fonctions bornées et continues par morceaux, donne :

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

9. d) Par 8. d), g est de classe C^1 et :

$$g'(u) = \frac{2}{3} \gamma' \left(\frac{u-x}{t} \right), \quad |g'(u)| \leq 1.$$

Donc $g \in W$ et, par linéarité, $|E(k_x(X)) - E(k_x(N))| \leq \frac{3}{2t} M_X$. Par ailleurs, $0 \leq k_x \leq 1$ et $\varphi \leq 1/\sqrt{2\pi}$, d'où :

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

La dernière inégalité utilise $\sqrt{2\pi} > 2$. En combinant avec 9. b), on obtient les deux majorations demandées.

10. On a $M_X \geq 0$, car l'hypothèse s'applique notamment à la fonction nulle. Si $M_X > 0$, la dérivée de $r(t) = \frac{3M_X}{2t} + \frac{t}{2}$ vaut $r'(t) = \frac{1}{2} - \frac{3M_X}{2t^2}$. Elle est négative pour $t < \sqrt{3M_X}$ et positive pour $t > \sqrt{3M_X}$. Le minimum est donc atteint en $\sqrt{3M_X}$ et vaut $\sqrt{3M_X}$.

Les deux sens d'inégalité de 9. d) donnent $|E(h_x(X)) - E(h_x(N))| \leq r(t)$ pour tout $t > 0$, donc la borne annoncée. Si $M_X = 0$, cette même différence est majorée par $t/2$ pour tout $t > 0$; elle est donc nulle. Enfin, par 6, elle vaut $|F_X(x) - \Phi(x)| = d_X(x)$. Cela établit (R_2) dans tous les cas.

Partie 3 - Estimation d'une densité

11. a) Question adaptée au programme actuel : les commandes pandas sont documentées dans le sujet, car leur syntaxe n'est pas exigible de mémoire. Les instructions sont :

Python

```
import pandas as pd

table = pd.read_csv("stats.csv")

echantillon = table["salaire"]
```

11. b) La condition doit traduire exactement l'intervalle ouvert à gauche et fermé à droite. L'indice Python i correspond à la donnée mathématique de rang $i + 1$. L'accès par position évite de dépendre des étiquettes de la série :

Python

```
import numpy as np

a = float(input("a="))

n = echantillon.count()

h = 1 / np.sqrt(n)

C = 0

for i in range(n):

    if echantillon.iloc[i] > a - h and echantillon.iloc[i] <= a + h:

        C += 1

print(C / (2 * n * h))
```

Le comptage est initialisé à zéro et traite chacune des n observations une fois. L'hypothèse de données numériques sans valeurs manquantes garantit que `count()` est la taille de l'échantillon. Question adaptée au programme actuel : cette hypothèse et l'accès par position sont explicités pour rendre le code non ambigu. Le programme calcule une estimation observée ; f_n reste, dans la théorie, une variable aléatoire.

12. Posons $I_{n,i} = 1$ si $a - h_n < X_i \leq a + h_n$, et $I_{n,i} = 0$ sinon. Les $I_{n,i}$ sont indépendantes, de même loi de Bernoulli, de paramètre :

$$p_n = P(a - h_n < X \leq a + h_n) = F(a + h_n) - F(a - h_n).$$

Comme

$$C_n = \sum_{i=1}^n I_{n,i},$$

C_n suit la loi binomiale (n, p_n) . Cette variable bornée possède une espérance np_n . Par linéarité, f_n possède une espérance égale à $p_n/(2h_n)$, soit la formule du sujet.

13. a) Méthode 1 — Dérivabilité de la fonction de répartition

Comme f est continue en a , F est dérivable en a et $F'(a) = f(a)$. On écrit :

$$E(f_n) = \frac{1}{2} \left(\frac{F(a + h_n) - F(a)}{h_n} + \frac{F(a) - F(a - h_n)}{h_n} \right).$$

Les deux quotients tendent vers $f(a)$, donc $E(f_n) \rightarrow f(a)$.

Méthode 2 — Encadrement de la moyenne intégrale

Par continuité de f en a , pour tout $\varepsilon > 0$ et tout n assez grand, $|f(u) - f(a)| \leq \varepsilon$ pour

$u \in [a - h_n, a + h_n]$. Comme $E(f_n) = \frac{1}{2h_n} \int_{a-h_n}^{a+h_n} f(u) du$, on obtient $|E(f_n) - f(a)| \leq \varepsilon$.

Donc $E(f_n) \rightarrow f(a)$. La première méthode utilise les taux d'accroissement de F ; la seconde explique la limite comme une moyenne de valeurs de f près de a .

13. b) La variance existe puisque C_n est bornée. La variance de la loi binomiale donne :

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

Par 13. a), $E(f_n)$ converge vers le réel $f(a)$ et est donc bornée ; $0 \leq 1 - p_n \leq 1$ et $nh_n \rightarrow +\infty$. Par encadrement, $\mathbb{V}(f_n) \rightarrow 0$.

14. a) Question adaptée au programme actuel : le sujet guide explicitement la soustraction des deux développements, sans supposer une technique générale de calcul sur les développements limités. Taylor-Young, appliqué directement à la fonction F de classe C^2 , donne :

$$F(a + h_n) = F(a) + h_n f(a) + \frac{h_n^2}{2} F''(a) + h_n^2 \varepsilon_n^+,$$

$$F(a - h_n) = F(a) - h_n f(a) + \frac{h_n^2}{2} F''(a) + h_n^2 \varepsilon_n^-,$$

où $\varepsilon_n^+ \rightarrow 0$ et $\varepsilon_n^- \rightarrow 0$. La soustraction annule le terme constant et les termes quadratiques. Ainsi :

$$p_n = 2h_n f(a) + h_n^2(\varepsilon_n^+ - \varepsilon_n^-) = 2h_n f(a) + o(h_n^2) = \theta_n^2 + o(h_n^2).$$

14. b) Puisque $f(a) > 0$, on peut diviser par $2h_n f(a)$. Le quotient du reste par cette quantité tend vers zéro, car il est de la forme $h_n \varepsilon_n / (2f(a))$, avec $\varepsilon_n \rightarrow 0$. Donc $p_n / (2h_n f(a)) \rightarrow 1$, soit $p_n \sim 2h_n f(a)$. Il en résulte $p_n \rightarrow 0$ et $np_n \sim 2f(a)nh_n \rightarrow +\infty$.

14. c) Comme $2h_n f(a) = \theta_n^2$, la définition de $\hat{f}_n$ se simplifie en :

$$\hat{f}_n = \frac{C_n}{\theta_n \sqrt{n}} - \theta_n \sqrt{n}.$$

En ajoutant et soustrayant $np_n / (\theta_n \sqrt{n})$, on obtient l'identité demandée. Le carré du premier coefficient est :

$$\left(\frac{\sigma_n}{\theta_n \sqrt{n}} \right)^2 = \frac{p_n(1 - p_n)}{2h_n f(a)} \rightarrow 1.$$

Ce coefficient est strictement positif, donc tend vers 1 par continuité de la racine carrée.

D'après 14. a), il existe $\varepsilon_n \rightarrow 0$ tel que $p_n - \theta_n^2 = h_n^2 \varepsilon_n$. Ainsi :

$$\sqrt{n} \left(\frac{p_n}{\theta_n} - \theta_n \right) = \frac{\sqrt{n}(p_n - \theta_n^2)}{\theta_n} = \frac{\sqrt{nh_n^3}}{\sqrt{2f(a)}} \varepsilon_n \rightarrow 0.$$

La dernière limite utilise l'hypothèse $nh_n^3 \rightarrow 0$.

15. a) Le quantile t_α est strictement positif puisque $1 - \alpha/2 > 1/2$. La convergence admise fournit :

$$P(-t_\alpha \leq \hat{f}_n \leq t_\alpha) \rightarrow \Phi(t_\alpha) - \Phi(-t_\alpha) = 1 - \alpha.$$

Comme $\theta_n^2 = 2h_n f(a)$ et $f(a) > 0$, l'événement considéré équivaut successivement à :

$$\hat{f}_n^2 \leq \eta_\alpha,$$

$$\frac{2nh_n}{f(a)}(f_n - f(a))^2 \leq \eta_\alpha,$$

$$(f_n - f(a))^2 - \frac{\eta_\alpha}{2nh_n} f(a) \leq 0.$$

En développant le carré, on retrouve exactement l'événement de la question ; sa probabilité tend donc vers $1 - \alpha$.

15. b) Posons $q_n = \eta_\alpha / (4nh_n) > 0$. Puisque $f_n \geq 0$, $(f_n + q_n)^2 - f_n^2 = 2q_n f_n + q_n^2 > 0$, donc Δ_n est bien définie. Pour un réel z , une complétion du carré donne :

$$z^2 - (2f_n + 2q_n)z + f_n^2 = (z - f_n - q_n)^2 - [(f_n + q_n)^2 - f_n^2].$$

Cette expression est négative ou nulle si et seulement si $z \in [f_n + q_n - \Delta_n, f_n + q_n + \Delta_n]$.

En prenant $z = f(a)$ et en utilisant 15. a), on obtient le résultat. Cet intervalle aléatoire a donc un niveau de couverture asymptotique égal à $1 - \alpha$; aucune formule générale d'intervalle pour une densité n'est supposée connue.

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

16. a) Par indépendance des X_i et lemme des coalitions, X_k et

$$Y_k = \sum_{j \neq k} X_j$$

sont indépendantes. Il en est de même de X_k^2 et $f'(Y_k)$. Les espérances existent par les hypothèses et admissions de l'énoncé. Donc :

$$E(X_k^2 f'(Y_k)) = E(X_k^2) E(f'(Y_k)) = v_k E(f'(Y_k)).$$

Le membre gauche demandé se simplifie en

$$\sum_{k=1}^n v_k E(f'(S_n)),$$

qui vaut $E(f'(S_n))$ car

$$\sum v_k = 1.$$

16. b) Par la même indépendance et le centrage de X_k ,

$$E(X_k f(Y_k)) = E(X_k) E(f(Y_k)) = 0. \text{ Ainsi :}$$

$$\sum_{k=1}^n E(X_k (f(S_n) - f(Y_k))) = \sum_{k=1}^n E(X_k f(S_n)) = E \left(\left(\sum_{k=1}^n X_k \right) f(S_n) \right).$$

Le dernier membre est $E(S_n f(S_n))$ par définition de S_n .

16. c) On soustrait l'égalité de 16. b) à celle de 16. a), puis on regroupe, pour chaque k , les termes :

$$E(X_k^2 f'(Y_k)) - E(X_k (f(S_n) - f(Y_k))) = E(X_k [X_k f'(Y_k) - (f(S_n) - f(Y_k))]).$$

Cette substitution donne précisément l'identité demandée.

17. a) La fonction $t \mapsto f(a + tb)$ est de classe C^1 sur $[0, 1]$ et a pour dérivée $bf'(a + tb)$.

Par le théorème fondamental de l'intégration :

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

Par ailleurs $bf'(a) = \int_0^1 bf'(a) dt$. La soustraction prouve l'égalité, pour tout signe de b , y compris $b = 0$. On n'utilise aucune formule générale de Taylor avec reste intégral.

17. b) Comme $|f''| \leq 2$, l'inégalité des accroissements finis appliquée à f' donne $|f'(a) - f'(a + tb)| \leq 2t|b|$ pour $0 \leq t \leq 1$. Donc :

$$|bf'(a) - (f(a + b) - f(a))| \leq \int_0^1 |b| 2t|b| dt = b^2.$$

17. c) D'une part, $|f'(S_n) - f'(Y_k)| \leq 2|S_n - Y_k| = 2|X_k|$. D'autre part, 17. b), appliquée ponctuellement avec $a = Y_k$ et $b = X_k$, donne :

$$|X_k f'(Y_k) - (f(S_n) - f(Y_k))| \leq X_k^2.$$

L'identité 16. c), l'inégalité triangulaire et la croissance de l'espérance fournissent donc :

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

Notons M ce dernier membre, fini par les hypothèses. Pour tout $h \in W$, la partie 1 donne une fonction f_h de classe C^2 avec $|f_h| \leq 1$, $|f_h''| \leq 2$, ainsi que :

$$|E(h(S_n)) - E(h(N))| = |E(f_h'(S_n) - S_n f_h'(S_n))| \leq M.$$

On peut donc appliquer (R_2) à S_n avec $M_{S_n} = M$. Cela donne (R_3) pour tout réel x . Il s'agit d'une borne démontrée pour la somme considérée, et non d'un théorème d'approximation invoqué sans preuve.

18. a) Les variables X_k sont indépendantes, centrées et possèdent un moment d'ordre trois, par transformation affine des Z_k . De plus :

$$v_k = \frac{\mathbb{V}(Z_k)}{\sigma^2 n} = \frac{1}{n}, \quad E(|X_k|) = \frac{s_1}{\sigma \sqrt{n}}, \quad E(|X_k|^3) = \frac{s_3}{\sigma^3 n^{3/2}}.$$

Ainsi

$$\sum v_k = 1,$$

ce qui vérifie toutes les hypothèses de (R_3) . Le membre sous la racine s'écrit :

$$3 \left(2 \frac{s_1}{\sigma \sqrt{n}} + \frac{s_3}{\sigma^3 \sqrt{n}} \right) = 3 \frac{2\sigma^2 s_1 + s_3}{\sigma^3 \sqrt{n}}.$$

On obtient donc la majoration demandée.

18. b) La borne de 18. a) ne dépend pas de x et tend vers zéro puisque s_1, s_3 et $\sigma > 0$ sont fixes. Par la définition du sujet, (S_n) converge uniformément en loi vers N , donc en loi vers N . Le théorème limite central donne aussi directement cette dernière conclusion :

$$S_n = \sqrt{n} \frac{\bar{Z}_n - E(Z_1)}{\sigma}, \quad \bar{Z}_n = \frac{1}{n} \sum_{k=1}^n Z_k,$$

et les Z_k sont indépendantes de même loi, de variance $\sigma^2 > 0$. Le théorème limite central n'exige pas, pour cette conclusion, le moment d'ordre trois utilisé par la borne quantitative.

19. a) Écrivons $p = p_n, q = 1 - p$ et $\sigma_n = \sqrt{npq} > 0$. La variable X_k vaut q/σ_n avec probabilité p , et $-p/\sigma_n$ avec probabilité q . Donc :

$$E(|X_k|) = \frac{pq + qp}{\sigma_n} = \frac{2pq}{\sigma_n} = \frac{2\sigma_n}{n},$$

$$E(|X_k|^3) = \frac{pq^3 + qp^3}{\sigma_n^3} = \frac{pq(p^2 + q^2)}{\sigma_n^3} = \frac{p^2 + q^2}{n\sigma_n} \leq \frac{2}{n\sigma_n}.$$

La dernière borne utilise $p^2 \leq 1, q^2 \leq 1$. On pourrait même utiliser $p^2 + q^2 \leq 1$ pour obtenir une borne plus fine ; la borne demandée suffit.

19. b) Les X_k sont indépendantes, centrées, à support fini. Leur variance est $v_k = pq/\sigma_n^2 = 1/n$, donc (R_3) s'applique. Avec 19. a) :

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

Par conséquent :

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

19. c) La somme

$$\sum_{k=1}^n Z_k$$

a la même loi que T_n . Donc S_n a la même loi que $(T_n - np_n)/\sigma_n$, et leurs fonctions de répartition coïncident. Puisque $p_n \rightarrow 0$ et $np_n \rightarrow +\infty$, $\sigma_n^2 = np_n(1 - p_n) \rightarrow +\infty$, d'où $1/\sigma_n \rightarrow 0$. En outre $p_n(1 - p_n) \leq 1/4$, donc :

$$0 \leq \frac{\sigma_n}{n} = \sqrt{\frac{p_n(1 - p_n)}{n}} \leq \frac{1}{2\sqrt{n}} \rightarrow 0.$$

La borne de 19. b) tend donc vers zéro indépendamment de x , ce qui est la convergence uniforme en loi demandée. On n'applique pas le théorème limite central usuel à une loi de Bernoulli dont le paramètre change avec n .

20. a) Comme $\alpha > 0$, les événements $[\alpha X + \beta \leq x]$ et $[X \leq (x - \beta)/\alpha]$ sont égaux. Il en est de même avec N . Par conséquent :

$$F_{\alpha X + \beta}(x) = F_X\left(\frac{x - \beta}{\alpha}\right), \quad F_{\alpha N + \beta}(x) = \Phi\left(\frac{x - \beta}{\alpha}\right).$$

La différence en valeur absolue est exactement $d_X((x - \beta)/\alpha)$.

20. b) Par hypothèse, il existe $\delta_n \rightarrow 0$ avec $d_{V_n}(u) \leq \delta_n$ pour tous réels u . La question 20. a), appliquée à $\alpha = a_n$, $\beta = b_n$, donne :

$$|P(a_n V_n + b_n \leq x) - P(a_n N + b_n \leq x)| \leq \delta_n \rightarrow 0.$$

La borne uniforme est essentielle : le point $(x - b_n)/a_n$ dépend de n .

20. c) Comme $a_n \rightarrow a > 0$ et $b_n \rightarrow b$, $(x - b_n)/a_n \rightarrow (x - b)/a$. La continuité de Φ fournit :

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

Avec 20. b), les fonctions de répartition de $a_n V_n + b_n$ convergent donc vers celle de $aN + b$ en tout réel. Cette dernière est continue ; c'est exactement la convergence en loi demandée. Par transformation affine d'une normale, $aN + b$ suit la loi normale de moyenne b et de variance a^2 . Le raisonnement repose entièrement sur les fonctions de répartition et des coefficients déterministes.

21. a) Par 12, C_n suit la loi binomiale (n, p_n) . La question 14. b) donne $p_n \rightarrow 0$ et $np_n \rightarrow +\infty$. Les hypothèses de 19. c) sont donc satisfaites ; ainsi $D_n = (C_n - np_n)/\sqrt{np_n(1 - p_n)}$ converge uniformément en loi vers N .

21. b) D'après 14. c), $\hat{f}_n = a_n D_n + b_n$, avec :

$$a_n = \frac{\sigma_n}{\theta_n \sqrt{n}} > 0, \quad b_n = \sqrt{n} \left(\frac{p_n}{\theta_n} - \theta_n \right), \quad a_n \rightarrow 1, \quad b_n \rightarrow 0.$$

La question 20 s'applique à $V_n = D_n$, $a = 1$, $b = 0$. Elle donne $\hat{f}_n \rightarrow N$ en loi, ce qui démontre le résultat admis dans la partie 3.

Relevé des adaptations

Question 7. b) : ajout de la documentation utile de NumPy et de matplotlib. Le tracé explicatif du corrigé est un SVG intégré, calculé exactement à partir de la formule de γ ; aucune figure n'était imprimée dans le sujet source.

Question 11 : ajout de la documentation pandas, de l'hypothèse de données numériques sans valeurs manquantes, et d'un accès par position explicite dans la réponse. La

question et ses blancs sont conservés. Le langage source étant déjà Python, aucune conversion d'un ancien langage n'a été nécessaire.

Question 14. a) : ajout d'une indication locale pour appliquer directement Taylor-Young à F et soustraire ses deux égalités. Aucun développement au-delà de l'ordre deux ni règle générale de calcul sur les développements limités n'est utilisé.