Sujet *

Partie I

1/ $G_{a,b}$ est dérivable sur $\mathbb{R}^+$ par composition avec :

$$\forall x \in \mathbb{R}^+, \quad G_{a,b}'(x) = -\underbrace{(a+bx)}_{>0} \exp\left(-ax - \frac{b}{2}x^2\right) < 0$$

La fonction $G_{a,b}$ est donc strictement décroissante.
De plus

$$G_{a,b}(x) \xrightarrow[+\infty]{} 0 \quad \text{et} \quad G_{a,b}(0) = 1.$$

Donc $G_{a,b}$ réalise une bijection de $\mathbb{R}^+$ dans $]0;1]$.

2/ Soient $x \in \mathbb{R}^+$, $u \in [0,1[$

$$G_{a,b}(x) = 1-u \iff -ax - \frac{b}{2}x^2 = \ln(1-u)$$
$$\iff \frac{b}{2}x^2 + ax + \ln(1-u) = 0$$

Cette équation polynomiale a pour discriminant

$$\Delta = a^2 - 4\left(\frac{b}{2}\right)(+\ln(1-u)) = a^2 - 2b\ln(1-u) > 0$$

Il y a deux solutions a priori

$$x_1 = \frac{-a-\sqrt{\Delta}}{b} < 0 \quad \text{et} \quad x_2 = \frac{-a+\sqrt{\Delta}}{b}$$

mais seule la deuxième est positive, il vient :

$$G_{a,b}^{-1}(1-u) = \frac{-a+\sqrt{a^2-2b\ln(1-u)}}{b}.$$

3/ La fonction $G_{a,b}$ est continue sur $\mathbb{R}^+$. On a donc une intégrale généralisée en $+\infty$. Or par les croissances comparées

$$x^2 G_{a,b}(x) \xrightarrow[x \to +\infty]{} 0 \qquad \text{puis} \quad G_{ab}(x) = \underset{+\infty}{O}\left(\frac{1}{x^2}\right).$$

Or l'intégrale de Riemann $\int_1^{+\infty} 1/x^2 dx$ est convergente. Par le critère de négligeabilité (avec $1/x^2 > 0$), on a bien la convergence de

$$\int_0^{+\infty} G_{a,b}(x)\, dx.$$

4/ Vérifier que :

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{1}{\sqrt{2\pi \cdot \frac{1}{b}}} \exp\left(-\frac{1}{2} \frac{(x-(-a/b))^2}{1/b}\right)$$

par identification, on reconnaît la densité d'une loi normale

$$\underline{\mathcal{N}\left(-\frac{a}{b} ; \frac{1}{b}\right)}.$$

5/ Vérifier que pour tout $x \in \mathbb{R}$,

$$f(x) = \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{a^2}{2b}\right) G_{a,b}(x).$$

On retrouve la convergence de $\int_0^{+\infty} G_{a,b}(x) dx$ et l'égalité

$$\int_0^{+\infty} G_{a,b}(x)\, dx = \sqrt{\frac{2\pi}{b}} \exp\left(\frac{a^2}{2b}\right) \int_0^{+\infty} f(x)\, dx.$$

Or $\displaystyle\int_0^{+\infty} f(x)\,dx = P(X \geqslant 0)$ où $X \hookrightarrow \mathcal{N}\left(-\frac{a}{b}, \frac{1}{b}\right)$

$$= 1 - P(X \leqslant 0) \quad X \text{ à densité}$$

$$= 1 - P\left(X - E(X) \leqslant a/b\right).$$

$$= 1 - P\left(\frac{X - E(X)}{\sqrt{Var(X)}} \leqslant \frac{a/b}{\sqrt{1/b}}\right)$$

$$= 1 - P(X^* \leqslant a/\sqrt{b}) \quad \text{où } X^* \hookrightarrow \mathcal{N}(0,1).$$

$$= 1 - \phi(a/\sqrt{b}).$$

$$\int_0^{+\infty} f(x)\,dx = \phi(-a/\sqrt{b}).$$

Concluons :

$$\underline{\int_0^{+\infty} G_{a,b}(x)\,dx = \sqrt{\frac{2\pi}{b}}\exp\left(\frac{a^2}{2b}\right)\Phi\left(-\frac{a}{\sqrt{b}}\right)}.$$

6/ → la fonction $f_{a,b}$ est positive sur $\mathbb{R}$.

→ La fonction $f_{a,b}$ est continue sur $\mathbb{R}^*$.

→ Pour $A \in \mathbb{R}^+$,

$$\int_{-\infty}^{A} f_{a,b}(x)\,dx = \int_0^A f_{a,b}(x)\,dx$$

$$= \int_0^A -G_{a,b}'(x)\,dx = G_{a,b}(0) - G_{ab}(A)$$

$$\xrightarrow[A \to +\infty]{} 1 - 0 = 1.$$

Ce qui conclut, $f_{a,b}$ est bien une densité.

7/ Intégrons par parties (les fonctions considérées sont C^1). Pour $A \in \mathbb{R}_*^+$,

$$\int_{-\infty}^{A} x f_{a,b}(x)\,dx = \int_0^A x f_{a,b}(x)\,dx$$
$$= \int_0^A x(a+bx)\exp\left(-ax-\frac{b}{2}x^2\right)dx$$
$$= \int_0^A -x\, G_{a,b}'(x)\,dx.$$
$$= \left[-x G_{a,b}(x)\right]_0^A + \int_0^A G_{a,b}(x)\,dx.$$

Comme le crochet est nul lorsque $A \to +\infty$, on en déduit la convergence et l'égalité

$$\int_{-\infty}^{+\infty} x f_{a,b}(x)\,dx = \int_0^{+\infty} G_{a,b}(x)\,dx.$$

Comme l'intégrande de gauche est positive, on a bien une convergence absolue. L'espérance est bien définie et l'égalité prouvée.

8/ On retrouve ici la méthode d'inversion.
$1-U$ est à valeurs dans $[0;1[$ et $G_{ab}^{-1}(1-U)$ à valeurs dans $\mathbb{R}^+$ (question 2).

Notons F_{ab}, la fonction de répartition de $G_{ab}^{-1}(1-U)$. On a donc F_{ab} nulle sur $\mathbb{R}_*^-$.

Soit $x \in \mathbb{R}^+$

$$F_{ab}(x) = P\left(G_{ab}^{-1}(1-U) \leq x\right).$$

$$= P\left(1-U \geqslant G_{ab}(x)\right) \quad \text{strict. décroiss de } G_{ab}.$$

$$= P\left(U \leq 1 - G_{ab}(x)\right).$$

$$= F_U\left(1 - G_{ab}(x)\right). \quad F_U \text{ fonction de rép. de } \mathcal{U}([0;1])$$

Or $G_{ab}(x) \in [0;1]$ et pour $t \in [0;1]$, $F_U(t) = t$.

On obtient

$$F_{ab}(x) = 1 - G_{ab}(x).$$

En reprenant le calcul de la question 7, on montre que

$$x \in \mathbb{R} \longmapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - G_{ab}(x) & \text{si } x \geqslant 0 \end{cases}$$

est justement la fonction de répartition d'une loi $\mathcal{E}_p(a,b)$. Comme la fonction de répartition caractérise la loi

$$G_{a,b}^{-1}(1-U) \hookrightarrow \mathcal{E}_p(a,b).$$

9. Reprenons la question 2

```
def reciproq(a, b, u):
    Delta = a**2 - 2*b*np.log(1-u)
    return (-a + Delta**(1/2)) / b.
```

```
def simuE(a, b):
    u = rd.random().
    return reciproq(a, b, u).
```

10/ M_n est à valeurs dans $\mathbb{R}^+$ puisque chaque X_i l'est. Soit $x \in \mathbb{R}^+$

$$P(M_n \geq x) = \prod_{i=1}^{n} P(X_i \geq x) \quad \text{par indépendance}$$

$$= P(X_1 \geq x)^n \quad \text{même loi}$$

$$= (1 - P(X_1 \leq x))^n \quad X_1 \text{ à densité}$$

$$= G_{ab}(x)^n$$

$$= \exp\left(-(an)x - \frac{(bn)}{2}x^2\right)$$

$$P(M_n \geq x) = G_{na,nb}(x).$$

puis $\quad P(M_n \leq x) = 1 - G_{na,nb}(x)$.

De plus $\quad P(M_n \leq x) = 0$ si $x \in \mathbb{R}^-$.

On reconnaît la fonction de répartition de $\mathcal{E}_p(a,b)$.
D'où

$$\underline{M_n \hookrightarrow \mathcal{E}_p(a,b).}$$

11/ La variable U_n est positive donc:

$$\forall x \in \mathbb{R}^-_*, \quad F_{U_n}(x) = 0.$$

De plus, U_n ne peut dépasser nh. (par définition du minimum), donc

$$\forall x \in [nh, +\infty[, \qquad F_{U_n}(x) = 1$$

Soit $x \in [0; nh]$

$$\begin{aligned}
F_{U_n}(x) &= P(U_n \leq x) = P(n H_n \leq x) \\
&= P\left(H_n \leq \frac{x}{n}\right) \\
&= P\left(\left[M_n \leq \frac{x}{n}\right] \cap \left[h \leq \frac{x}{n}\right]\right) \\
&= P\left(M_n \leq \frac{x}{n}\right) \\
&= 1 - G_{an, bn}\left(\frac{x}{n}\right) \\
&= 1 - \exp\left(-an \cdot \frac{x}{n} - \frac{bn}{2}\left(\frac{x}{n}\right)^2\right)
\end{aligned}$$

On a bien vérifié les trois égalités.

12. Non, U_n n'est pas une variable à densité car

$$P(U_n = hn) = P\left(\bigcap_{i=1}^{n} [X_i \geqslant h]\right) = P(M_n \geqslant h) \neq 0$$

On peut aussi vérifier que la fonction de répartition n'est pas continue en nh car

$$\lim_{x \to nh^-} F_{U_n}(x) = 1 - \exp(*) < 1 = F_{U_n}(nh).$$

13.b) Pour 100 valeurs x ~~espacées~~ régulièrement espacées entre 0 et 4., le programme donne une approximation de

$$P(U_n \leq x).$$

(On effectue 2000 simulations et on regarde la fréquence d'apparition de l'événement $[U_n \leq x]$).

On trace ensuite une approximation de la courbe de F_{U_n}.

On voit bien un saut de discontinuité en $nh = 2{,}1$.

La variable U_n n'est pas à densité.

13.a)

```
def simuU(a,b,h,n):
    H = np.zeros(n+1)
    H(0) = h
    for k in range(1,n+1):
        H[k] = simuE(a,b)
    return n * np.min(H).
```

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

Si $x \leq 0$ alors on a directement

$$F_{U_n}(x) = 0 \xrightarrow[n\to+\infty]{} 0.$$

Si $x \geq 0$, pour n suffisament grand ($x < nh$) et

$$F_{U_n}(x) = 1 - \exp\left(-ax - \frac{b}{2n}x^2\right) \xrightarrow[n\to+\infty]{} 1 - \exp(-ax).$$

On constate que si on note F_a, la fonction de répartition de la loi exponentielle $\mathcal{E}(a)$, on a :

$$\forall x \in \mathbb{R}, \qquad F_{U_n}(x) \xrightarrow[n\infty]{} F_a(x).$$

C'est la définition de la convergence en loi de $(U_n)_{n\in\mathbb{N}^*}$ vers une variable X de loi $\mathcal{E}(a)$.

15. On sait que

$$\forall x \in \mathbb{R}, \quad F_Y(x) = (1-e^{-x})\,\mathbb{1}_{[0;+\infty[}(x).$$

Ainsi

$$\begin{aligned}
& P(c \leq Y \leq d) = 1-\alpha \\
\Leftrightarrow \quad & P(c < Y \leq d) = 1-\alpha \qquad (Y \text{ à densité}) \\
\Leftrightarrow \quad & F_Y(d) - F_Y(c) = 1-\alpha \\
\Leftrightarrow \quad & -e^{-d} + e^{-c} = 1-\alpha \qquad (*)
\end{aligned}$$

et

$$\begin{aligned}
& P(Y \leq c) = \frac{\alpha}{2} \\
\Leftrightarrow \quad & 1 - e^{-c} = \frac{\alpha}{2} \\
\Leftrightarrow \quad & 1 - \frac{\alpha}{2} = e^{-c} \quad \Leftrightarrow \quad -\ln\left(1-\frac{\alpha}{2}\right) = c
\end{aligned}$$

En revenant à $(*)$, il vient

$$1 - \frac{\alpha}{2} - e^{-d} = 1-\alpha \quad \Leftrightarrow \quad \frac{\alpha}{2} = e^{-d}$$
$$\Leftrightarrow \quad -\ln\left(\frac{\alpha}{2}\right) = d.$$

On pose donc

$$c = -\ln\left(1-\frac{\alpha}{2}\right) \text{ et } d = -\ln\left(\frac{\alpha}{2}\right).$$

16. On a

$$P\left(a \in \left[\frac{c}{U_n} ; \frac{d}{U_n}\right]\right) = P\left(\frac{c}{U_n} \leq a \leq \frac{d}{U_n}\right).$$
$$= P(c \leq aU_n \leq d).$$

On a vu que $(U_n) \xrightarrow[n\to+\infty]{\mathcal{L}} X$ où $X \hookrightarrow \mathcal{E}(a)$, c'est-à-dire

$$\forall x \in \mathbb{R} \quad F_{U_n}(x) \xrightarrow[n\infty]{} F_X(x).$$

puis

$$\forall x \in \mathbb{R} \quad P(aU_n \leq x) = P\left(U_n \leq \frac{x}{a}\right) \xrightarrow[n\infty]{} F_X\left(\frac{x}{a}\right) = F_Y(x)$$

où $Y \hookrightarrow \mathcal{E}(1)$.

Ainsi

$$P(c \leq aU_n \leq d) = P(c < aU_n \leq d)$$
$$= P(aU_n \leq d) - P(aU_n \leq c).$$
$$\xrightarrow[n\infty]{} F_Y(d) - F_Y(c) = 1 - \alpha.$$

Ce qui conclut.

17. La variable S_i ne prend que 0 et 1 comme valeurs.
S_i suit donc une loi de Bernoulli de paramètre

$$p = P(X_i \geq h) = 1 - P(X_i \leq h) \quad X_i \text{ à densité}$$
$$= 1 - (1 - G_{a,b}(h)) \quad \text{car } X_i \hookrightarrow \mathcal{E}_p(a,b).$$
$$= G_{ab}(h)$$

. Dans le cas d'une variable de Bernoulli, l'espérance se confond avec le paramètre

$$E(S_i) = G_{ab}(h).$$

* De même $S_i D_i$ est une variable de Bernoulli mais ici on a

$$P(S_i D_i = 1) = P([X_i \geq h] \cap [X_i \leq 1])$$
$$= 0$$

car on a $h \geq 2$. (voir juste avant la partie II).

$S_i D_i$ est donc presque surement nulle

$$E(S_i D_i) = 0.$$

18. D'une part

$$E(S_i D_i) = 0 \neq \underbrace{E(S_i)}_{\neq 0} \underbrace{E(D_i)}_{1 - G_{ab}(1) \neq 0}$$

Ainsi S_i, D_i ne peuvent être indépendantes car dans le cas d'indépendance, le produit des espérances est l'épérance du produit.

Par contre, pour $i \neq j$, les variables

$$S_i = \mathbb{1}_{[X_i \geq h]} \quad , \quad D_j = \mathbb{1}_{[X_j \leq 1]}$$

sont indépendantes car X_i et X_j le sont (lemme des coalitions).

19. Par bilinéarité de la covariance

$$\mathrm{Cov}(\overline{S_n}, \overline{D_n}) = \mathrm{Cov}\left(\sum_{i=1}^{n} \frac{1}{n} S_i, \sum_{j=1}^{n} \frac{1}{n} D_j\right).$$

$$= \sum_{i,j=1}^{n} \frac{1}{n^2} \mathrm{Cov}(S_i, D_j)$$

$$= \sum_{i=1}^{n} \frac{1}{n^2} \mathrm{Cov}(S_i, D_i)$$

pour $i \neq j$, S_i et D_j sont indépendants

$$= \sum_{i=1}^{n} \frac{1}{n^2} \left(\underbrace{E(S_i D_i)}_{=0} - E(S_i)E(D_i)\right)$$

$$= - \sum_{i=1}^{n} \frac{1}{n^2} G_{ab}(h)\,(1 - G_{ab}(1)).$$

$$\mathrm{Cov}(\overline{S_n}, \overline{D_n}) = -\frac{1}{n} G_{ab}(h)\,(1 - G_{ab}(1)).$$

20. Les variables (S_i) sont de même loi, indépendantes et avec une variance. D'après la loi faible des grands nombres

$$\overline{S_n} \xrightarrow[n \to +\infty]{P} E(S_1) = G_{ab}(h).$$

Dit autrement, $\overline{S_n}$ est convergent vers $G_{ab}(h)$.

De plus, $\overline{S_n}$ est sans biais car $E(\overline{S_n}) = G_{ab}(h)$.
On a de même avec $\overline{D_n}$ avec $1 - G_{ab}(1)$.

21. On a par indépendance des (X_i)

$$V(\overline{S_n}) = \frac{1}{n} V(S_1) = \frac{1}{n} G_{ab}(h)\,(1 - G_{ab}(h)) \leq \frac{1}{4n}.$$

puis pour $\varepsilon > 0$

$$P(|\overline{S_n} - G_{ab}(h)| > \varepsilon) \leq \frac{V(\overline{S_n})}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

Par passage au complémentaire

$$P\left(|\overline{S_n} - G_{ab}(h)| \leq \varepsilon\right) \geq 1 - \frac{1}{4n\varepsilon^2}.$$

avec le choix $\varepsilon = \frac{1}{2\sqrt{n\alpha}}$, il vient

$$P\left(|\overline{S_n} - G_{ab}(h)| \leq \varepsilon\right) \geq 1 - \alpha.$$

Un intervalle de confiance au niveau de confiance $1-\alpha$ est donc

$$\left[\overline{S_n} - \frac{1}{2\sqrt{n\alpha}} \, ; \, \overline{S_n} + \frac{1}{2\sqrt{n\alpha}}\right].$$

22/ Comme $\frac{1}{n} \xrightarrow[n\infty]{} 0$, on vérifie que

$$1 + \frac{1}{n} - \overline{D_n} \xrightarrow[n\to+\infty]{P} 1 - \underbrace{(1 - G_{ab}(1))}_{= G_{ab}(1)} \quad \text{et} \quad \overline{S_n} + \frac{1}{n} \xrightarrow[n\infty]{P} G_{ab}(h).$$

La fonction logarithme étant continue sur $\mathbb{R}_*^+$,

$$\ln\left(1 + \frac{1}{n} - \overline{D_n}\right) \xrightarrow[n\infty]{P} G_{ab}(1) \quad \text{et} \quad \ln\left(\overline{S_n} + \frac{1}{n}\right) \xrightarrow[n\infty]{P} G_{ab}(h)$$

On retrouve la définition des estimateurs convergents.

23/ À partir de l'inégalité triangulaire, on a.

$$\left[|(\lambda Z_n - \mu R_n) - (\lambda z(a,b) - \mu r(a,b)| \geq \varepsilon\right]$$

$$\subset \left[\lambda |Z_n - z(a,b)| + \mu |R_n - r(a,b)| \geq \varepsilon\right].$$

$$\subset \left[|Z_n - z(a,b)| \geqslant \frac{\varepsilon}{2\lambda} \right] \cup \left[|R_n - r(a,b)| \geqslant \frac{\varepsilon}{2\mu} \right]$$

En utilisant la formule du crible.

$$\forall A, B \in \mathcal{A} \quad P(A \cup B) \leq P(A) + P(B)$$

D'où

$$P\left(|(\lambda Z_n - \mu R_n) - (\lambda z(a,b) - \mu r(a,b))| \geqslant \varepsilon \right)$$

$$\leq P\left(\left[|Z_n - Z(a,b)| \geqslant \frac{\varepsilon}{2\lambda} \right] \cup \left[|R_n - r(a,b)| \geqslant \frac{\varepsilon}{2\mu} \right] \right)$$

$$\leq P\left(|Z_n - z(a,b)| \geqslant \frac{\varepsilon}{2\lambda} \right) + P\left(|R_n - r(a,b)| \geqslant \frac{\varepsilon}{2\mu} \right).$$

24. Reprenons l'inégalité précédente avec

$$\lambda = \frac{2}{h-1} \quad \text{et} \quad \mu = \frac{2}{h(h-1)}$$

de sorte que $\lambda z(a,b) - \mu r(a,b)$

$$= \frac{2}{h-1}\left(-a - \frac{b}{2}\right) - \frac{2}{h(h-1)}\left(-ah - \frac{b}{2} h^2\right) = b.$$

On a alors par convergence de Z_n (resp R_n) vers $z(a,b)$ (resp $r(a,b)$).

$$B_n \xrightarrow[n \to +\infty]{P} b.$$