Lycée Clemenceau — **Jeudi 14 février 2026**
ECG 2 — **Durée : 4 heures**

Corrigé du devoir Surveillé de Mathématiques $n^o 6$

1. Sujet type EDHEC

Corrigé de l'exercice 1. On considère deux variables aléatoires réelles X et Y définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et suivant toutes deux la loi normale centrée réduite (de densité notée φ et de fonction de répartition notée Φ). On pose $Z = \max\{X, Y\}$ et l'on se propose de déterminer la loi de Z, son espérance et sa variance.

(1) (a) Montrons que Z est une variable aléatoire à densité. Pour ce faire, on va calculer la fonction de répartition de Z. Pour tout $x \in \mathbb{R}$, on trouve que :

$$F_Z(x) = P([Z \leq x]) = P([\max\{X, Y\} \leq x]) = P([X \leq x] \cap [Y \leq x]).$$

Comme X et Y sont indépendantes et suivent la loi normale centrée réduite, on obtient que :

$$F_Z(x) = P([X \leq x])P([Y \leq x]) = \Phi(x)\Phi(x) = \Phi^2(x).$$

Comme Φ est la fonction de répartition de la loi normale centrée réduite, la fonction Φ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$. En particulier, la fonction $F_Z = \Phi^2$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ comme carré d'une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Par conséquent :

$$\boxed{Z \text{ est une variable à densité.}}$$

(b) Vérifions que Z admet pour densité la fonction f définie pour tout $x \in \mathbb{R}$ par $f(x) = 2\varphi(x)\Phi(x)$. D'après la question précédente, une densité de Z est donnée par la dérivée de $F_Z = \Phi^2$. Dès lors, une densité f de Z est donnée pour tout $x \in \mathbb{R}$ par $f(x) = (\Phi^2)'(x) = 2\Phi'(x)\Phi(x)$, et donc :

$$\boxed{f(x) = 2\varphi(x)\Phi(x).}$$

(2) (a) Rappelons la valeur de l'intégrale $\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt$. D'après le cours, on sait que :

$$\boxed{\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt = \sqrt{2\pi}.}$$

(b) Etablissons la convergence et déterminons la valeur de $\int_{-\infty}^{+\infty} e^{-t^2}\,dt$. Pour ce faire, on pose $u = \sqrt{2}t$. Alors u est une bijection strictement croissante de $\mathbb{R}$ dans $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$. De plus, on voit que $du = \sqrt{2}dt$, que u tend vers $-\infty$ quand t tend vers $-\infty$ et que u tend vers $+\infty$ quand t tend vers $+\infty$. Dès lors, par changement de variable, on trouve que :

$$\int_{-\infty}^{+\infty} e^{-t^2}\,dt = \int_{-\infty}^{+\infty} e^{-(\frac{u}{\sqrt{2}})^2}\,\frac{du}{\sqrt{2}} = \frac{1}{\sqrt{2}}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\,du.$$

En particulier, on obtient par linéarité que l'intégrale $\int_{-\infty}^{+\infty} e^{-t^2}\,dt$ converge si et seulement si l'intégrale $\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\,du$. Mais comme cette dernière converge d'après le cours, il s'ensuit que l'intégrale $\int_{-\infty}^{+\infty} e^{-t^2}\,dt$ converge, et que de plus :

$$\int_{-\infty}^{+\infty} e^{-t^2}\,dt = \frac{1}{\sqrt{2}}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\,du = \frac{1}{\sqrt{2}} \times \sqrt{2\pi}.$$

Par conséquent, on en déduit que :

$$\boxed{\text{l'intégrale } \int_{-\infty}^{+\infty} e^{-t^2}\,dt \text{ converge et vaut } \sqrt{\pi}.}$$

(c) Vérifions que $\varphi'(x) = -x\varphi(x)$ pour tout $x \in \mathbb{R}$. Pour tout $x \in \mathbb{R}$, on trouve que :

$$\varphi'(x) = \left(e^{-\frac{x^2}{2}}\right)' = -xe^{-\frac{x^2}{2}} = -x\varphi(x).$$

Par conséquent, on en déduit que :

$$\boxed{\forall x \in \mathbb{R},\ \ \varphi'(x) = -x\varphi(x).}$$

(d) A l'aide d'une intégration par parties, montrons que :
$$\int_0^{+\infty} xf(x)dx = \frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_0^{+\infty} e^{-t^2}dt.$$
Pour ce faire, on se fixe un réel $c > 0$, et l'on pose $u(x) = x\varphi(x) = -\varphi'(x)$ et $v(x) = \Phi(x)$ pour tout $x \in [0, c]$. Alors u et v sont de classe $\mathcal{C}^1$ sur $[0, c]$. De plus, par intégration par parties et linéarité de l'intégrale, on trouve que :
$$\begin{aligned}
\int_0^c xf(x)dx &= \int_0^c 2x\varphi(x)\Phi(x)dx \\
&= -2\int_0^c \varphi'(x)\Phi(x)dx \\
&= -2\int_0^c u'(x)v(x)dx \\
&= -2\left[u(x)v(x)\right]_0^c + 2\int_0^c u(x)v'(x)dx \\
&= -2\left[\varphi(x)\Phi(x)\right]_0^c + 2\int_0^c \varphi(x)\Phi'(x)dx \\
&= -2\varphi(c)\Phi(c) + 2\varphi(0)\Phi(0) + 2\int_0^c \left(\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right)^2 dx \\
&= -2\varphi(c)\Phi(c) + 2\varphi(0)\Phi(0) + 2\int_0^c \frac{e^{-x^2}}{2\pi}dx \\
&= -2\varphi(c)\Phi(c) + 2\varphi(0)\Phi(0) + \frac{1}{\pi}\int_0^c e^{-x^2}dx.
\end{aligned}$$
Comme $\varphi(c) = \frac{1}{\sqrt{2\pi}}e^{-\frac{c^2}{2}}$ tend vers 0 quand c tend vers $+\infty$, que $\Phi(0) = \frac{1}{2}$ et que $\Phi(c)$ tend vers 1 quand c tend vers $+\infty$, il s'ensuit que :
$$\int_0^c xf(x)dx \underset{c\to+\infty}{\longrightarrow} 2 \times \frac{1}{\sqrt{2\pi}} \times \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} e^{-x^2}dx.$$
Par conséquent, on en déduit que :
$$\boxed{\int_0^{+\infty} xf(x)dx = \frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_0^{+\infty} e^{-t^2}dt.}$$

(e) Montrons de même l'égalité suivante :
$$\int_{-\infty}^0 xf(x)dx = -\frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_{-\infty}^0 e^{-t^2}dt.$$
Pour ce faire, on se fixe un réel $c < 0$, et l'on pose $u(x) = x\varphi(x) = -\varphi'(x)$ et $v(x) = \Phi(x)$ pour tout $x \in [c, 0]$. Alors u et v sont de classe $\mathcal{C}^1$ sur $[c, 0]$. De plus, par intégration par parties et linéarité de l'intégrale, on trouve comme à la question précédente que :
$$\int_c^0 xf(x)dx = -2\varphi(0)\Phi(0) + 2\varphi(c)\Phi(c) + \frac{1}{\pi}\int_c^0 e^{-x^2}dx.$$
Comme $\varphi(c) = \frac{1}{\sqrt{2\pi}}e^{-\frac{c^2}{2}}$ tend vers 0 quand c tend vers $-\infty$, que $\Phi(0) = \frac{1}{2}$ et que $\Phi(c)$ tend vers 0 quand c tend vers $-\infty$, il s'ensuit que :
$$\int_c^0 xf(x)dx \underset{c\to-\infty}{\longrightarrow} -2 \times \frac{1}{\sqrt{2\pi}} \times \frac{1}{2} + \frac{1}{\pi}\int_{-\infty}^0 e^{-x^2}dx.$$
Par conséquent, on en déduit que :
$$\boxed{\int_{-\infty}^0 xf(x)dx = -\frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_{-\infty}^0 e^{-t^2}dt.}$$

(f) Montrons que Z admet une espérance et donnons sa valeur. D'après ce qui précède, les intégrales $\int_{-\infty}^{0} xf(x)dx$ et $\int_{0}^{+\infty} xf(x)dx$ convergent. Dès lors, l'intégrale $\int_{-\infty}^{+\infty} xf(x)dx$ converge absolument, et donc Z admet une espérance. De plus, d'après la relation de Chasles et les questions précédentes, on trouve que :

$$\begin{aligned} E(Z) &= \int_{-\infty}^{+\infty} xf(x)dx \\ &= \int_{-\infty}^{0} xf(x)dx + \int_{0}^{+\infty} xf(x)dx \\ &= -\frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_{-\infty}^{0} e^{-t^2}dt + \frac{1}{\sqrt{2\pi}} + \frac{1}{\pi}\int_{0}^{+\infty} e^{-t^2}dt \\ &= \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-t^2}dt \quad = \quad \frac{1}{\sqrt{\pi}}. \end{aligned}$$

Par conséquent, on en déduit que :

$$\boxed{Z \text{ admet une espérance et } E(Z) = \frac{1}{\sqrt{\pi}}.}$$

(3) (a) Montrons que X^2 et Z^2 suivent la même loi. Pour ce faire, on commence par calculer la fonction de répartition F_{X^2} de X^2, puis une densité f_{X^2} de X^2. Pour tout $x \in \mathbb{R}$, on a :

$$F_{X^2}(x) = P([X^2 \leq x]).$$

Si $x < 0$, alors on voit que $[X^2 \leq x] = \emptyset$, et donc $F_{X^2}(x) = 0$. Si maintenant $x > 0$, alors on a :

$$F_{X^2}(x) = P([X^2 \leq x]) = P([-\sqrt{x} \leq X \leq \sqrt{x}]) = F_X(\sqrt{x}) - F_X(-\sqrt{x}).$$

Par dérivation, on obtient que, pour tout $x > 0$:

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}}\left[f_X(\sqrt{x}) + f_X(-\sqrt{x})\right] = \frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}.$$

Dès lors, il s'ensuit qu'une densité de X^2 est définie pour tout $x \in \mathbb{R}$ par :

$$\boxed{f_{X^2}(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}} & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{cases}.}$$

A présent, calculons la fonction de répartition F_{Z^2} de Z^2, puis une densité f_{Z^2} de Z^2. Pour tout $x \in \mathbb{R}$, on trouve que :

$$F_{Z^2}(x) = P([Z^2 \leq x]).$$

Si $x < 0$, alors on voit que $[Z^2 \leq x] = \emptyset$, et donc $F_{Z^2}(x) = 0$. Si maintenant $x > 0$, alors on a :

$$F_{Z^2}(x) = P([Z^2 \leq x]) = P([-\sqrt{x} \leq Z \leq \sqrt{x}]) = F_Z(\sqrt{x}) - F_Z(-\sqrt{x}).$$

D'après la question (1)(b), on obtient par dérivation que, pour tout $x > 0$:

$$\begin{aligned} f_{Z^2}(x) &= \frac{1}{2\sqrt{x}}\left[f_Z(\sqrt{x}) + f_Z(-\sqrt{x})\right] \\ &= \frac{1}{2\sqrt{x}}\left[2\varphi(\sqrt{x})\Phi(\sqrt{x}) + 2\varphi(-\sqrt{x})\Phi(-\sqrt{x})\right] \\ &= \frac{1}{2\sqrt{x}}\left[2 \times \frac{1}{\sqrt{2\pi}} \times e^{-\frac{(\sqrt{x})^2}{2}}\Phi(\sqrt{x}) + 2 \times \frac{1}{\sqrt{2\pi}} \times e^{-\frac{(-\sqrt{x})^2}{2}}\Phi(-\sqrt{x})\right] \\ &= \frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}\left[\Phi(\sqrt{x}) + \Phi(-\sqrt{x})\right]. \end{aligned}$$

Comme la densité φ de la loi normale centrée réduite est une fonction paire, on voit que la fonction $\Phi - \Phi(0)$ est impaire, et donc :

$$\Phi(-\sqrt{x}) - \Phi(0) = -[\Phi(\sqrt{x}) - \Phi(0)],$$

ce qui entraine après calculs que, pour tout $x > 0$:

$$\Phi(-\sqrt{x}) + \Phi(\sqrt{x}) = 2\Phi(0) = 2 \times \frac{1}{2} = 1.$$

Dès lors, il s'ensuit qu'une densité de Z^2 est définie pour tout $x \in \mathbb{R}$ par :

$$\boxed{f_{Z^2}(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}} & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{cases}.}$$

Comme X^2 et Z^2 ont même densité, on en déduit que :

$$\boxed{X^2 \text{ et } Z^2 \text{ suivent la même loi.}}$$

(b) Calculons $E(Z^2)$, puis donnons la valeur de la variance de Z. Comme X^2 et Z^2 suivent la même loi, elles ont même espérance, et donc $E(Z^2) = E(X^2)$. Mais comme X est une variable normale centrée réduite, on sait que $E(X) = 0$ et $V(X) = 1$. Dès lors, d'après la formule de Koenig-Huygens, on trouve que :

$$E(X^2) = V(X) + E(X)^2 = 1 + 0^2 = 1.$$

Par conséquent, on en déduit que :

$$\boxed{E(Z^2) = 1.}$$

Toujours d'après la formule de Koenig-Huygens et d'après la question (2)(f), on trouve que :

$$V(Z) = E(Z^2) - E(Z)^2 = 1 - \left(\frac{1}{\sqrt{\pi}}\right)^2.$$

Par conséquent, on en déduit que :

$$\boxed{V(Z) = 1 - \frac{1}{\pi}.}$$

(4) Ecrivons une fonction en Python qui, étant donné un entier $n \geq 1$, réalise et affiche n simulations de la variable aléatoire Z. Pour ce faire, on pourra utiliser la commande `rd.normal`, et ce comme suit :

```python
import numpy as np
import numpy.random as rd

def simulz(n):
    x=rd.normal(0,1,n)
    y=rd.normal(0,1,n)
    z=np.zeros(n)
    for i in range(n):
        u=np.array([x[i],y[i]])
        z[i]=np.max(u)
    return z
```

Corrigé de l'exercice 2. Soit a un réel > 0. On considère deux variables aléatoires réelles X et Y définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et suivant toutes deux la loi uniforme sur $[0, a[$. On pose $Z = |X - Y|$, et on admet que $-Y, X - Y, Z$ sont des variables à densité définies sur $(\Omega, \mathcal{A}, P)$.

(1) (a) Déterminons une densité de $-Y$. Pour ce faire, on commence par calculer la fonction de répartition F_{-Y} de $-Y$. Pour tout $x \in \mathbb{R}$, on a :

$$F_{-Y}(x) = P([-Y \leq x]) = P([Y \geq -x]) = 1 - P([Y < -x]).$$

Comme Y est une variable à densité, on sait que $P([Y < -x]) = P([Y \leq -x])$, et donc :

$$F_{-Y}(x) = 1 - P([Y \leq -x]) = 1 - F_Y(-x).$$

Par dérivation, on obtient que $f_{-Y}(x) = f_Y(-x)$ pour tout $x \neq 0, -a$. Mais comme Y suit la loi uniforme sur $[0, a[$, on en déduit qu'une densité de $-Y$ est donnée par :

$$\boxed{f_{-Y}(x) = \begin{cases} \frac{1}{a} & \text{si} \quad x \in [-a, 0] \\ 0 & \text{si} \quad x \notin [-a, 0] \end{cases}.}$$

(b) Montrons que $X - Y$ admet pour densité la fonction g définie pour tout $x \in \mathbb{R}$ par :

$$g(x) = \begin{cases} \dfrac{a - |x|}{a^2} & \text{si} \quad x \in [-a, a] \\ 0 & \text{si} \quad x \notin [-a, a] \end{cases}.$$

Comme X et Y sont indépendantes, X et $-Y$ le sont aussi d'après le lemme des coalitions. Dès lors, comme X et $-Y$ sont des variables à densité, de densités bornées sur $\mathbb{R}$, on sait d'après le cours que leur somme $X - Y$ admet une densité g définie pour tout $x \in \mathbb{R}$ par le produit de convolution suivant :

$$g(x) = \int_{-\infty}^{+\infty} f_X(t) f_{-Y}(x-t)dt.$$

Comme X suit la loi uniforme sur $[0, a[$, on voit que $f_X(t) = \frac{1}{a}$ si $t \in [0, a[$ et que $f_X(t) = 0$ si $t \notin [0, a[$, ce qui entraine que :

$$g(x) = \int_0^a f_X(t) f_{-Y}(x-t)dt = \frac{1}{a}\int_0^a f_{-Y}(x-t)dt.$$

Posons alors $u = x - t$. Alors u est une bijection strictement décroissante de $\mathbb{R}$ dans $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$. De plus, on voit que $u = x$ si $t = 0$, que $u = x - a$ si $t = a$ et que $du = -dt$. Par changement de variables, on obtient que :

$$g(x) = \frac{1}{a}\int_x^{x-a} -f_{-Y}(u)du = \frac{1}{a}\int_{x-a}^{x} f_{-Y}(u)du.$$

Supposons d'abord que $x < -a$. Alors on voit que $x - a < -a$ et que $f_{-Y}(u) = 0$ pour tout $u \in [x-a, x]$, ce qui entraine que :

$$g(x) = \frac{1}{a}\int_{x-a}^{x} 0du = 0.$$

Supposons maintenant que $x \in [-a, 0]$. Alors on voit que $-a \leq x \leq 0$, que $-2a \leq x - a \leq -a$, que $f_{-Y}(u) = \frac{1}{a}$ pour tout $u \in [-a, x]$ et que $f_{-Y}(u) = 0$ pour tout $u \notin [-a, x]$, et donc :

$$g(x) = \frac{1}{a}\int_{-a}^{x} \frac{1}{a}du = \frac{1}{a^2}[x-(-a)] = \frac{1}{a^2}[a-(-x)] = \frac{a-|x|}{a^2}.$$

A présent, supposons que $x \in [0, a]$. Alors on voit que $0 \leq x \leq a$, que $-a \leq x - a \leq 0$, que $f_{-Y}(u) = \frac{1}{a}$ pour tout $u \in [-a, 0]$ et que $f_{-Y}(u) = 0$ pour tout $u \notin [-a, 0]$, et donc :

$$g(x) = \frac{1}{a}\int_{x-a}^{0} \frac{1}{a}du = \frac{1}{a^2}[0-(x-a)] = \frac{1}{a^2}[a-x] = \frac{a-|x|}{a^2}.$$

Enfin, supposons que $x > a$. Alors on voit que $x - a > 0$ et que $f_{-Y}(u) = 0$ pour tout $u \in [x-a, x]$, ce qui entraine que :

$$g(x) = \frac{1}{a}\int_{x-a}^{x} 0du = 0.$$

Par conséquent, $X - Y$ admet pour densité la fonction g définie pour tout $x \in \mathbb{R}$ par :

$$\boxed{g(x) = \begin{cases} \dfrac{a-|x|}{a^2} & \text{si} \quad x \in [-a, a] \\ 0 & \text{si} \quad x \notin [-a, a] \end{cases}.}$$

(2) On désigne par G la fonction de répartition de $X - Y$.

(a) Exprimons la fonction de répartition H de Z en fonction de G. Pour tout $x \in \mathbb{R}$, on a :

$$H(x) = P([Z \leq x]) = P([|X-Y| \leq x]).$$

Si $x < 0$, alors on voit que $[|X-Y| \leq x] = \emptyset$, et donc $H(x) = 0$. Si maintenant $x \geq 0$, alors :

$$H(x) = P([|X-Y| \leq x]) = P([-x \leq X - Y \leq x]) = G(x) - G(-x).$$

Par conséquent, on en déduit que :

$$\boxed{H(x) = \begin{cases} G(x) - G(-x) & \text{si} \quad x \geq 0 \\ 0 & \text{si} \quad x < 0 \end{cases}.}$$

(b) Montrons qu'une densité de Z est donnée par la fonction h définie pour tout $x \in \mathbb{R}$ par :

$$h(x) = \begin{cases} \dfrac{2(a-x)}{a^2} & \text{si} \quad x \in [0, a] \\ 0 & \text{si} \quad x \notin [0, a] \end{cases}.$$

Pour ce faire, il suffit de remarquer que la fonction de répartition H de Z est dérivable sur $\mathbb{R} \setminus \{0, -a, a\}$ comme différence de fonctions dérivables. De plus, comme $Z = |X - Y|$ et que X, Y sont des variables uniformes sur $[0, a[$, on voit que $0 \leq X \leq a$ et $0 \leq Y \leq a$, de sorte que $-a \leq X - Y \leq a$, et donc $0 \leq Z \leq a$. En particulier, le support de Z est contenu dans $[0, a]$, et donc $H(x) = 0$ pour tout $x < -a$ et $H(x) = 1$ pour tout $x > a$. Par dérivation, il s'ensuit que $H'(x) = 0$ pour tout $x \notin [0, a]$. De plus, pour tout $x \in]0, a[$, on a :

$$H'(x) = G'(x) + G'(-x) = \frac{a - |x|}{a^2} + \frac{a - |-x|}{a^2} = \frac{2(a-x)}{a^2}.$$

Par conséquent, on en déduit que Z admet pour densité :

$$\boxed{h(x) = \begin{cases} \dfrac{2(a-x)}{a^2} & \text{si} \quad x \in [0, a] \\ 0 & \text{si} \quad x \notin [0, a] \end{cases}.}$$

(3) Montrons que Z admet une espérance et une variance et calculons-les. Par définition, on sait que Z admet une variance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t^2 h(t)dt$ converge absolument. Mais comme h est nulle en dehors de $[0, a]$, il s'ensuit que Z admet une variance si et seulement si l'intégrale $\int_0^a t^2 h(t)dt$ converge absolument, c'est-à-dire converge (vu que $th(t) \geq 0$ pour tout $t \in [0, a]$). Comme la fonction h est continue sur $[0, a]$, il s'ensuit que $t \longmapsto t^2 h(t)$ est continue sur $[0, a]$, et donc l'intégrale $\int_0^a t^2 h(t)dt$ est faussement impropre. En particulier, elle converge, ce qui entraine que Z admet une variance (et donc aussi une espérance). Par conséquent :

$$\boxed{Z \text{ admet une espérance et une variance.}}$$

A présent, calculons $E(Z)$. D'après les questions précédentes, on a :

$$E(Z) = \int_0^a th(t)dt = \int_0^a \frac{2t(a-t)}{a^2}dt = \int_0^a \frac{2at - 2t^2}{a^2}dt.$$

Par des calculs simples, on trouve que :

$$E(Z) = \int_0^a \frac{2at - 2t^2}{a^2}dt = \left[\frac{t^2}{a} - \frac{2t^3}{3a^2}\right]_0^a = \frac{a^2}{a} - \frac{2a^3}{3a^2} - 0.$$

Par conséquent, on en déduit après simplification que :

$$\boxed{E(Z) = \frac{a}{3}.}$$

Enfin, calculons $V(Z)$. D'après les questions précédentes, on a :

$$E(Z^2) = \int_0^a t^2 h(t)dt = \int_0^a \frac{2t^2(a-t)}{a^2}dt = \int_0^a \frac{2at^2 - 2t^3}{a^2}dt.$$

Par des calculs simples, on trouve que :

$$E(Z^2) = \int_0^a \frac{2at^2 - 2t^3}{a^2}dt = \left[\frac{2t^3}{3a} - \frac{2t^4}{4a^2}\right]_0^a = \frac{2a^2}{3} - \frac{a^2}{2} = \frac{a^2}{6}.$$

D'après la formule de Koenig-Huygens, on obtient que :

$$V(Z) = E(Z^2) - E(Z)^2 = \frac{a^2}{6} - \left(\frac{a}{3}\right)^2 = \frac{a^2}{6} - \frac{a^2}{9}.$$

Par conséquent, on en déduit après simplification que :

$$\boxed{V(Z) = \frac{a^2}{18}.}$$

(4) Ecrivons une fonction en Python qui, étant donnés un réel $a > 0$ et deux entiers $n, m \geq 1$, réalise et affiche n simulations de la variable aléatoire Z, puis affiche l'histogramme correspondant pour une subdivision de l'intervalle $[0, 2a]$ en m classes de même amplitude. Pour ce faire, on utilise la commande `rd.random` pour simuler des variables uniformes, ainsi que la commande `plt.hist` pour construire l'histogramme, et ce comme suit :

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

def simul2(a,n,m):
    z=np.zeros(n)
    for i in range(n):
        z[i]=(a*rd.random())+(a*rd.random())
    print(z)
    c=np.linspace(0,(2*a),m)
    plt.hist(x,c)
    plt.show()
```

Corrigé de l'exercice 3. Soit $n \in \mathbb{N}$ et soit $E = \mathbb{R}_{2n+1}[x]$. Pour tout $k \in \{0, ..., 2n+1\}$, on admet que l'expression $x \longmapsto x^{2n+1} \times \frac{1}{x^k}$ désigne le polynôme $x \longmapsto x^{2n+1-k}$. On désigne par Id l'endomorphisme identique de E et par f l'application qui, à tout élément P de E, associe le polynôme $f(P) : x \longmapsto x^{2n+1} P\left(\frac{1}{x}\right)$.

(1) Montrons que f est un endomorphisme de E. Tout d'abord, on peut remarquer que, pour tout polynôme $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k \in E$ et pour tout $x \in \mathbb{R}^*$, on a :

$$f(P)(x) = x^{2n+1} \sum_{k=0}^{2n+1} a_k \left(\frac{1}{x}\right)^k = \sum_{k=0}^{2n+1} a_k x^{2n+1-k}.$$

Si l'on effectue le changement d'indices $l = 2n+1-k$, alors on trouve que, pour tout $x \in \mathbb{R}$:

$$f(P)(x) = \sum_{l=0}^{2n+1} a_{2n+1-l} x^l.$$

En particulier, on voit que $f(P)$ appartient à E, et donc f est une application de E dans E. Pour montrer que f est un endomorphisme de E, il reste à vérifier que f est linéaire. Soient $P, Q \in E$ et soient $\lambda, \mu \in \mathbb{R}$. Pour tout $x \in \mathbb{R}^*$, on trouve que :

$$f(\lambda P+\mu Q)(x) = x^{2n+1}(\lambda P+\mu Q)\left(\frac{1}{x}\right) = \lambda \left[x^{2n+1} P\left(\frac{1}{x}\right)\right] + \mu \left[x^{2n+1} Q\left(\frac{1}{x}\right)\right] = \lambda f(P) + \mu f(Q)(x),$$

d'où il s'ensuit que $f(\lambda P + \mu Q) = \lambda f(P) + \mu f(Q)$, et donc f est linéaire. Par conséquent :

$$\boxed{f \text{ est un endomorphisme de } E.}$$

(2) (a) Vérifions que $f \circ f = \mathrm{Id}$. Pour tout $P \in E$, on trouve que :

$$f \circ f(P) = f(f(P)) = f\left(X^{2n+1} P\left(\frac{1}{X}\right)\right) = X^{2n+1} \times \left(\frac{1}{X}\right)^{2n+1} \times P\left(\frac{1}{\frac{1}{X}}\right) = P(X).$$

Comme ceci est vrai pour tout $P \in E$, on en déduit que :

$$\boxed{f \circ f = \mathrm{Id}.}$$

(b) Comme $f \circ f = \mathrm{Id}$, le polynôme $X^2 - 1 = (X-1)(X+1)$ est annulateur de f, et donc :

$$\boxed{1 \text{ et } -1 \text{ sont les deux valeurs propres possibles de } f.}$$

(3) Soit $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k$ un élément quelconque de $\ker(f - \mathrm{Id})$.

(a) Montrons que les a_k ($0 \leq k \leq 2n+1$) sont solutions du système : $\forall k \in \{0, ..., n\}$, $a_k = a_{2n+1-k}$. Comme P appartient à $\ker(f - \mathrm{Id})$, on voit que $f(P) - P = 0$ et $f(P) = P$, ce qui entraine avec les calculs de la question (1) que, pour tout $x \in \mathbb{R}$:

$$f(P)(x) = \sum_{l=0}^{2n+1} a_{2n+1-l} x^l = P(x) = \sum_{k=0}^{2n+1} a_k x^k.$$

Comme deux polynômes sont égaux si et seulement si les coefficients de leurs écritures réduites sont égaux, on obtient que $a_{2n+1-k} = a_k$ pour tout $k \in \{0, ..., 2n+1\}$. En particulier, on a :

$$\boxed{\forall k \in \{0, ..., n\}, \quad a_k = a_{2n+1-k}.}$$

(b) Déterminons une base de $\ker(f-\mathrm{Id})$. Pour ce faire, considérons un élément $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k$ de $\ker(f-\mathrm{Id})$, que l'on commence par écrire sous la forme suivante, pour tout $x \in \mathbb{R}$:

$$P = \sum_{k=0}^{n} a_k x^k + \sum_{k=n+1}^{2n+1} a_k x^k.$$

Si l'on effectue le changement d'indices $l = 2n+1-k$ dans la deuxième somme de droite, alors on trouve que, pour tout $x \in \mathbb{R}$:

$$P(x) = \sum_{k=0}^{n} a_k x^k + \sum_{l=0}^{n} a_{2n+1-l} x^{2n+1-l} = \sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} a_{2n+1-k} x^{2n+1-k}.$$

Comme P appartient à $\ker(f-\mathrm{Id})$, on sait d'après la question précédente que $a_k = a_{2n+1-k}$ pour tout $k \in \{0, ..., n\}$, ce qui nous donne que, pour tout $x \in \mathbb{R}$:

$$P(x) = \sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} a_k x^{2n+1-k} = \sum_{k=0}^{n} a_k \left(x^k + x^{2n+1-k}\right).$$

Si l'on pose $P_k : x \longmapsto x^k + x^{2n+1-k}$ pour tout $k \in \{0, ..., n\}$, alors on voit que tout élément P de $\ker(f-\mathrm{Id})$ est combinaison linéaire des P_k, et donc :

$$\ker(f-\mathrm{Id}) \subset \mathrm{Vect}(P_0, ..., P_n).$$

Par ailleurs, on trouve que, pour tout $k \in \{0, ..., n\}$ et pour tout $x \in \mathbb{R}^*$:

$$f(P_k)(x) = x^{2n+1}\left[\left(\frac{1}{x}\right)^k + \left(\frac{1}{x}\right)^{2n+1-k}\right] = x^{2n+1-k} + x^k = P_k(x),$$

ce qui entraine que P_k appartient à $\ker(f-\mathrm{Id})$ pour tout $k \in \{0, ..., n\}$, et donc :

$$\ker(f-\mathrm{Id}) = \mathrm{Vect}(P_0, ..., P_n).$$

En particulier, la famille $(P_0, ..., P_n)$ est génératrice dans $\ker(f-\mathrm{Id})$. Mais comme P_k est de degré $2n+1-k$ pour tout $k \in \{0, ..., n\}$, la famille $(P_0, ..., P_n)$ est constituée de polynômes de degrés deux à deux distincts, et donc elle est libre. Par conséquent :

$$\boxed{(x \mapsto 1 + x^{2n+1}, ..., x \mapsto x^k + x^{2n+1-k}, ..., x \mapsto x^n + x^{n+1}) \text{ est une base de } \ker(f-\mathrm{Id}).}$$

(4) En procédant exactement comme à la question (3), on montre aussi que :

$$\boxed{(x \mapsto 1 - x^{2n+1}, ..., x \mapsto x^k - x^{2n+1-k}, ..., x \mapsto x^n - x^{n+1}) \text{ est une base de } \ker(f+\mathrm{Id}).}$$

(5) Pour tout $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k \in E$ et tout $Q : x \longmapsto \sum_{k=0}^{2n+1} b_k x^k \in E$, on pose :

$$\varphi(P, Q) = \sum_{k=0}^{2n+1} a_k b_k.$$

(a) Montrons que φ est un produit scalaire sur E. Pour ce faire, on va montrer que φ est une forme bilinéaire symétrique définie positive, et ce en plusieurs étapes :

Première étape : φ est symétrique.

En effet, pour tout $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k \in E$ et tout $Q : x \longmapsto \sum_{k=0}^{2n+1} b_k x^k \in E$, on a :

$$\varphi(P, Q) = \sum_{k=0}^{2n+1} a_k b_k = \sum_{k=0}^{2n+1} b_k a_k = \varphi(Q, P),$$

d'où il s'ensuit que φ est symétrique.

Deuxième étape : φ est bilinéaire.

En effet, pour tous éléments $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k$, $Q : x \longmapsto \sum_{k=0}^{2n+1} b_k x^k$, $R : x \longmapsto \sum_{k=0}^{2n+1} c_k x^k$ de E et pour tous $\lambda, \mu \in \mathbb{R}$, on trouve par linéarité de la somme que :

$$\varphi(\lambda P + \mu Q, R) = \sum_{k=0}^{2n+1} (\lambda a_k + \mu b_k) c_k = \lambda \sum_{k=0}^{2n+1} a_k c_k + \mu \sum_{k=0}^{2n+1} b_k c_k = \lambda \varphi(P, R) + \mu \varphi(Q, R),$$

ce qui entraine que φ est linéaire à gauche, et donc bilinéaire par symétrie.

Troisième étape : φ est définie positive.

En effet, pour tout élément $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k$ de E, on a :

$$\varphi(P,P) = \sum_{k=0}^{2n+1} a_k^2 \geq 0,$$

et donc φ est positive. De plus, si $\varphi(P,P) = 0$, alors on voit que $a_k = 0$ pour tout $k \in \{0, ..., 2n+1\}$, et donc $P = 0$. En particulier, la forme bilinéaire φ est définie positive.

Par conséquent, on en déduit que :

$$\boxed{\varphi \text{ est un produit scalaire sur } E.}$$

(b) Etablissons que f est un endomorphisme symétrique de E. Soient P, Q deux éléments de E, de la forme $P : x \longmapsto \sum_{k=0}^{2n+1} a_k x^k$ et $Q : x \longmapsto \sum_{k=0}^{2n+1} b_k x^k$. D'après la question (1), on sait que, pour tout $x \in \mathbb{R}$:

$$f(P)(x) = \sum_{k=0}^{2n+1} a_{2n+1-k} x^k \quad \text{et} \quad f(Q)(x) = \sum_{k=0}^{2n+1} b_{2n+1-k} x^k.$$

Dès lors, on obtient par définition du produit scalaire que :

$$\varphi(f(P), Q) = \sum_{k=0}^{2n+1} a_{2n+1-k} b_k.$$

Si l'on effectue le changement d'indices $l = 2n+1-k$ dans la somme de droite, alors on a :

$$\varphi(f(P), Q) = \sum_{l=0}^{2n+1} a_l b_{2n+1-l} = \sum_{k=0}^{2n+1} b_{2n+1-k} a_k = \varphi(f(Q), P) = \varphi(P, f(Q)).$$

Mais comme ceci est vrai pour tous $P, Q \in E$, on en déduit que :

$$\boxed{f \text{ est un endomorphisme symétrique de } E.}$$

(c) Montrons que $\ker(f + \mathrm{Id})$ et $\ker(f - \mathrm{Id})$ sont supplémentaires orthogonaux. Comme f est un endomorphisme symétrique, l'espace vectoriel E est somme directe des sous-espaces propres de f, qui sont de plus deux à deux orthogonaux. Mais comme les seules valeurs de f sont 1 et -1 d'après les questions précédentes, il s'ensuit que $E_{-1}(f) = \ker(f + \mathrm{Id})$ et $E_1(f) = \ker(f - \mathrm{Id})$ sont orthogonaux et que de plus $E = E_{-1}(f) \oplus E_1(f)$. En particulier :

$$\boxed{\ker(f + \mathrm{Id}) \text{ et } \ker(f - \mathrm{Id}) \text{ sont supplémentaires orthogonaux dans } E.}$$

Corrigé du problème 1. On considère l'espace euclidien $\mathbb{R}^3$ muni du produit scalaire canonique, lequel est défini pour tout $u = (x, y, z) \in \mathbb{R}^3$ et tout $u' = (x', y', z') \in \mathbb{R}^3$ par $\langle u, u' \rangle = xx' + yy' + zz'$. La norme du vecteur u est définie par $\|u\| = \sqrt{\langle u, u \rangle}$. On désigne par $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$, et on rappelle que cette base est orthonormée pour le produit scalaire $\langle , \rangle$. Le but de ce problème est de montrer que l'on peut trouver une famille $\mathcal{F} = (e_1, ..., e_n)$ de cardinal maximal, formée de n vecteurs unitaires deux à deux distincts de $\mathbb{R}^3$, ainsi qu'un réel α tels que, pour tout couple d'entiers (i, j) vérifiant $1 \leq i < j \leq n$, on ait : $\langle u_i, u_j \rangle = \alpha$. La partie 1 permet d'obtenir un résultat d'algèbre linéaire utile pour la suite, la partie 2 étudie les propriétés d'une telle famille et la partie 3 propose la construction d'une famille solution du problème pour $n = 4$ (cette valeur est d'ailleurs la valeur maximale possible de n mais ce résultat ne sera pas démontré ici).

Partie 1 : Soit n est un entier ≥ 2. Pour tout $a \in \mathbb{R}$, on désigne par M_a la matrice de $\mathcal{M}_n(\mathbb{R})$ dont les éléments diagonaux sont tous égaux à 1, les autres étant égaux à a. On note I la matrice identité de $\mathcal{M}_n(\mathbb{R})$ et J la matrice de $\mathcal{M}_n(\mathbb{R})$ dont tous les coefficients valent 1.

(1) (a) Comme la matrice J a tous ses coefficients égaux à 1, elle est symétrique réelle. Dès lors, d'après le théorème spectral, il s'ensuit que :

$$\boxed{J \text{ est diagonalisable.}}$$

(b) Calculons J^2 et déterminons les deux valeurs propres de J. Par des calculs simples, on a :

$$J^2 = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} = \begin{pmatrix} n & \dots & n \\ \vdots & & \vdots \\ n & \dots & n \end{pmatrix}.$$

Par conséquent, on en déduit que :

$$\boxed{J^2 = nJ.}$$

Dès lors, comme le polynôme $P : x \longmapsto x^2 - nx$ est annulateur de J, on voit que 0 et n sont les seules valeurs propres possibles de J. Reste à vérifier que ce sont bien des valeurs propres de J. Par des calculs simples, on trouve que :

$$\begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \neq 0,$$

et donc N est bien valeur propre de J. De la même manière, on obtient que :

$$\begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \neq 0,$$

et donc 0 est aussi valeur propre de J. Par conséquent :

$$\boxed{0 \text{ et } n \text{ sont les seules valeurs propres de } J.}$$

(2) (a) Calculons les valeurs propres de M_a. Pour ce faire, on va déterminer une base de vecteurs propres de J. Comme 0 et n sont les seules valeurs propres de J, on commence par calculer une base de $E_0(J)$. Soit X un vecteur colonne de composantes $x_1, ..., x_n$. Alors :

$$X \in E_0(J) \iff JX = 0X \iff \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

En termes de coordonnées, cela nous donne que $(x_1, ..., x_n)$ est solution du système linéaire :

$$\left\{ \begin{array}{ccccccccc} x_1 & + & x_2 & + & \dots & + & x_n & = & 0 \end{array} \right. .$$

Si l'on choisit $x_2, ..., x_n$ comme paramètres, alors on trouve que $x_1 = -x_2 - ... - x_n$, et donc :

$$X \in E_0(J) \iff \exists x_2, ..., x_n \in \mathbb{R}, \ X = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + ... + x_n \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$\iff X \in \operatorname{Vect}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right).$$

Pour tout $k \in \{2, ..., n\}$, on désigne par E_k le vecteur de composantes $-1, 0, ...0, 1, 0, ..., 0$ (où 1 est placé en k-ème position). D'après ce qui précède, on voit que :

$$E_0(J) = \operatorname{Vect}(E_2, ..., E_n).$$

Dès lors, il s'ensuit que $(E_2, ..., E_n)$ est une famille génératrice de $E_0(J)$. Comme ces vecteurs forment un système réduit (pour la méthode du pivot de Gauss), on voit que $\operatorname{rg}(E_2, ..., E_n) = n-1$, et donc $E_0(J)$ est de dimension $(n-1)$. Mais comme $(E_2, ..., E_n)$ est une famille génératrice à $(n-1)$ éléments d'un espace vectoriel de dimension $(n-1)$, il s'ensuit que :

$$\boxed{(E_2, ..., E_n) \text{ est une base de } E_0(J).}$$

A présent, calculons une base de $E_n(J)$. Comme n est une valeur propre de J, on voit que $\dim E_n(J) \geq 1$. De plus, comme $E_n(J)$ et $E_0(J)$ sont en somme directe, on trouve que :

$$\dim E_n(J) \leq n - \dim E_0(J) = 1,$$

et donc $E_n(J)$ est de dimension 1. En outre, si l'on désigne par E_1 le vecteur de composantes $1, ..., 1$, alors on voit d'après la question précédente que $JE_1 = E_1$, et donc E_1 appartient à $E_n(J)$. Comme E_1 est non nul, il constitue une famille libre de $E_n(J)$, et donc une base de $E_n(J)$ car $\dim E_n(J) = 1$. En particulier :

$$\boxed{(E_1) \text{ est une base de } E_n(J).}$$

En résumé, on obtient que :

$$\boxed{(E_1, E_2, ..., E_n) \text{ est une base de vecteurs propres de } J.}$$

Passons maintenant au calcul des valeurs propres de M_a. Tout d'abord, on voit que :

$$M_a E_1 = \begin{pmatrix} 1 & a & \dots & a \\ a & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ a & \dots & a & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + (n-1)a \\ \vdots \\ 1 + (n-1)a \end{pmatrix} = [1 + (n-1)a] \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

d'où il s'ensuit que $1 + (n-1)a$ est valeur propre de M_a. De la même façon, on trouve que, pour tout $k \in \{2, ..., n\}$:

$$M_a E_k = \begin{pmatrix} 1 & a & \dots & a \\ a & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ a & \dots & a & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -1 + a \\ 0 \\ \vdots \\ 0 \\ a - 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (a-1) \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

d'où il s'ensuit que $a - 1$ est aussi valeur propre de M_a. Mais comme la famille $(E_1, ..., E_n)$ est une base de vecteurs propres de M_a, on en déduit que $1 + (n-1)a$ et $a - 1$ sont les seules valeurs propres de M_a. Par conséquent :

$$\boxed{a - 1 \text{ et } 1 + (n-1)a \text{ sont les seules valeurs propres de } M_a.}$$

(b) Montrons que M_a est inversible si et seulement si : $a \neq 1$ et $a \neq -\frac{1}{n-1}$. Par définition, la matrice M_a est inversible si et seulement si 0 n'est pas valeur propre de M_a. Mais comme $a - 1$ et $1 + (n-1)a$ sont les seules valeurs propres de M_a, il s'ensuit que M_a est inversible si et seulement si $a - 1 \neq 0$ et $1 + (n-1)a \neq 0$, et donc :

$$\boxed{M_a \text{ est inversible si et seulement si : } a \neq 1 \text{ et } a \neq -\frac{1}{n-1}.}$$

(3) Ecrivons une fonction en Python qui, étant donnés un réel a et un entier $n \geq 2$, affiche la matrice M_a puis détermine si cette matrice est inversible ou pas. Pour ce faire, on part du fait que la matrice M_a se décompose sous la forme $M_a = aJ + (1-a)I$, et on utilise les questions précédentes comme suit :

```
import numpy as np

def matrice(a,n):
    m=np.zeros([n,n])
    for i in range(n):
        for j in range(n):
            if i==j:
                m[i,j]=1
            else:
                m[i,j]=a
    print(m)
    if a==1 or a==-1/(n-1):
        print('la matrice n est pas inversible')
    else:
        print('la matrice est inversible')
```

Partie 2 : On suppose que l'on a trouvé une famille $(u_1, ..., u_n)$ formée de n vecteurs unitaires et deux à deux distincts de $\mathbb{R}^3$ et un réel α solutions du problème.

(1) Soient $\lambda_1, ..., \lambda_n$ des réels tels que $\sum_{k=1}^n \lambda_k u_k = 0$. Montrons que :

$$M_\alpha \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0.$$

Pour tout $l \in \{1, ..., n\}$, on obtient par bilinéarité du produit scalaire que :

$$\left\langle \sum_{k=1}^n \lambda_k u_k, u_l \right\rangle = \sum_{k=1}^n \lambda_k \langle u_k, u_l \rangle = 0.$$

Comme les u_i sont des vecteurs unitaires, on voit que $\langle u_l, u_l \rangle = 1$ pour tout l. De plus, comme $\langle u_i, u_j \rangle = \alpha$ pour tout couple d'entiers (i, j) tels que $1 \leq i < j \leq n$, et que le produit scalaire est symétrique, on trouve que $\langle u_k, u_l \rangle = \alpha$ si $k \neq l$. Dès lors, il s'ensuit que, pour tout $l \in \{1, ..., n\}$:

$$\sum_{k=1}^n \lambda_k \langle u_k, u_l \rangle = \alpha\lambda_1 + ... + \alpha\lambda_{l-1} + \lambda_l + \alpha\lambda_{l+1} + ... + \alpha\lambda_n = 0.$$

Comme M_α est la matrice de $\mathcal{M}_n(\mathbb{R})$ dont les éléments diagonaux sont tous égaux à 1, les autres étant égaux à α, l'égalité ci-dessus n'est ni plus ni moins que la l-ème composante du vecteur :

$$U = M_\alpha \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Par conséquent, on en déduit que :

$$\boxed{M_\alpha \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0.}$$

(2) Déterminons la valeur maximale de n lorsque $\alpha \neq 1$ et $\neq -\frac{1}{n-1}$. Dans ces conditions, on sait d'après la question $(2)(b)$ de la partie 1 que la matrice M_α est inversible. Dès lors, d'après la question précédente, on obtient que :

$$\sum_{k=1}^n \lambda_k u_k = 0 \quad \Longrightarrow \quad M_\alpha \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0 \quad \Longrightarrow \quad \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0 \quad \Longrightarrow \quad \lambda_1 = ... = \lambda_n = 0.$$

En particulier, la famille $(u_1, ..., u_n)$ est libre. Mais comme il s'agit d'une famille de $\mathbb{R}^3$, elle doit compter au plus 3 éléments, et donc la valeur maximale de n est donnée par :

$$\boxed{n = 3.}$$

(3) Etude du cas $\alpha = 1$.

(a) Ecrivons l'inégalité de Cauchy-Schwarz pour u_i et u_j avec $i \neq j$. D'après le cours, on voit que :

$$|\langle u_i, u_j\rangle| \leq \|u_i\| \times \|u_j\|.$$

Comme les u_i sont unitaires et que $\langle u_i, u_j\rangle = \alpha = 1$ pour tous i, j avec $i \neq j$, on obtient que :

$$\boxed{|\langle u_i, u_j\rangle| = \|u_i\| \times \|u_j\|.}$$

De plus, on sait d'après le cours qu'il y a égalité dans l'inégalité de Cauchy-Schwarz si et seulement si les vecteurs u_i et u_j sont colinéaires. Dès lors, il existe un réel $a_{i,j}$ tel que $u_i = a_{i,j}u_j$, et donc :

$$\langle u_i, u_j\rangle = \langle a_{i,j}u_j, u_j\rangle = a_{i,j}\|u_j\|^2 = a_{i,j} \times 1 = \alpha = 1.$$

Par conséquent, on en déduit que, si $i \neq j$:

$$\boxed{u_i = u_j.}$$

(b) D'après ce qui précède, tous les vecteurs u_i sont égaux. Mais comme la famille $(u_1, ..., u_n)$ est formée de vecteurs deux à deux distincts, cette famille n'a qu'un seul élément, et donc :

$$\boxed{n = 1.}$$

(4) Dans cette question, on admet qu'il existe une famille (u_1, u_2, u_3, u_4) formée de 4 vecteurs unitaires et deux à deux distincts de $\mathbb{R}^3$ solution du problème.

(a) Donnons la valeur de α. Comme la famille (u_1, u_2, u_3, u_4) comporte 4 éléments de $\mathbb{R}^3$, elle n'est pas libre. Dès lors, d'après la question (1) de la partie 2, on sait que la matrice M_α n'est pas inversible. D'après la question (2)(b) de la partie 1, il s'ensuit que $\alpha = 1$ ou $\alpha = -\frac{1}{3}$. Mais d'après la question (3)(b) de la partie 2, on sait que, si $\alpha = 1$, alors $n = 1$, ce qui est impossible car $n = 4$ par hypothèse. Par conséquent :

$$\boxed{\alpha = -\frac{1}{3}.}$$

(b) Montrons que (u_1, u_2, u_3) est une base de $\mathbb{R}^3$. Comme cette famille comporte 3 éléments et que $\mathbb{R}^3$ est de dimension 3, il suffit de montrer que la famille (u_1, u_2, u_3) est libre. Soient $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 0$. Alors on voit que $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 + 0u_4 = 0$, ce qui entraine d'après la question (1) de la partie 2 que :

$$M_\alpha \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ 0 \end{pmatrix} = 0.$$

D'après la question (2)(a) de la partie 1 et comme $\alpha = -\frac{1}{3}$, on sait que le noyau de M_α (c'est-à-dire le sous-espace propre $E_0(M_\alpha)$) est engendré par le vecteur colonne E_1, dont toutes les composantes sont égales à 1. Dès lors, il existe un réel θ tel que :

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ 0 \end{pmatrix} = \theta \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

d'où il s'ensuit que $\theta = 0$, et donc $\lambda_1 = \lambda_2 = \lambda_3 = 0$. En particulier, la famille (u_1, u_2, u_3) est libre, et donc :

$$\boxed{(u_1, u_2, u_3) \text{ est une base de } \mathbb{R}^3.}$$

(c) Calculons les coordonnées de u_4 dans cette base. Soient $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ les coordonnées de u_4 dans la base $\mathcal{C} = (u_1, u_2, u_3)$. Par définition, on voit que $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = u_4$, et donc $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 - u_4 = 0$. D'après la question (1) de la partie 2, ceci entraine que :

$$M_\alpha \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ -1 \end{pmatrix} = 0.$$

D'après la question (2)(a) de la partie 1 et comme $\alpha = -\frac{1}{3}$, on sait que le noyau de M_α est engendré par le vecteur colonne E_1. Dès lors, il existe un réel θ tel que :

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ -1 \end{pmatrix} = \theta \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

d'où il s'ensuit que $\theta = -1$, et donc $\lambda_1 = \lambda_2 = \lambda_3 = -1$. En particulier, la matrice colonne des coordonnées de u_4 dans la base $\mathcal{C}$ est donnée par :

$$\boxed{\mathfrak{mat}_{\mathcal{C}}(u_4) = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.}$$

Partie 3 : On se propose de trouver des familles solutions du problème dans certains cas.

(1) Donnons une famille solution du problème posé pour $n = 3$ et $\alpha = 0$. Si $n = 3$ et $\alpha = 0$, il s'agit de trouver une famille de 3 vecteurs deux à deux orthogonaux et tous unitaires, c'est-à-dire une famille orthonormale de $\mathbb{R}^3$ à 3 éléments. On voit alors que la base canonique de $\mathbb{R}^3$ convient. Dès lors :

$$\boxed{\text{la base canonique de } \mathbb{R}^3 \text{ est solution du problème pour } n = 3 \text{ et } \alpha = 0.}$$

(2) On pose $v_1 = e_1$, $v_2 = -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$ et $v_3 = -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2$.

(a) Montrons que (v_1, v_2, v_3) est solution du problème posé pour $\alpha = -\frac{1}{2}$. Par définition, on voit que ces 3 vecteurs sont deux à deux distincts. De plus, ils sont unitaires car :

$$\|v_1\| = 1, \quad \|v_2\| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2} = 1, \quad \|v_3\| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(-\frac{\sqrt{3}}{2}\right)^2} = 1$$

Reste à vérifier que $\langle v_1, v_2\rangle = \langle v_1, v_3\rangle = \langle v_2, v_3\rangle = -\frac{1}{2}$. Par bilinéarité du produit scalaire et vu que (e_1, e_2) est une famille orthonormée, on trouve que :

$$\langle v_1, v_2\rangle = \left\langle e_1, -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2 \right\rangle = -\frac{1}{2}\langle e_1, e_1\rangle + \frac{\sqrt{3}}{2}\langle e_1, e_2\rangle = -\frac{1}{2}.$$

De la même façon, on obtient que :

$$\langle v_1, v_3\rangle = \left\langle e_1, -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2 \right\rangle = -\frac{1}{2}\langle e_1, e_1\rangle - \frac{\sqrt{3}}{2}\langle e_1, e_2\rangle = -\frac{1}{2}.$$

Par des calculs analogues, on trouve que :

$$\begin{aligned} \langle v_2, v_3\rangle &= \left\langle -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2, -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2 \right\rangle \\ &= \left(-\frac{1}{2}\right)^2 \langle e_1, e_1\rangle - \left(\frac{\sqrt{3}}{2}\right)^2 \langle e_2, e_2\rangle \\ &= \frac{1}{4} - \frac{3}{4} \quad = \quad -\frac{1}{2}. \end{aligned}$$

Par conséquent, on en déduit que :

$$\boxed{\text{la famille } (v_1, v_2, v_3) \text{ est solution du problème pour } \alpha = -\frac{1}{2}.}$$

(b) Déterminons deux réels λ, μ tels que la famille $(u_1, u_2, u_3, u_4) = (e_3, \lambda v_1 + \mu e_3, \lambda v_2 + \mu e_3, \lambda v_3 + \mu e_3)$ soit solution du problème posé pour $n = 4$. Par définition, cela signifie que $\alpha = -\frac{1}{3}$, que les vecteurs de cette famille sont deux à deux distincts (et donc $\lambda \neq 0$), et que de plus :

$$\begin{cases} \langle u_1, u_1\rangle = \langle u_2, u_2\rangle = \langle u_3, u_3\rangle = \langle u_4, u_4\rangle = 1 \\ \langle u_1, u_2\rangle = \langle u_1, u_3\rangle = \langle u_1, u_4\rangle = \langle u_2, u_3\rangle = \langle u_2, u_4\rangle = \langle u_3, u_4\rangle = -\dfrac{1}{3} \end{cases}.$$

Après calculs, on trouve que λ, μ sont solutions du système :

$$\left\{\begin{array}{ccccc} \lambda^2 & + & \mu^2 & = & 1 \\ -\frac{1}{2}\lambda^2 & + & \mu^2 & = & -\frac{1}{3} \\ & & \mu & = & -\frac{1}{3} \end{array}\right. ,$$

d'où il s'ensuit après résolution que $\lambda^2 = \frac{8}{9}$, et donc $\lambda = \pm\frac{2\sqrt{2}}{3}$. Par conséquent :

$$\boxed{(u_1, u_2, u_3, u_4) \text{ est solution du problème pour } n = 4,\ \lambda = \pm\frac{2\sqrt{2}}{3} \text{ et } \mu = -\frac{1}{3}.}$$

(3) Ecrivons une fonction en Python qui, étant donnés 3 vecteurs unitaires et distincts u, v, w de $\mathbb{R}^3$, détermine si la famille (u, v, w) est solution du problème ou pas, et affiche la valeur de α dans ce cas. Pour ce faire, on doit vérifier que $\langle u, v\rangle = \langle u, w\rangle = \langle v, w\rangle$, ce qui revient à contrôler que $\langle u, v\rangle = \langle u, w\rangle$ et $\langle u, v\rangle = \langle v, w\rangle$, et dans ce cas $\alpha = \langle u, v\rangle$. Dès lors, on peut procéder comme suit :

```
import numpy as np

def vecteurs(u,v,w):
    x=np.sum(u*v)
    y=np.sum(u*w)
    z=np.sum(v*w)
    if x==y and y==z:
        print('les vecteurs sont solutions du probleme')
        print(x)
    else:
        print('les vecteurs ne sont pas solutions du probleme')
```

2. Sujet type HEC-ESCP Maths I

Corrigé du problème 2. Dans ce problème, on s'intéresse à des opérations de transport dans des situations déterministes ou aléatoires, modélisées de manière discrète ou continue, dans le but de trouver un programme de transport optimal dont le coût serait le plus faible possible. Les parties I, II et III sont largement indépendantes. Toutes les variables aléatoires considérées dans ce problème sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. Sous réserve d'existence, on note $E(Z)$ l'espérance d'une variable aléatoire Z. Enfin, pour tout entier $N \geq 1$, on note $\mathcal{E}_N$ l'ensemble des applications de $[\![1, N]\!]$ dans $[\![1, N]\!]$.

Préliminaire

(1) Soit p un réel vérifiant $0 < p < 1$. On considère une variable aléatoire X suivant la loi exponentielle de paramètre 1. Pour tout $\omega \in \Omega$, on pose $Y(\omega) = \lfloor pX(\omega) \rfloor$, où $\lfloor\ \rfloor$ désigne la fonction partie entière.

(a) Vérifions tout d'abord que Y est une variable aléatoire discrète. Comme X suit une loi exponentielle, on voit que $X(\Omega) = \mathbb{R}_+$, ce qui entraine que $(pX)(\Omega) = \mathbb{R}_+$, et donc $Y(\Omega) \subset \mathbb{N}$. En particulier, l'ensemble $Y(\Omega)$ est au plus dénombrable. De plus, pour tout $n \in \mathbb{N}$, on voit que :

$$[Y = n] = [\lfloor pX \rfloor = n] = [n \leq pX < n+1] = \left[\frac{n}{p} \leq X < \frac{n+1}{p}\right] = \left[X \in \left[\frac{n}{p}, \frac{n+1}{p}\right[\right].$$

Comme X est une variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$ et que $[\frac{n}{p}, \frac{n+1}{p}[$ est un intervalle, il s'ensuit que l'ensemble $\left[X \in [\frac{n}{p}, \frac{n+1}{p}[\right[$ est un élément de $\mathcal{A}$, et donc l'ensemble $[Y = n]$ appartient à $\mathcal{A}$ pour tout $n \in \mathbb{N}$. Par conséquent, on en déduit que :

$$\boxed{Y \text{ est une variable aléatoire discrète.}}$$

A présent, calculons $P([Y = n])$ pour tout $n \in \mathbb{N}$. Pour tout $n \in \mathbb{N}$, on a d'après ce qui précède :

$$P([Y = n]) = P\left(\left[\frac{n}{p} \leqslant X < \frac{n+1}{p}\right]\right).$$

Comme X est une variable à densité, ceci nous donne que, pour tout $n \in \mathbb{N}$:

$$P([Y = n]) = F_X\left(\frac{n+1}{p}\right) - F_X\left(\frac{n}{p}\right).$$

Comme X suit la loi exponentielle de paramètre 1, il s'ensuit que, pour tout $n \in \mathbb{N}$:

$$P([Y = n]) = \left(1 - e^{-\frac{n+1}{p}}\right) - \left(1 - e^{-\frac{n}{p}}\right) = e^{-\frac{n}{p}} - e^{-\frac{n+1}{p}} = e^{-\frac{n}{p}}\left(1 - e^{-\frac{1}{p}}\right).$$

Par conséquent, on en déduit que, pour tout $n \in \mathbb{N}$:

$$\boxed{P([Y = n]) = e^{\frac{n}{p}}\left(1 - e^{-\frac{1}{p}}\right).}$$

(b) Montrons que la variable aléatoire $Y + 1$ suit une loi géométrique dont on précisera le paramètre. Pour ce faire, on pose $Z = Y + 1$. Comme $Y(\Omega) \subset \mathbb{N}$ et que $P([Y = n]) \neq 0$ pour tout $n \in \mathbb{N}$ d'après la question précédente, on voit que $Y(\Omega) = \mathbb{N}$, et donc $Z(\Omega) = \mathbb{N}^*$. De plus, on constate d'après la question précédente que, pour tout $n \in \mathbb{N}^*$:

$$P([Z = n]) = P([Y = n - 1]) = e^{-\frac{n-1}{p}}\left(1 - e^{-\frac{1}{p}}\right) = \left(e^{-\frac{1}{p}}\right)^{n-1}\left(1 - e^{-\frac{1}{p}}\right).$$

Par conséquent, on en déduit que :

$$\boxed{Y + 1 \text{ suit la loi géométrique de paramètre } 1 - e^{-\frac{1}{p}}.}$$

(c) Etablissons les inégalités strictes : $0 < E(Y) < p$. Comme $Y+1$ suit la loi géométrique de paramètre $1 - e^{-\frac{1}{p}}$ d'après la question précédente, on sait d'après le cours que $Y + 1$ admet une espérance égale à $E(Y+1) = \frac{1}{1-e^{-\frac{1}{p}}}$. Comme $Y = (Y + 1) - 1$, ceci entraine par linéarité de l'espérance que Y admet une espérance et que :

$$E(Y) = E(Y+1) - 1 = \frac{1}{1 - e^{-\frac{1}{p}}} - 1 = \frac{e^{-\frac{1}{p}}}{1 - e^{-\frac{1}{p}}} = \frac{1}{e^{\frac{1}{p}} - 1}.$$

Comme $\frac{1}{p} > 0$, on voit que $e^{\frac{1}{p}} > 1$, et donc $E(Y) > 0$. Posons maintenant $f(x) = e^x - 1 - x$ pour tout $x \in \mathbb{R}_+$. Alors la fonction f est dérivable sur $\mathbb{R}_+$ et de plus, on a $f'(x) = e^x - 1 > 0$ pour tout $x \in \mathbb{R}_+^*$. En particulier, la fonction f est strictement croissante sur $[0, +\infty[$, ce qui entraine que $f(x) > f(0)$ pour tout $x > 0$, et donc $e^x - 1 > x$ pour tout $x > 0$. Comme $\frac{1}{p} > 0$, ceci entraine que :

$$e^{\frac{1}{p}} - 1 > \frac{1}{p} > 0.$$

Par stricte décroissance de la fonction $t \longmapsto \frac{1}{t}$ sur $]0, +\infty[$, il s'ensuit que $\frac{1}{e^{\frac{1}{p}}-1} < p$, et donc :

$$\boxed{0 < E(Y) < p.}$$

(2) (a) Pour tout couple $(r, s) \in \mathbb{N}^2$, montrons que l'intégrale $\int_0^1 x^r (\ln x)^s dx$ est convergente. Dans ce qui suit, on pose $I_{r,s} = \int_0^1 x^r (\ln x)^s dx$ pour tout $(r, s) \in \mathbb{N}^2$. Fixons alors un couple $(r, s) \in \mathbb{N}^2$. Comme la fonction $x \longmapsto x^r (\ln(x))^s$ est continue sur $]0, 1]$, l'intégrale $I_{r,s}$ est impropre en 0. De plus, on obtient par croissances comparées que :

$$x^{r+1/2} (\ln x)^s \underset{x \to 0^+}{\longrightarrow} 0.$$

En particulier, ceci entraine que :

$$x^r (\ln x)^s \underset{x \to 0^+}{=} o\left(\frac{1}{\sqrt{x}}\right).$$

Comme l'intégrale de Riemann $\int_0^1 \frac{dx}{\sqrt{x}}$ converge et que $\frac{1}{\sqrt{x}} \geq 0$ pour tout $x \in]0, 1]$, le critère de négligeabilité pour les intégrales entraine que, pour tout $(r, s) \in \mathbb{N}^2$:

$$\boxed{\text{l'intégrale } I_{r,s} = \int_0^1 x^r (\ln x)^s dx \text{ converge.}}$$

(b) Etablissons pour tout couple $(r, s) \in \mathbb{N}^2$, l'égalité : $\int_0^1 x^r (\ln x)^s dx = \frac{(-1)^s s!}{(r+1)^{s+1}}$. Pour cela, on conserve les notations de la question précédente. Si l'on pose $u = -\ln x$, c'est-à-dire $x = e^{-u}$, alors on constate que l'application $u \longmapsto e^{-u}$ est de classe $\mathcal{C}^1$ sur $]0, +\infty[$ et bijective de $]0, +\infty[$ sur $]0, 1[$ et que, de plus on a $dx = -e^{-u} du$, u tend vers $+\infty$ quand x tend vers 0 et u tend vers 0 quand x tend vers 1. Par changement de variable, on trouve que :

$$I_{r,s} = \int_{+\infty}^0 \left(e^{-u}\right)^r (-u)^s \times \left(-e^{-u}\right) du = \int_0^{+\infty} (-1)^s u^s e^{-(r+1)u} du.$$

On effectue à nouveau un changement de variable en posant $v = (r + 1)u$. A noter que ce dernier est bien licite car la fonction $u \longmapsto (r + 1)u$ est affine (vu que $r + 1 > 0$). Comme $dv = (r + 1)du$, que v tend vers $+\infty$ quand u tend vers $+\infty$ et que v tend vers 0 quand u tend vers 0, on obtient par changement de variable que :

$$I_{r,s} = \int_0^{+\infty} (-1)^s \left(\frac{v}{r+1}\right)^s e^{-v} \frac{dv}{r+1} = \int_0^{+\infty} \frac{(-1)^s}{(r+1)^{s+1}} v^{s+1-1} e^{-v} dv.$$

Comme $s + 1$ est un entier > 0 et que la fonction $\Gamma : x \longmapsto \int_0^{+\infty} v^{x-1} e^{-v} dv$ est définie sur $\mathbb{R}_+^*$, il s'ensuit par linéarité de l'intégrale que :

$$I_{r,s} = \frac{(-1)^s}{(r+1)^{s+1}} \Gamma(s+1) = \frac{(-1)^s}{(r+1)^{s+1}} s!.$$

Par conséquent, on en déduit que, pour tout $(r, s) \in \mathbb{N}^2$:

$$\boxed{I_{r,s} = \int_0^1 x^r (\ln x)^s dx = \frac{(-1)^s}{(r+1)^{s+1}} s!.}$$

Partie I. Transport dans une situation aléatoire.

On dit que la loi d'une variable aléatoire Y est *accessible* depuis une variable aléatoire X, s'il existe une application $T : X(\Omega) \longrightarrow \mathbb{R}$ telle que la variable aléatoire $T(X)$ suit la même loi que Y. L'application T est alors appelée une *fonction de transport* de la variable aléatoire X vers la loi de Y. On associe à T un *coût de transport* $C(T)$ défini, sous réserve d'existence, par : $C(T) = E\big((X - T(X))^2\big)$. Dans toute cette partie, X

désigne une variable aléatoire vérifiant $X(\Omega) =]0,1[$ et suivant la loi uniforme sur $]0,1[$, c'est-à-dire admettant pour densité la fonction f_X définie par :

$$f_X(x) = \begin{cases} 1 & \text{si } x \in]0,1[\\ 0 & \text{sinon} \end{cases}.$$

(1) Soit p un réel vérifiant $0 < p < 1$. Pour tout réel $a \in [0, 1-p]$, on note dans cette question T_a la fonction définie sur $]0,1[$ par :

$$T_a(x) = \begin{cases} 1 & \text{si } x \in]a, a+p[\\ 0 & \text{sinon} \end{cases}$$

(a) Calculons la probabilité $P([T_a(X) = 1])$ et montrons que les fonctions T_a sont des fonctions de transport de X vers une même loi que l'on précisera. Comme X est une variable à densité, on trouve que :

$$P([T_a(X) = 1]) = P([a < X < a+p]) = F_X(a+p) - F_X(a).$$

Comme X suit la loi uniforme sur $]0,1[$ et que $a, a+p$ appartiennent à $]0,1[$, ceci nous donne que :

$$\boxed{P([T_a(X) = 1]) = a + p - a = p.}$$

Mais comme $T_a(X)$ ne prend que les valeurs 0 et 1 par construction, il s'ensuit que $T_a(X)$ suit la loi de Bernoulli de paramètre p. Par conséquent, on en déduit que :

$$\boxed{\text{les fonctions } T_a \text{ sont des fonctions de transport de } X \text{ vers la loi } \mathcal{B}(p).}$$

(b) Vérifions que le coût de transport $C(T_a)$ est égal à $\frac{1}{3} + p(1-p) - 2ap$. Comme X est à densité, on sait d'après le théorème de transfert que la variable aléatoire $(X - T_a(X))^2$ admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} (t - T_a(t))^2 f_X(t)dt$ converge absolument. Comme X suit la loi uniforme sur $]0,1[$, X admet une densité nulle sur $]-\infty, 0] \cup [1, +\infty[$ et ceci revient à vérifier que l'intégrale $\int_0^1 (t - T_a(t))^2 dt$ converge absolument. Comme $0 \leq a < a+p \leq 1$ par hypothèse, on trouve par des calculs simples que :

$$(t - T_a(t))^2 = \begin{cases} (t-1)^2 & \text{si } t \in]a, a+p[\\ t^2 & \text{sinon} \end{cases}.$$

Comme les fonctions $t \longmapsto (t-1)^2$ et $t \longmapsto t^2$ sont continues sur les segments $[a, a+p]$, $[0,a]$, $[a+p, 1]$, les intégrales $\int_a^{a+p} (t - T_a(t))^2 dt$, $\int_0^a (t - T_a(t))^2 dt$ et $\int_{a+p}^1 (t - T_a(t))^2 dt$ convergent absolument, et donc l'intégrale $\int_0^1 (t - T_a(t))^2 dt$ converge absolument d'après la relation de Chasles, d'où l'existence du coût de transport $C(T_a)$. De plus, d'après le théorème de transfert, la relation de Chasles et la formule du binôme, on trouve que :

$$\begin{aligned}
C(T_a) &= \int_{-\infty}^{+\infty} (t - T_a(t))^2 f_X(t)dt \\
&= \int_0^1 (t - T_a(t))^2 f_X(t)dt \\
&= \int_0^a (t - T_a(t))^2 dt + \int_a^{a+p} (t - T_a(t))^2 dt + \int_{a+p}^1 (t - T_a(t))^2 dt \\
&= \int_0^a t^2 dt + \int_a^{a+p} (t-1)^2 dt + \int_{a+p}^1 t^2 dt \\
&= \left[\frac{t^3}{3}\right]_0^a + \left[\frac{(t-1)^3}{3}\right]_a^{a+p} + \left[\frac{t^3}{3}\right]_{a+p}^1 \\
&= \frac{a^3}{3} + \frac{(a+p-1)^3}{3} - \frac{(a-1)^3}{3} + \frac{1}{3} - \frac{(a+p)^3}{3} \\
&= \frac{a^3 + (a+p-1)^3 - (a-1)^3 + 1 - (a+p)^3}{3} \\
&= \frac{a^3 + (a-1)^3 + 3p(a-1)^2 + 3p^2(a-1) + p^3 - (a-1)^3 + 1 - (a+p)^3}{3} \\
&= \frac{a^3 + 3p(a-1)^2 + 3p^2(a-1) + p^3 + 1 - (a+p)^3}{3} \\
&= \frac{a^3 + 3pa^2 - 6pa + 3p + 3p^2a - 3p^2 + p^3 + 1 - a^3 - 3pa^2 - 3p^2a - p^3}{3} \\
&= \frac{-6pa + 3p - 3p^2 + 1}{3}.
\end{aligned}$$

Par conséquent, on en déduit après simplification que :

$$\boxed{C(T_a) = \frac{1}{3} + p(1-p) - 2ap.}$$

(c) Déterminons la valeur de a qui minimise $C(T_a)$ et exprimons le coût minimal correspondant en fonction de p. Comme la fonction affine $a \longmapsto \frac{1}{3} + p\,(1-p) - 2ap$ est strictement décroissante sur l'intervalle $[0, 1-p]$, elle atteint son minimum sur $[0, 1-p]$ en $a = 1-p$. En particulier, on voit avec la question précédente que $C(T_a)$ est minimal si et seulement si $a = 1-p$, et dans ce cas :

$$C(T_{1-p}) = \frac{1}{3} + p(1-p) - 2p(1-p) = \frac{1}{3} - p(1-p).$$

Par conséquent, on en déduit que :

$$\boxed{C\,(T_a) \text{ est minimal pour } a = 1-p \text{ et le coût minimal vaut } C\,(T_{1-p}) = \frac{1}{3} - p\,(1-p)\,.}$$

(2) Soit T_1 et T_2 les applications définies sur $]0, 1[$ par $T_1(x) = -\ln x$ et $T_2(x) = -\ln(1-x)$.

(a) Vérifions que T_1 et T_2 sont des fonctions de transport de X vers une loi que l'on précisera. Comme X suit la loi uniforme sur $]0, 1[$, on sait que $X(\Omega) =]0, 1[$, ce qui entraine que $T_1(X)(\Omega) \subset \mathbb{R}_+^*$ et $T_2(X)(\Omega) \subset \mathbb{R}_+^*$, et donc $F_{T_1(X)}(x) = F_{T_2(X)}(x) = 0$ pour tout $x \leq 0$. De plus, pour tout $x > 0$, on trouve par croissance de l'exponentielle que :

$$F_{T_1(X)}(x) = P([-\ln(X) \leq x]) = P([\ln(X) \geq -x]) = P([X \geq e^{-x}]).$$

Comme X est à densité, ceci nous donne que, pour tout $x > 0$:

$$F_{T_1(X)}(x) = P([X \geq e^{-x}]) = 1 - P([X < e^{-x}]) = 1 - F_X(e^{-x}).$$

Comme X suit la loi uniforme sur $]0, 1[$ et que e^{-x} appartient à $]0, 1[$ car $x > 0$, ceci entraine que :

$$F_{T_1(X)}(x) = 1 - F_X(e^{-x}) = 1 - e^{-x}.$$

En d'autres termes, on vient de trouver que :

$$F_{T_1(X)}(x)=\left\{\begin{array}{cc} 1-e^{-x} & \text{si } x>0 \\ 0 & \text{sinon}\end{array}\right. .$$

De même, pour tout $x>0$, on trouve par croissance de l'exponentielle que :

$$F_{T_2(X)}(x)=P([-\ln(1-X)\leq x])=P([\ln(1-X)\geq -x])=P([1-X\geq e^{-x}]).$$

Comme X suit la loi uniforme sur $]0,1[$ et que e^{-x} appartient à $]0,1[$ car $x>0$, ceci entraine que :

$$F_{T_2(X)}(x)=P([1-X\geq e^{-x}])=P([X\leq 1-e^{-x}])=F_X(1-e^{-x})=1-e^{-x}.$$

En d'autres termes, on vient de trouver que :

$$F_{T_2(X)}(x)=\left\{\begin{array}{cc} 1-e^{-x} & \text{si } x>0 \\ 0 & \text{sinon}\end{array}\right. .$$

D'après les calculs ci-dessus, on voit que $T_1(X)$ et $T_2(X)$ suivent la loi exponentielle de paramètre 1. Par conséquent, on en déduit que :

$$\boxed{T_1 \text{ et } T_2 \text{ sont des fonctions de transport de } X \text{ vers la loi } \mathcal{E}(1).}$$

(b) Comparons les coûts de transport $C(T_1)$ et $C(T_2)$. Pour ce faire, on commence par calculer $C(T_1)$. Comme X est à densité, on sait d'après le théorème de transfert que la variable aléatoire $(X-T_1(X))^2$ admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty}(t-T_1(t))^2 f_X(t)dt$ converge absolument. Comme X suit la loi uniforme sur $]0,1[$, X admet une densité nulle sur $]-\infty,0]\cup[1,+\infty[$ et ceci revient à vérifier que l'intégrale $\int_0^1(t-T_1(t))^2dt=\int_0^1(t+\ln(t))^2dt$ converge absolument, c'est-à-dire converge car l'intégrande est positive. Or on voit par des calculs simples que $(t+\ln(t))^2=t^2+2t\ln(t)+\ln^2(t)$ pour tout $t\in]0,1[$. Comme les intégrales du type $\int_0^1 t^r\ln^s(t)dt$ convergent pour tout $(r,s)\in\mathbb{N}^2$ d'après la question (2) du préliminaire, ceci entraine que les intégrales $\int_0^1 t^2dt$, $\int_0^1 t\ln(t)dt$ et $\int_0^1\ln^2(t)dt$ convergent, et donc l'intégrale $\int_0^1(t+\ln(t))^2dt$ converge par linéarité de l'intégration. En particulier, la variable aléatoire $(X-T_1(X))^2$ admet bien une espérance et le coût de transport $C(T_1)$ est bien défini. De plus, d'après le théorème de transfert, on trouve que :

$$C(T_1)=\int_{-\infty}^{+\infty}(t-T_1(t))^2 f_X(t)dt=\int_0^1(t-T_1(t))^2 f_X(t)dt=\int_0^1(t+\ln(t))^2dt.$$

Par linéarité de l'intégration et d'après la question (2) du préliminaire, ceci nous donne que :

$$\begin{aligned} C(T_1) &= \int_0^1(t+\ln(t))^2dt \\ &= \int_0^1 t^2dt+2\int_0^1 t\ln(t)dt+\int_0^1\ln^2(t)dt \\ &= \frac{(-1)^0}{(2+1)^{0+1}}0!+2\frac{(-1)^1}{(1+1)^{1+1}}1!+\frac{(-1)^2}{(0+1)^{2+1}}2! \\ &= \frac{1}{3}-2\times\frac{1}{4}+2. \end{aligned}$$

Par conséquent, on en déduit après simplification que :

$$\boxed{C(T_1)=\frac{11}{6}.}$$

Passons maintenant au calcul de $C(T_2)$. Comme X suit la loi uniforme sur $]0,1[$, $Y=1-X$ suit aussi la loi uniforme sur $]0,1[$. De plus, on voit que $X-T_2(X)=X+\ln(1-X)=1-Y+\ln(Y)=1-Y+T_1(Y)$. Comme Y est à densité, on sait d'après le théorème de transfert que la variable aléatoire $(X-T_2(X))^2=(1-Y+T_1(Y))^2$ admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty}(1-t+T_1(t))^2 f_Y(t)dt$ converge absolument. Comme Y suit la loi uniforme sur $]0,1[$, Y admet une densité nulle sur $]-\infty,0]\cup[1,+\infty[$ et ceci revient à vérifier que l'intégrale $\int_0^1(1-t+T_1(t))^2dt=\int_0^1(1-t+\ln(t))^2dt$ converge absolument, c'est-à-dire converge car l'intégrande est positive. Or on voit par des calculs simples que, pour tout $t\in]0,1[$:

$$(1-t+\ln(t))^2=1+t^2+\ln^2(t)-2t+2\ln(t)-2t\ln(t).$$

Comme les intégrales du type $\int_0^1 t^r \ln^s(t)dt$ convergent pour tout $(r,s) \in \mathbb{N}^2$ d'après la question (2) du préliminaire, ceci entraine que toutes les intégrales sur $[0,1]$ des fonctions ci-dessus convergent, et donc l'intégrale $\int_0^1 (1-t+\ln(t))^2 dt$ converge par linéarité de l'intégration. En particulier, la variable aléatoire $(X - T_2(X))^2$ admet bien une espérance et le coût de transport $C(T_2)$ est bien défini. De plus, d'après le théorème de transfert, on trouve que :

$$C(T_2) = \int_{-\infty}^{+\infty} (1-t+T_1(t))^2 f_Y(t)dt = \int_0^1 (1-t+T_1(t))^2 f_Y(t)dt = \int_0^1 (1-t+\ln(t))^2 dt.$$

Par linéarité de l'intégration et d'après la question (2) du préliminaire, ceci nous donne que :

$$\begin{aligned} C(T_2) &= \int_0^1 (1-t+\ln(t))^2 dt \\ &= \int_0^1 1.dt + \int_0^1 t^2 dt + \int_0^1 \ln^2(t)dt - 2\int_0^1 tdt + 2\int_0^1 \ln(t)dt - 2\int_0^1 t\ln(t)dt \\ &= 1 + \frac{1}{3} + \frac{(-1)^2 2!}{(0+1)^{2+1}} - 2 \times \frac{1}{2} + 2 \times \frac{(-1)^1 1!!}{(0+1)^{1+1}} - 2 \times \frac{(-1)^1 1!}{(1+1)^{1+1}} \\ &= 1 + \frac{1}{3} + 2 - 1 - 2 + \frac{1}{2}. \end{aligned}$$

Par conséquent, on en déduit après simplification que :

$$\boxed{C(T_2) = \frac{5}{6}.}$$

D'après les calculs faits ci-avant, on en déduit que :

$$\boxed{C(T_2) < C(T_1).}$$

(c) Montrons que toutes les lois géométriques sont accessibles depuis X. Pour ce faire, fixons un réel $a \in]0,1[$. D'après la question (2)(a) de la partie I, on sait que $T_1(X) = -\ln(X)$ suit la loi exponentielle de paramètre 1. De plus, d'après la question (2) du préliminaire, on voit que $\lfloor aT_1(X) \rfloor + 1$ suit la loi géométrique de paramètre $1 - e^{-\frac{1}{a}}$. Or, pour tout $p \in]0,1[$, on constate que :

$$1 - e^{-\frac{1}{a}} = p \quad \Longleftrightarrow \quad 1 - p = e^{-\frac{1}{a}} \quad \Longleftrightarrow \quad \frac{1}{a} = -\ln(1-p) \quad \Longleftrightarrow \quad a = -\frac{1}{\ln(1-p)}.$$

Si l'on pose $q = 1-p$ et $a = -\frac{1}{\ln(q)}$, alors il s'ensuit que la variable aléatoire $Y = \lfloor -a\ln(X) \rfloor + 1$ suit la loi géométrique de paramètre p. En particulier, la fonction :

$$T : \left\{ \begin{array}{ccc}]0,1[& \longrightarrow & \mathbb{R} \\ t & \longmapsto & \left\lfloor \frac{-\ln(t)}{-\ln(q)} \right\rfloor + 1 \end{array} \right.$$

permet d'accéder à la loi $\mathcal{G}(p)$ depuis X. Comme ceci est vrai pour tout $p \in]0,1[$, on en déduit que :

$$\boxed{\text{toutes les lois géométriques sont accessibles depuis } X.}$$

(3) Dans cette question, Y désigne une variable aléatoire admettant une densité f_Y continue et strictement positive sur $\mathbb{R}$.

(a) Justifions que la fonction de répartition F_Y de Y réalise une bijection de $\mathbb{R}$ sur l'intervalle $]0,1[$. Comme Y est une variable aléatoire admettant une densité f_Y continue et strictement positive sur $\mathbb{R}$, la fonction de répartition F_Y est dérivable sur $\mathbb{R}$ et de plus, on a $(F_Y)'(t) = f_Y(t) > 0$ pour tout $t \in \mathbb{R}$. En particulier, la fonction F_Y est continue et strictement croissante sur $\mathbb{R}$. D'après le théorème de la bijection, la fonction F_Y réalise une bijection de $\mathbb{R}$ sur $F_Y(\mathbb{R})$. Comme F_Y est continue sur $\mathbb{R}$, l'image $F_Y(\mathbb{R})$ est un intervalle d'après le théorème des valeurs intermédiaires. Mais comme F_Y est strictement croissante sur $\mathbb{R}$, qu'elle tend vers 0 en $-\infty$ et vers 1 en $+\infty$ en tant que fonction de répartition, il s'ensuit que $F_Y(\mathbb{R}) =]0,1[$. Par conséquent, on en déduit que :

$$\boxed{\text{la fonction } F_Y \text{ réalise une bijection de } \mathbb{R} \text{ sur }]0,1[.}$$

(b) On note F_Y^{-1} la bijection réciproque de F_Y. Montrons que F_Y^{-1} est une fonction de transport de la variable aléatoire X vers la loi de Y. Pour ce faire, on pose $Z = F_Y^{-1}(X)$. Comme F_Y est strictement croissante sur $\mathbb{R}$ d'après la question précédente, on trouve que, pour tout $x \in \mathbb{R}$:

$$F_Z(x) = P\left([F_Y^{-1}(X) \leq x]\right) = P([X \leq F_Y(x)]).$$

Comme $F_Y(x)$ appartient à $]0,1[$ pour tout $x \in \mathbb{R}$ et que X suit la loi uniforme sur $]0,1[$, ceci nous donne que, pour tout $x \in \mathbb{R}$:

$$F_Z(x) = P([X \leq F_Y(x)]) = F_X(F_Y(x)) = F_Y(x),$$

et donc Y et $Z = F_Y^{-1}(X)$ suivent la même loi. Par conséquent, on en déduit que :

$$\boxed{F_Y^{-1} \text{ est une fonction de transport de } X \text{ vers la loi de } Y.}$$

(4) *Cas particulier : on suppose que Y suit la loi normale centrée réduite.* On note F_Y la fonction de répartition de Y et φ la densité continue sur $\mathbb{R}$ de Y.

(a) Etablissons tout d'abord la convergence de l'intégrale $\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy$. Comme F_Y est une fonction de répartition, elle est à valeurs dans $[0,1]$, ce qui entraine que $0 \leq |yF_Y(y)\varphi(y)| \leq |y|\varphi(y)$ pour tout $y \in \mathbb{R}$. Comme Y suit la loi normale centrée réduite, Y admet une espérance d'après le cours et l'intégrale $\int_{-\infty}^{+\infty} |y|\varphi(y)dy$ converge. D'après le critère de comparaison des intégrales de fonctions positives, l'intégrale $\int_{-\infty}^{+\infty} |yF_Y(y)\varphi(y)|dy$ converge. En particulier, l'intégrale $\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy$ converge absolument, et donc :

$$\boxed{\text{l'intégrale } \int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy \text{ converge.}}$$

A présent, montrons que :

$$\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy = \frac{1}{2\sqrt{\pi}}.$$

Pour ce faire, fixons deux réels a, b tels que $a < b$, et posons $u(y) = F_Y(y)$ et $v(y) = -e^{-\frac{y^2}{2}}$ pour tout $y \in [a,b]$. Alors u et v sont de classe $\mathcal{C}^1$ sur $[a,b]$ et de plus, on a $u'(y) = \varphi(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$ et $v'(y) = ye^{-\frac{y^2}{2}}$ pour tout $y \in [a,b]$. Par intégration par parties, on trouve que :

$$\begin{aligned}
\int_a^b yF_Y(y)\varphi(y)dy &= \left[-\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}F_Y(y)\right]_a^b + \frac{1}{2\pi}\int_a^b e^{-y^2}dy \\
&= -\frac{1}{\sqrt{2\pi}}e^{-\frac{b^2}{2}}F_Y(b) + \frac{1}{\sqrt{2\pi}}e^{-\frac{a^2}{2}}F_Y(a) + \frac{1}{2\pi}\int_a^b e^{-y^2}dy.
\end{aligned}$$

Comme F_Y tend vers 0 en $-\infty$ et vers 1 en $+\infty$ en tant que fonction de répartition, que $e^{-\frac{b^2}{2}}$ tend vers 0 quand b tend vers $+\infty$ et que $e^{-\frac{a^2}{2}}$ tend vers 0 quand a tend vers $-\infty$, on obtient par passage à la limite quand a tend vers $-\infty$ et quand b tend vers $+\infty$ dans l'égalité ci-dessus que :

$$\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-y^2}dy. \quad (*)$$

Par ailleurs, considérons la fonction :

$$h : y \longmapsto \frac{1}{\sqrt{\pi}}\exp(-y^2) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(\frac{y}{\sqrt{\frac{1}{2}}}\right)^2\right).$$

Alors on voit que h est une densité d'une variable aléatoire qui suit la loi $\mathcal{N}\left(0, \frac{1}{2}\right)$, et donc l'intégrale $\int_{-\infty}^{+\infty} h(y)dy$ converge et vaut 1. En particulier, il s'ensuit par linéarité de l'intégration que :

$$\int_{-\infty}^{+\infty} e^{-y^2}dy = \sqrt{\pi}.$$

Par conséquent, on en déduit avec l'égalité $(*)$ que :

$$\boxed{\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy = \frac{1}{2\pi} \times \sqrt{\pi} = \frac{1}{2\sqrt{\pi}}.}$$

(b) Montrons que l'intégrale $\int_{-\infty}^{+\infty}(y - F_Y(y))^2\varphi(y)dy$ est convergente et calculons-la. Pour tout $y \in \mathbb{R}$, on commence par remarquer que $(y - F_Y(y))^2\varphi(y) = y^2\varphi(y) - 2yF_Y(y)\varphi(y) + (F_Y(y))^2\varphi(y)$ $(*)$.

Comme Y suit la loi normale centrée réduite, elle admet un moment d'ordre 2 et donc l'intégrale $\int_{-\infty}^{+\infty} y^2\varphi(y)\,dy$ converge. De plus, on trouve d'après la formule de Koenig-Huygens que :

$$\int_{-\infty}^{+\infty} y^2\varphi(y)\,dy = E(Y^2) = V(Y) + E(Y)^2 = 1 + 0^2 = 1.$$

En outre, on sait d'après la question précédente que l'intégrale $\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy$ converge et vaut $\frac{1}{2\sqrt{\pi}}$. Par ailleurs, comme $(F_Y)' = \varphi$, on obtient que, pour tous réels a, b tels que $a < b$:

$$\int_a^b (F_Y(y))^2\varphi(y)dy = \left[\frac{1}{3}(F_Y(y))^3\right]_a^b = \frac{1}{3}(F_Y(b))^3 - \frac{1}{3}(F_Y(a))^3.$$

Comme F_Y est une fonction de répartition, elle tend vers 0 en $-\infty$ et vers 1 en $+\infty$. Par passage à la limite quand a tend vers $-\infty$ et quand b tend vers $+\infty$ dans l'égalité ci-dessus, ceci entraine que l'intégrale $\int_{-\infty}^{+\infty}(F_Y(y))^2\varphi(y)dy$ converge et vaut $\frac{1}{3}$. Dès lors, il s'ensuit d'après la relation $(*)$ et par linéarité de l'intégration que l'intégrale $\int_{-\infty}^{+\infty}(y - F_Y(y))^2\varphi(y)dy$ converge. De plus, toujours par linéarité, on trouve que :

$$\begin{aligned}\int_{-\infty}^{+\infty}(y - F_Y(y))^2\varphi(y)dy &= \int_{-\infty}^{+\infty} y^2\varphi(y)dy - 2\int_{-\infty}^{+\infty} yF_Y(y)\varphi(y)dy + \int_{-\infty}^{+\infty}(F_Y(y))^2\varphi(y)dy \\ &= 1 - 2\times\frac{1}{2\sqrt{\pi}} + \frac{1}{3}.\end{aligned}$$

Par conséquent, on en déduit après simplification que :

$$\boxed{\text{l'intégrale } \int_{-\infty}^{+\infty}(y - F_Y(y))^2\varphi(y)dy \text{ converge et vaut } \frac{4}{3} - \frac{1}{\sqrt{\pi}}.}$$

(c) Montrons que le coût de transport $C(F_Y^{-1})$ est égal à $\frac{4}{3} - \frac{1}{\sqrt{\pi}}$. Comme X est à densité, on sait d'après le théorème de transfert que la variable aléatoire $(X - F_Y^{-1}(X))^2$ admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty}(t - F_Y^{-1}(t))^2 f_X(t)dt$ converge absolument. Comme X suit la loi uniforme sur $]0, 1[$, X admet une densité nulle sur $]-\infty, 0] \cup [1, +\infty[$ et ceci revient à vérifier que l'intégrale $I = \int_0^1(t - F_Y^{-1}(t))^2 f_X(t)dt = \int_0^1(t - F_Y^{-1}(t))^2dt$ converge absolument, c'est-à-dire converge car l'intégrande est positive. Posons alors $t = F_Y(y)$. D'après la question (3)(a) de la partie I, la fonction F_Y est de classe $\mathcal{C}^1$ et bijective de $\mathbb{R}$ sur $]0, 1[$, et de plus on a $dt = \varphi(y)dy$, y tend vers $-\infty$ quand t tend vers 0 et y tend vers $+\infty$ quand t tend vers 1. Par changement de variable, l'intégrale I est de de même nature que l'intégrale :

$$J = \int_{-\infty}^{+\infty}(F_Y(y) - F_Y^{-1}\circ F_Y(y))^2\varphi(y)dy = \int_{-\infty}^{+\infty}(y - F_Y(y))^2\varphi(y)dy,$$

et de plus, elles sont égales en cas de convergence. Comme l'intégrale J converge et vaut $\frac{4}{3} - \frac{1}{\sqrt{\pi}}$ d'après la question précédente, il s'ensuit que l'intégrale I converge et vaut aussi $\frac{4}{3} - \frac{1}{\sqrt{\pi}}$. Par conséquent, on en déduit que :

$$\boxed{\text{le coût de transport } C(F_Y^{-1}) \text{ est égal à } \frac{4}{3} - \frac{1}{\sqrt{\pi}}.}$$

Partie II. Transport optimal dans une situation déterministe.

Dans toute cette partie, N désigne un entier supérieur ou égal à 2. On considère N réels $d_1, d_2, \ldots, d_N$ (appelés points de départ) et N réels $a_1, a_2, \ldots, a_N$ (appelés points d'arrivée) vérifiant $d_1 < d_2 < \cdots < d_N$ et $a_1 < a_2 < \cdots < a_N$. On pose $D = \{d_1, d_2, \ldots, d_N\}$ et $A = \{a_1, a_2, \ldots, a_N\}$.

(1) (a) Montrons que pour tout couple $(k, l) \in [\![1, N]\!]^2$, on a : $d_ka_k \geq d_ka_l + d_la_k - d_la_l$. Pour ce faire, fixons un couple $(k, l) \in [\![1, N]\!]^2$. Si $k \leq l$, alors on voit que $d_k \leq d_l$ et $a_k \leq a_l$ car les suites finies (a_i) et (d_i) sont croissantes, ce qui entraine que $d_k - d_l \leq 0$ et $a_k - a_l \leq 0$, et donc $(d_k - d_l)(a_k - a_l) \geq 0$. Si maintenant $k \geq l$, alors on voit que $d_k \geq d_l$ et $a_k \geq a_l$ car les suites finies (a_i) et (d_i) sont croissantes, ce qui entraine que $d_k - d_l \geq 0$ et $a_k - a_l \geq 0$, et donc $(d_k - d_l)(a_k - a_l) \geq 0$. Dans tous les cas, on constate que, pour tout $(k, l) \in [\![1, N]\!]^2$:

$$(d_k - d_l)(a_k - a_l) \geq 0.$$

et donc $d_ka_k - d_ka_l - d_la_k + d_la_l \geq 0$. Par conséquent, on en déduit que, pour tout $(k, l) \in [\![1, N]\!]^2$:

$$\boxed{d_ka_k \geq d_ka_l + d_la_k - d_la_l.}$$

(b) Montrons que, pour tout N-uplet $(p_1,p_2,\dots,p_N)\in\mathbb{R}_+^N$ tel que $\sum_{k=1}^N p_k=1$, on a :

$$\sum_{k=1}^N p_k d_k a_k \geq \left(\sum_{k=1}^N p_k d_k\right)\times\left(\sum_{k=1}^N p_k a_k\right) \qquad (1)$$

Soit $(p_1,p_2,...,p_N)\in(\mathbb{R}_+)^N$ tel que $\sum_{k=1}^N p_k=1$. Comme $d_k a_k \geq d_k a_l + d_l a_k - d_l a_l$ pour tout $(k,l)\in[\![1,N]\!]^2$ d'après la question précédente et que tous les p_i sont positifs par hypothèse, on obtient par produit que, pour tout $(k,l)\in[\![1,N]\!]^2$:

$$p_k p_l d_k a_k \geq p_k p_l d_k a_l + p_k p_l d_l a_k - p_k p_l d_l a_l.$$

Par double sommation sur les entiers k et l, on trouve que :

$$\sum_{k=1}^N\sum_{l=1}^N p_k p_l d_k a_k \geq \sum_{k=1}^N\sum_{l=1}^N \left(p_k p_l d_k a_l + p_k p_l d_l a_k - p_k p_l d_l a_l\right),$$

ce que l'on peut réécrire sous la forme suivante (et ce par linéarité de la somme) :

$$\left(\sum_{l=1}^N p_l\right)\left(\sum_{k=1}^N p_k d_k a_k\right) \geq \left(\sum_{k=1}^N p_k d_k\right)\left(\sum_{l=1}^N p_l a_l\right)+\left(\sum_{l=1}^N p_l d_l\right)\left(\sum_{k=1}^N p_k a_k\right)-\left(\sum_{k=1}^N p_k\right)\left(\sum_{l=1}^N p_l d_l a_l\right).$$

Comme $\sum_{k=1}^N p_k=\sum_{l=1}^N p_l=1$, ceci entraine que :

$$\sum_{k=1}^N p_k d_k a_k \geq \left(\sum_{k=1}^N p_k d_k\right)\left(\sum_{l=1}^N p_l a_l\right)+\left(\sum_{l=1}^N p_l d_l\right)\left(\sum_{k=1}^N p_k a_k\right)-\sum_{l=1}^N p_l d_l a_l.$$

En d'autres termes, on vient de trouver que :

$$\sum_{k=1}^N p_k d_k a_k \geq 2\left(\sum_{k=1}^N p_k d_k\right)\left(\sum_{l=1}^N p_l a_l\right)-\sum_{l=1}^N p_l d_l a_l.$$

En faisant passer la somme de droite ci-dessus de l'autre côté, il s'ensuit que :

$$2\sum_{k=1}^N p_k d_k a_k \geq 2\left(\sum_{k=1}^N p_k d_k\right)\left(\sum_{l=1}^N p_l a_l\right).$$

Par conséquent, on en déduit après division par 2 que :

$$\boxed{\sum_{k=1}^N p_k d_k a_k \geq \left(\sum_{k=1}^N p_k d_k\right)\left(\sum_{l=1}^N p_l a_l\right).}$$

(2) Soit $t\in\mathcal{E}_N$. On réordonne la liste $\big(t(1),t(2),\dots,t(N)\big)$ selon les valeurs croissantes et on note alors $\big(\widehat{t}(1),\widehat{t}(2),\dots,\widehat{t}(N)\big)$ la liste ordonnée obtenue. On a donc $\widehat{t}(1)\leq\widehat{t}(2)\leq\cdots\leq\widehat{t}(N)$.

(a) Justifions pour tout $n\in[\![1,N]\!]$ l'inégalité : $\sum_{k=n}^N a_{t(k)} \leq \sum_{k=n}^N a_{\widehat{t}(k)}$. Pour ce faire, fixons un entier $n\in[\![1,N]\!]$. Comme la suite finie (a_i) est croissante et que $\widehat{t}(1)\leq\widehat{t}(2)\leq\cdots\leq\widehat{t}(N)$ par construction, l'expression $\sum_{k=n}^N a_{\widehat{t}(k)}$ correspond à la somme des $N-n+1$ plus grands éléments de la suite $\big(a_{t(1)},...,a_{t(N)}\big)$. En particulier, on voit que l'expression $\sum_{k=n}^N a_{\widehat{t}(k)}$ est supérieure ou égale à la somme de $N-n+1$ éléments quelconques de la liste $\big(a_{t(1)},...,a_{t(N)}\big)$, et donc :

$$\boxed{\sum_{k=n}^N a_{t(k)} \leq \sum_{k=n}^N a_{\widehat{t}(k)}.}$$

(b) On pose $d_0 = 0$. Justifions l'égalité : $\sum_{n=1}^N d_n a_{t(n)} = \sum_{n=1}^N \left((d_n - d_{n-1}) \sum_{k=n}^N a_{t(k)}\right)$. Par interversion des sommes, on trouve que :

$$\begin{aligned}
\sum_{n=1}^N \left((d_n - d_{n-1}) \sum_{k=n}^N a_{t(k)}\right) &= \sum_{n=1}^N \sum_{k=n}^N \left((d_n - d_{n-1})\, a_{t(k)}\right) \\
&= \sum_{k=1}^N \sum_{n=1}^k \left((d_n - d_{n-1})\, a_{t(k)}\right) \\
&= \sum_{k=1}^N \left(a_{t(k)} \sum_{n=1}^k (d_n - d_{n-1})\right).
\end{aligned}$$

Par télescopage, on obtient que :

$$\sum_{n=1}^N \left((d_n - d_{n-1}) \sum_{k=n}^N a_{t(k)}\right) = \sum_{k=1}^N \left(a_{t(k)} \sum_{n=1}^k (d_n - d_{n-1})\right) = \sum_{k=1}^N a_{t(k)}(d_k - d_0).$$

Mais comme $d_0 = 0$ par convention, on en déduit que :

$$\boxed{\sum_{n=1}^N \left((d_n - d_{n-1}) \sum_{k=n}^N a_{t(k)}\right) = \sum_{k=1}^N a_{t(k)} d_k.}$$

(c) Etablissons l'inégalité : $\sum_{n=1}^N d_n a_{t(n)} \leq \sum_{n=1}^N d_n a_{\widehat{t}(n)}$ (2). Pour ce faire, on pose $d_0 = d_1$ et $c_k = d_k - d_0$ pour tout $k \in [\![0, N]\!]$. Par construction, la suite finie $(c_k)_{0 \leq k \leq N}$ est croissante, et donc on a $c_n - c_{n-1} \geq 0$ pour tout $n \in [\![1, N]\!]$. Comme $\sum_{k=n}^N a_{t(k)} \leq \sum_{k=n}^N a_{\widehat{t}(k)}$ pour tout $n \in [\![1, N]\!]$ d'après la question (2)(a) de la partie II, on trouve par produit que, pour tout $n \in [\![1, N]\!]$:

$$(c_n - c_{n-1}) \sum_{k=n}^N a_{t(k)} \leq (c_n - c_{n-1}) \sum_{k=n}^N a_{\widehat{t}(k)}.$$

Par sommation sur l'entier n, ceci nous donne avec la question précédente que :

$$\begin{aligned}
\sum_{n=1}^N c_n a_{t(n)} &= \sum_{n=1}^N \left((c_n - c_{n-1}) \sum_{k=n}^N a_{t(k)}\right) \\
&\leq \sum_{n=1}^N \left((c_n - c_{n-1}) \sum_{k=n}^N a_{\widehat{t}(k)}\right) \\
&\leq \sum_{n=1}^N c_n a_{\widehat{t}(n)}.
\end{aligned}$$

En revenant à la définition des c_i, ceci entraine que :

$$\sum_{n=1}^N (d_n - d_1) a_{t(n)} \leq \sum_{n=1}^N (d_n - d_1) a_{\widehat{t}(n)}.$$

Par linéarité de la somme, on obtient que :

$$\sum_{n=1}^N d_n a_{t(n)} - d_1 \sum_{n=1}^N a_{t(n)} \leq \sum_{n=1}^N d_n a_{\widehat{t}(n)} - d_1 \sum_{n=1}^N a_{\widehat{t}(n)}. \quad (*)$$

Comme $\left(\widehat{t}(1), \widehat{t}(2), \ldots, \widehat{t}(N)\right)$ est la liste ordonnée selon les valeurs croissantes et obtenue à partir de la liste $\left(t(1), t(2), \ldots, t(N)\right)$, ces deux listes comptabilisent les mêmes éléments de $[\![1, N]\!]$ le même nombre de fois, et donc :

$$\sum_{n=1}^N a_{t(n)} = \sum_{n=1}^N a_{\widehat{t}(n)}.$$

Par différence dans l'inégalité $(*)$, on en déduit que :

$$\boxed{\sum_{n=1}^N d_n a_{t(n)} \leq \sum_{n=1}^N d_n a_{\widehat{t}(n)}.}$$

On appelle *programme de transport* toute bijection T de D sur A et *coût* d'un programme de transport T la somme $c(T)$ définie par $c(T) = \sum_{k=1}^{N} \left(d_k - T(d_k)\right)^2$.

(3) Soit $\widehat{T}$ le programme de transport défini pour tout $k \in [\![1, N]\!]$ par $\widehat{T}(d_k) = a_k$. Montrons que le programme $\widehat{T}$ est optimal, c'est-à-dire que, pour tout programme de transport T, on a : $c(T) \geq c(\widehat{T})$. Pour ce faire, fixons une bijection T de D sur A. Pour tout $k \in [\![1, N]\!]$, on désigne par $t(k)$ l'unique indice de $[\![1, N]\!]$ tel que $T(d_k) = a_{t(k)}$. Ce faisant, on définit bien une application t de $[\![1, N]\!]$ dans $[\![1, N]\!]$, qui est de plus bijective car T l'est. En particulier, on a :

$$\left(\widehat{t}(1), \widehat{t}(2), ..., \widehat{t}(N)\right) = (1, 2, ..., N).$$

Par des calculs simples, on trouve que :

$$c(T) = \sum_{k=1}^{N} (d_k - T(d_k))^2 = \sum_{k=1}^{N} d_k^2 - 2\sum_{k=1}^{N} d_k T(d_k) + \sum_{k=1}^{N} (T(d_k))^2. \quad (*)$$

Or on voit d'après la question précédente que :

$$\sum_{k=1}^{N} d_k T(d_k) = \sum_{k=1}^{N} d_k a_{t(k)} \leq \sum_{k=1}^{N} d_k a_{\widehat{t}(k)} = \sum_{k=1}^{N} d_k a_k = \sum_{k=1}^{N} d_k \widehat{T}(d_k),$$

ce qui entraine que $-2\sum_{k=1}^{N} d_k T(d_k) \geq -2\sum_{k=1}^{N} d_k \widehat{T}(d_k)$. En reportant cette inégalité dans $(*)$, ceci nous donne que :

$$c(T) \geq \sum_{k=1}^{N} d_k^2 - 2\sum_{k=1}^{N} d_k \widehat{T}(d_k) + \sum_{k=1}^{N} (T(d_k))^2. \quad (**)$$

Par ailleurs, on voit par construction que :

$$\sum_{k=1}^{N} (T(d_k))^2 = \sum_{k=1}^{N} \left(a_{t(k)}\right)^2.$$

En particulier, comme l'application t est une bijection de $[\![1, N]\!]$, la somme des carrés des $a_{t(k)}$ est égale à celle des carrés des a_k (vu que les listes $(a_{t(1)}, ..., a_{t(N)})$ et $(a_1, ..., a_N)$ sont formées des mêmes réels comptés le même nombre de fois mais dans des ordres éventuellement différents). En d'autres termes, on a par définition de $\widehat{T}$:

$$\sum_{k=1}^{N} (T(d_k))^2 = \sum_{k=1}^{N} \left(a_{t(k)}\right)^2 = \sum_{k=1}^{N} a_k^2 = \sum_{k=1}^{N} \left(\widehat{T}(d_k)\right)^2.$$

En reportant cette égalité dans $(**)$, il s'ensuit que :

$$c(T) \geq \sum_{k=1}^{N} d_k^2 - 2\sum_{k=1}^{N} d_k \widehat{T}(d_k) + \sum_{k=1}^{N} \left(\widehat{T}(d_k)\right)^2 = \sum_{k=1}^{N} \left(d_k - \widehat{T}(d_k)\right)^2 = c\left(\widehat{T}\right).$$

Par conséquent, on en déduit que, pour tout programme de transport T :

$$\boxed{c(T) \geq c\left(\widehat{T}\right).}$$

(4) *Interprétation probabiliste des inégalités* (1) *et* (2). Soit h une application croissante de $\mathbb{R}$ dans $\mathbb{R}$.

(a) Etablissons pour toute variable aléatoire discrète X ne prenant qu'un nombre fini de valeurs, l'inégalité : $E(Xh(X)) \geq E(X)E(h(X))$. Pour ce faire, considérons une variable aléatoire discrète X telle que $X(\Omega) = \{d_1, d_2, ..., d_N\}$, où les d_i ont été rangés en ordre strictement croissant. Avec les notations de la question précédente, on voit que $X(\Omega) = D$. Par la suite, on pose $p_k = P([X = d_k])$ et $h(d_k) = a_k$ pour tout $k \in [\![1, N]\!]$. Comme la fonction h est croissante et que les d_i sont rangés en ordre croissant, on voit que $a_1 \leq ... \leq a_N$. De plus, on trouve d'après le théorème de transfert que :

$$E(Xh(X)) = \sum_{k=1}^{N} d_k h(d_k) P([X = d_k]) = \sum_{k=1}^{N} d_k h(d_k) p_k = \sum_{k=1}^{N} d_k a_k p_k.$$

Comme la famille $([X = d_k])_{1 \leq k \leq N}$ est un système complet d'événements, on a $\sum_{k=1}^{N} p_k = 1$. Comme de plus tous les p_k sont ≥ 0 (vu que ce sont des probabilités), on obtient d'après la question (1)(b)

de la partie II que :

$$E(Xh(X)) = \sum_{k=1}^{N} d_k a_k p_k \geq \left(\sum_{k=1}^{N} p_k d_k\right)\left(\sum_{k=1}^{N} p_k a_k\right). \quad (*)$$

A noter que, dans la démonstration de la question (1)(b) de la partie II, on s'est juste servi de la croissance de la suite finie $(a_k)_{1\leq k\leq N}$ et non de la stricte croissance de cette suite pour établir la relation (1). En d'autres termes, l'inégalité (1) reste valable si l'on suppose seulement la suite $(a_k)_{1\leq k\leq N}$ croissante. Par ailleurs, on a par définition de l'espérance que :

$$E(X) = \sum_{k=1}^{N} d_k P([X = d_k]) = \sum_{k=1}^{N} p_k d_k. \quad (**)$$

D'après le théorème de transfert, on voit aussi que :

$$E(h(X)) = \sum_{k=1}^{N} h(d_k) P\left([X = d_k]\right) = \sum_{k=1}^{N} p_k a_k. \quad (***)$$

En reportant les égalités $(**)$ et $(***)$ dans l'inégalité $(*)$, on en déduit que :

$$\boxed{E(Xh(X)) \geq E(X)E(h(X)).}$$

(b) Montrons que le coefficient de corrélation linéaire de X et $h(X)$ est ≥ 0 lorsque les variances de X et $h(X)$ sont > 0. D'après la formule de Koenig-Huygens et la question précédente, on trouve que :

$$\operatorname{cov}(X, h(X)) = E(Xh(X)) - E(X)E(h(X)) \geq 0.$$

Comme les écart-types de X et $h(X)$ sont > 0 par hypothèse, on en déduit que :

$$\boxed{\rho(X, h(X)) = \frac{\operatorname{cov}(X, h(X))}{\sigma_X \sigma_{h(X)}} \geq 0.}$$

(c) Montrons que, si X est une variable aléatoire discrète suivant la loi uniforme sur $[\![1, N]\!]$ et si t est un élément de $\mathcal{E}_N$, alors on a : $E(h(X)t(X)) \leq E(h(X)\widehat{t}(X))$. Comme X suit la loi uniforme sur $[\![1, N]\!]$, on voit que $X(\Omega) = [\![1, N]\!]$ et $P([X = k]) = \frac{1}{N}$ pour tout $k \in [\![1, N]\!]$. D'après le théorème de transfert, on trouve que :

$$E(h(X)t(X)) = \sum_{k=1}^{N} h(k)t(k)P([X = k]) = \sum_{k=1}^{N} h(k)t(k)\frac{1}{N} = \frac{1}{N}\sum_{k=1}^{N} h(k)t(k).$$

De la même façon, on obtient que $E(h(X)\widehat{t}(X)) = \frac{1}{N}\sum_{k=1}^{N} h(k)\widehat{t}(k)$. Posons alors $d_k = h(k)$ et $a_k = k$ pour tout $k \in [\![1, N]\!]$. Comme la fonction h est croissante, on voit que la suite finie $(d_k)_{1\leq k\leq N}$ est croissante. Comme la suite $(a_k)_{1\leq k\leq N}$ est strictement croissante par construction, on obtient avec l'inégalité (2) que :

$$\frac{1}{N}\sum_{k=1}^{N} h(k)t(k) = \frac{1}{N}\sum_{k=1}^{N} h(k)a_{t(k)} \leq \frac{1}{N}\sum_{k=1}^{N} h(k)a_{\widehat{t}(k)} = \frac{1}{N}\sum_{k=1}^{N} h(k)\widehat{t}(k). \quad (*)$$

A noter que, dans la démonstration de l'inégalité (2), on s'est juste servi de la croissance de la suite finie $(d_k)_{1\leq k\leq N}$ et non de la stricte croissance de cette suite. En d'autres termes, l'inégalité (2) reste valable si l'on suppose seulement la suite $(d_k)_{1\leq k\leq N}$ croissante. En retraduisant l'inégalité $(*)$ en termes d'espérances, il s'ensuit que :

$$E(h(X)t(X)) = \frac{1}{N}\sum_{k=1}^{N} h(k)t(k) \leq \frac{1}{N}\sum_{k=1}^{N} h(k)\widehat{t}(k) = E(h(X)\widehat{t}(X)).$$

Par conséquent, on en déduit que :

$$\boxed{E(h(X)t(X)) \leq E(h(X)\widehat{t}(X)).}$$

Partie III. Transport optimal dans une situation aléatoire

Les définitions de fonction de transport et de coût de transport sont identiques à celles données dans le préambule de la partie I. Dans toute cette partie, U désigne une variable aléatoire vérifiant $U(\Omega) = [0, 1]$ et suivant la loi uniforme sur le segment $[0, 1]$. Soit Y une variable aléatoire admettant une densité f_Y nulle hors d'un segment $[\alpha, \beta]$ ($\alpha < \beta$) et dont la restriction à ce segment est continue et strictement positive. On note F_Y la fonction de répartition de Y. On suppose l'existence d'une fonction g de classe $\mathcal{C}^1$ sur $[0, 1]$, à valeurs dans $[\alpha, \beta]$, telle que la variable aléatoire $Z = g(U)$ suit la même loi que Y.

(1) Pour tout entier $N \geq 1$, on pose pour tout $\omega \in \Omega$:

$$X_N(\omega) = \begin{cases} \lfloor 1 + NU(\omega) \rfloor & \text{si} \quad 0 \leq U(\omega) < 1 \\ N & \text{si} \quad U(\omega) = 1 \end{cases} \qquad \text{et} \quad Y_N(\omega) = g\left(\frac{X_N(\omega)}{N}\right).$$

(a) Trouvons la loi de la variable aléatoire X_N. Comme U suit la loi uniforme sur $[0,1]$, on voit que, pour tout $\omega \in \Omega$, on a soit $0 \leq U(\omega) < 1$, soit $U(\omega) = 1$. Si $0 \leq U(\omega) < 1$, alors on a $0 \leq NU(\omega) < N$, ce qui entraine que $1 \leq 1 + NU(\omega) < N+1$, et donc $X_N(\omega) = \lfloor 1 + NU(\omega) \rfloor$ appartient à $[\![1, N]\!]$. Si maintenant $U(\omega) = 1$, alors $X_N(\omega) = N$ par construction. Dans tous les cas, on constate que $X_N(\omega)$ appartient à $[\![1, N]\!]$ pour tout $\omega \in \Omega$, et donc :

$$X_N(\Omega) \subset [\![1, N]\!].$$

De plus, pour tout $k \in [\![1, N-1]\!]$, on trouve par des calculs simples que :

$$P([X_N = k]) = P([\lfloor 1 + NU \rfloor = k]) = P([k \leq 1 + NU < k+1]) = P\left(\left[\frac{k-1}{N} \leq U < \frac{k}{N}\right]\right).$$

Comme U est une variable à densité, ceci nous donne que, pour tout $k \in [\![1, N-1]\!]$:

$$P([X_N = k]) = P\left(\left[\frac{k-1}{N} \leq U < \frac{k}{N}\right]\right) = F_U\left(\frac{k}{N}\right) - F_U\left(\frac{k-1}{N}\right).$$

Comme U suit la loi uniforme sur $[0,1]$ et que $\frac{k}{N}$, $\frac{k-1}{N}$ appartiennent à $[0,1]$ pour tout $k \in [\![1, N-1]\!]$, il s'ensuit que, pour tout $k \in [\![1, N-1]\!]$:

$$P([X_N = k]) = F_U\left(\frac{k}{N}\right) - F_U\left(\frac{k-1}{N}\right) = \frac{k}{N} - \frac{k-1}{N} = \frac{1}{N}.$$

Par ailleurs, comme $X_N(\Omega) \subset [\![1, N]\!]$ d'après ce qui précède, la famille $([X_N = k])_{1 \leq k \leq N}$ est un système complet d'événements, et donc :

$$P([X_N = N]) = 1 - \sum_{k=1}^{n-1} P([X_N = k]) = 1 - \sum_{k=1}^{n-1} \frac{1}{N} = 1 - \frac{N-1}{N} = \frac{1}{N}.$$

Comme $X_N(\Omega) \subset [\![1, N]\!]$ et que $P([X_N = k]) = \frac{1}{N}$ pour tout $k \in [\![1, N]\!]$, on en déduit que :

$$\boxed{X_N \hookrightarrow \mathcal{U}([\![1, N]\!]).}$$

(b) Etablissons l'existence d'un réel $\lambda > 0$ indépendant de N tel que : $\forall \omega \in \Omega$, $|Z(\omega) - Y_N(\omega)| \leq \frac{\lambda}{N}$. Comme la fonction g est de classe $\mathcal{C}^1$ sur $[0,1]$, sa dérivée g' est continue sur le segment $[0,1]$, et donc elle est bornée et atteint ses bornes sur $[0,1]$. Posons alors $\lambda = \max_{x \in [0,1]} |g'(x)| + 1$. Par construction, on voit que $\lambda > 0$ et de plus, on a $|g'(x)| \leq \lambda$ pour tout $x \in [0,1]$. D'après l'inégalité des accroissements finis, on obtient que, pour tous $x, y \in [0,1]$:

$$|g(x) - g(y)| \leq \lambda |x - y|.$$

Comme $U(\omega)$ et $\frac{X_N(\omega)}{N}$ appartiennent à $[0,1]$ pour tout $\omega \in \Omega$ d'après la question précédente, ceci entraine que, pour tout $\omega \in \Omega$:

$$\left| g(U(\omega)) - g\left(\frac{X_N(\omega)}{N}\right) \right| \leq \lambda \left| U(\omega) - \frac{X_N(\omega)}{N} \right|,$$

ce qui, par définition de Z et Y_N, se réécrit sous la forme suivante pour tout $\omega \in \Omega$:

$$|Z(\omega) - Y_N(\omega)| \leq \lambda \left| U(\omega) - \frac{X_N(\omega)}{N} \right|. \quad (*)$$

Pour tout $\omega \in \Omega$, on distingue deux cas. Si $0 \leq U(\omega) < 1$, alors on sait que $X_N(\omega) = \lfloor 1 + NU(\omega) \rfloor$. Comme $x - 1 < \lfloor x \rfloor \leq x$ pour tout $x \in \mathbb{R}$, on obtient que $NU(\omega) < X_N(\omega) \leq 1 + NU(\omega)$, ce qui entraine que $U(\omega) < \frac{X_N(\omega)}{N} \leq \frac{1}{N} + U(\omega)$, et donc :

$$\left| U(\omega) - \frac{X_N(\omega)}{N} \right| \leq \frac{1}{N}.$$

Par application de l'inégalité $(*)$, il s'ensuit que :

$$|Z(\omega) - Y_N(\omega)| \leq \lambda \left| U(\omega) - \frac{X_N(\omega)}{N} \right| \leq \frac{\lambda}{N}.$$

Si maintenant $U(\omega) = 1$, alors on voit que :

$$\left| U(\omega) - \frac{X_N(\omega)}{N} \right| = \left| 1 - \frac{N}{N} \right| = 0 \leq \frac{1}{N}.$$

Toujours par application de l'inégalité $(*)$, il s'ensuit que :

$$|Z(\omega) - Y_N(\omega)| \leq \lambda \left| U(\omega) - \frac{X_N(\omega)}{N} \right| \leq \frac{\lambda}{N}.$$

Dans tous les cas, on vient de montrer que $|U(\omega) - Y_N(\omega)| \leq \frac{\lambda}{N}$ pour tout $\omega \in \Omega$, et donc :

$$\boxed{\exists \lambda > 0,\ \forall \omega \in \Omega,\ |Z(\omega) - Y_N(\omega)| \leq \frac{\lambda}{N}.}$$

(c) Montrons que pour tout réel y, on a : $F_Y\big(y - \frac{\lambda}{N}\big) \leq P([Y_N < y])$. D'après la question précédente, on sait que $|Z(\omega) - Y_N(\omega)| \leq \frac{\lambda}{N}$ pour tout $\omega \in \Omega$, et donc :

$$-\frac{\lambda}{N} \leq Z(\omega) - Y_N(\omega) \leq \frac{\lambda}{N}.$$

En particulier, on a $Y_N(\omega) \leq Z(\omega) + \frac{\lambda}{N}$ pour tout $\omega \in \Omega$. Dès lors, pour tout $y \in \mathbb{R}$, on obtient que, si l'événement $[Z + \frac{\lambda}{N} < y]$ est réalisé, l'événement $[Y_N < y]$ l'est aussi, et donc on a l'inclusion $\big[Z + \frac{\lambda}{N} < y\big] \subset [Y_N < y]$. Par croissances des probabilités, on trouve que, pour tout $y \in \mathbb{R}$:

$$P\left(\left[Z + \frac{\lambda}{N} < y\right]\right) \leq P([Y_N < y]).$$

De plus, comme Y et Z suivent la même loi et que Y est à densité, la variable aléatoire Z est aussi à densité, et donc on a pour tout $y \in \mathbb{R}$:

$$F_Z\left(y - \frac{\lambda}{N}\right) = P\left(\left[Z \leq y - \frac{\lambda}{N}\right]\right) = P\left(\left[Z + \frac{\lambda}{N} \leq y\right]\right) = P\left(\left[Z + \frac{\lambda}{N} < y\right]\right) \leq P([Y_N < y]).$$

Comme Y et Z suivent la même loi, il s'ensuit que, pour tout $y \in \mathbb{R}$:

$$F_Y\left(y - \frac{\lambda}{N}\right) = F_Z\left(y - \frac{\lambda}{N}\right) \leq P([Y_N < y]).$$

Par conséquent, on en déduit que, pour tout $y \in \mathbb{R}$:

$$\boxed{F_Y\left(y - \frac{\lambda}{N}\right) \leq P([Y_N < y]).}$$

(2) Pour tout $k \in [\![1, N]\!]$, on pose $t_N(k) = g\big(\frac{k}{N}\big)$. On définit alors $\widehat{t_N}$ à partir de t_N, comme $\widehat{t}$ à partir de t dans la question (2) de la partie II.

(a) Etablissons pour tout $k \in [\![1, N]\!]$ les inégalités : $F_Y\big(\widehat{t_N}(k) - \frac{\lambda}{N}\big) \leq P([Y_N < \widehat{t_N}(k)]) < \frac{k}{N}$. D'après la question précédente, on sait que, pour tout $y \in \mathbb{R}$:

$$F_Y\left(y - \frac{\lambda}{N}\right) \leq P([Y_N < y]).$$

En particulier, on obtient en remplaçant y par $\widehat{t_N}(k)$ que, pour tout $k \in [\![1, N]\!]$:

$$F_Y\left(\widehat{t_N}(k) - \frac{\lambda}{N}\right) \leq P\left(\left[Y_N < \widehat{t_N}(k)\right]\right).$$

ce qui nous donne la première inégalité à démontrer. Pour la deuxième, on commence par remarquer que, pour $k = 1$:

$$P\left(\left[Y_N < \widehat{t_N}(1)\right]\right) = P\left(\left[g\left(\frac{X_N}{N}\right) < \widehat{t_N}(1)\right]\right).$$

Comme X_N suit la loi uniforme sur $[\![1, N]\!]$ d'après la question (1)(a) de la partie III, $\frac{X_N(\omega)}{N}$ est toujours de la forme $\frac{i}{N}$ avec $i \in [\![1, N]\!]$ pour tout $\omega \in \Omega$, et donc $g\left(\frac{X_N}{N}\right)$ est à valeurs dans l'ensemble $\{g(\frac{1}{N}), ..., g(\frac{N}{N})\} = \{t_N(1), ..., t_N(N)\}$. Comme les $\widehat{t_N}(i)$ sont obtenus à partir des $t_N(i)$ en les rangeant en ordre croissant, on voit que $\{t_N(1), ..., t_N(N)\} = \{\widehat{t_N}(1), ..., \widehat{t_N}(N)\}$, et donc $g\left(\frac{X_N}{N}\right)$ est à valeurs dans l'ensemble $\{\widehat{t_N}(1), ..., \widehat{t_N}(N)\}$. Comme les $\widehat{t_N}(i)$ sont rangés en ordre croissant, on constate que $\widehat{t_N}(1)$ est le plus petit élément de $\{\widehat{t_N}(1), ..., \widehat{t_N}(N)\}$. En particulier, il s'ensuit que $g\left(\frac{X_N}{N}\right) \geq \widehat{t_N}(1)$, et donc :

$$P\left(\left[Y_N < \widehat{t_N}(1)\right]\right) = P\left(\left[g\left(\frac{X_N}{N}\right) < \widehat{t_N}(1)\right]\right) = 0 < \frac{1}{N}.$$

Dès lors, l'inégalité de droite à démontrer est vraie pour $k=1$. A présent, fixons un entier $k \in [\![2,N]\!]$. Comme précédemment, on trouve que :

$$P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) = P\left(\left[g\left(\frac{X_N}{N}\right) < \widehat{t_N}(k)\right]\right). \quad (*)$$

Comme $\{t_N(1),...,t_N(N)\} = \{\widehat{t_N}(1),...,\widehat{t_N}(N)\}$, il existe pour tout $r \in [\![1,N]\!]$ un indice $i_r \in [\![1,N]\!]$ tel que $\widehat{t_N}(r) = t_N(i_r) = g\left(\frac{i_r}{N}\right)$. Comme $t_N(i) = g\left(\frac{i}{N}\right)$ pour tout $i \in [\![1,N]\!]$ par construction, il existe donc des indices $i_1, i_2,..., i_N$ tels que $\widehat{t_N}(r) = t_N(i_r) = g\left(\frac{i_r}{N}\right)$ pour tout $r \in [\![1,N]\!]$. A noter que, comme les $\widehat{t_N}(i)$ sont obtenus à partir des $t_N(i)$ en les rangeant en ordre croissant, les listes $t_N(1),...,t_N(N)$ et $\widehat{t_N}(1),...,\widehat{t_N}(N)$ comptabilisent les mêmes réels comptés le même nombre de fois. En particulier, on peut supposer que les indices $i_1,...,i_N$ sont deux à deux distincts, ce que l'on fera désormais. Par construction, on voit que :

$$g\left(\frac{X_N}{N}\right)(\Omega) = \{t_N(1),...,t_N(N)\} = \{\widehat{t_N}(1),...,\widehat{t_N}(N)\} = \left\{g\left(\frac{i_1}{N}\right),...,g\left(\frac{i_N}{N}\right)\right\}.$$

Fixons alors un élément $\omega \in \Omega$ tel que $g\left(\frac{X_N(\omega)}{N}\right) < \widehat{t_N}(k)$, et montrons que $X_N(\omega)$ appartient à $\{i_1, i_2, ..., i_{k-1}\}$. Pour ce faire, on raisonne par l'absurde et on suppose que $X_N(\omega) \notin \{i_1, i_2, ..., i_{k-1}\}$. Alors il existe un indice $r \in [\![k,N]\!]$ tel que $g\left(\frac{X_N(\omega)}{N}\right) = g\left(\frac{i_r}{N}\right) = t_N(i_r) = \widehat{t_N}(r)$. Comme les $\widehat{t_N}(i)$ sont rangés en ordre croissant, il s'ensuit que $g\left(\frac{X_N(\omega)}{N}\right) \geq \widehat{t_N}(k)$, ce qui est impossible car $g\left(\frac{X_N(\omega)}{N}\right) < \widehat{t_N}(k)$ par hypothèse, et donc on a bien $X_N(\omega) \in \{i_1, i_2, ..., i_{k-1}\}$. Partant de là, on a obtenu l'inclusion suivante :

$$\left[g\left(\frac{X_N}{N}\right) < \widehat{t_N}(k)\right] \subset \bigcup_{r=1}^{k-1} [X_N = i_r].$$

Par croissance des probabilités, ceci nous donne avec la relation $(*)$ que :

$$P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) = P\left(\left[g\left(\frac{X_N}{N}\right) < \widehat{t_N}(k)\right]\right) \leq P\left(\bigcup_{r=1}^{k-1}[X_N = i_r]\right).$$

Par incompatibilité, ceci entraine que :

$$P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) \leq \sum_{r=1}^{k-1} P([X_N = i_r]).$$

Comme X_N suit la loi uniforme sur $[\![1,N]\!]$ d'après la question (1)(a) de la partie III, il vient :

$$P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) \leq \sum_{r=1}^{k-1} P([X_N = i_r]) = \frac{k-1}{N},$$

et donc $P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) < \frac{k}{N}$ pour tout $k \in [\![2,N]\!]$, ce qui conclut la démonstration de l'inégalité de droite. Par conséquent, on en déduit que, pour tout $k \in [\![1,N]\!]$:

$$\boxed{F_Y\left(\widehat{t_N}(k) - \frac{\lambda}{N}\right) \leq P\left(\left[Y_N < \widehat{t_N}(k)\right]\right) < \frac{k}{N}.}$$

(b) On note F_Y^{-1} la fonction réciproque de la restriction à $[\alpha,\beta]$ de la fonction F_Y. Montrons que, pour tout entier $N \geq 1$, on a :

$$\frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right).$$

Montrons tout d'abord que F_Y^{-1} est une bijection de $[0,1]$ sur $[\alpha,\beta]$. Pour ce faire, on désigne par G la restriction de F_Y à $[\alpha,\beta]$. Comme f_Y est continue et strictement positive sur $[\alpha,\beta]$ par hypothèse, la fonction G est de classe $\mathcal{C}^1$ sur $[\alpha,\beta]$ et de plus, on a $G'(x) = f_Y(x) > 0$ pour tout $x \in [\alpha,\beta]$. En particulier, la fonction G est continue et strictement croissante sur $[\alpha,\beta]$, et donc G réalise une bijection de $[\alpha,\beta]$ sur $G([\alpha,\beta])$ d'après le théorème de la bijection. Comme G est continue, l'ensemble $G([\alpha,\beta])$ est un intervalle d'après le théorème des valeurs intermédiaires. Comme f_Y est nulle en dehors de $[\alpha,\beta]$, la fonction F_Y est constante sur les intervalles $]-\infty,\alpha]$ et $[\beta,+\infty[$. Comme F_Y tend vers 0 en $-\infty$ et vers 1 en $+\infty$ en tant que fonction de répartition, on obtient que F_Y est égale à 0 sur $]-\infty,\alpha]$ et égale à 1 sur $[\beta,+\infty[$. En particulier, on voit que $G(\alpha) = F_Y(\alpha) = 0$ et

$G(\beta) = F_Y(\beta) = 1$. Comme G est croissante sur $[\alpha, \beta]$ et que $G([\alpha, \beta])$ est un intervalle, il s'ensuit que $G([\alpha, \beta]) = [0, 1]$. En particulier, la fonction G réalise une bijection de $[\alpha, \beta]$ dans $[0, 1]$, et donc F_Y^{-1} est bien une bijection de $[0, 1]$ sur $[\alpha, \beta]$.

Passons maintenant à l'inégalité à démontrer. D'après la question (2)(b) de la partie II, on obtient en prenant $D = A = \{1, 2, ..., N\}$ (et donc $d_k = a_k = k$ pour tout $k \in [\![1, N]\!]$) que :

$$\sum_{k=1}^{N} k t_N(k) = \sum_{k=1}^{N} d_k a_{t_N(k)} \leq \sum_{k=1}^{N} d_k a_{\widehat{t_N}(k)} = \sum_{k=1}^{N} k \widehat{t_N}(k).$$

Comme $t_N(k) = g\left(\frac{k}{N}\right)$ pour tout $k \in [\![1, N]\!]$ par hypothèse, ceci nous donne que :

$$\sum_{k=1}^{N} k g\left(\frac{k}{N}\right) \leq \sum_{k=1}^{N} k \widehat{t_N}(k).$$

En divisant le tout par N^2, ceci entraine par linéarité de la somme que :

$$\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} \widehat{t_N}(k). \quad (*)$$

Pour tout $k \in [\![1, N]\!]$, on sait d'après la question précédente que $0 \leq F_Y(\widehat{t_N}(k) - \frac{\lambda}{N}) < \frac{k}{N} \leq 1$. Comme g est à valeurs dans $[\alpha, \beta]$ par hypothèse, on voit que $t_N(k) = g(\frac{k}{N})$ appartient à $[\alpha, \beta]$ pour tout $k \in [\![1, N]\!]$. Comme les $\widehat{t_N}(k)$ correspondent aux $t_N(k)$ rangés en ordre croissant, on obtient que $\widehat{t_N}(k)$ appartient aussi à $[\alpha, \beta]$ pour tout $k \in [\![1, N]\!]$. Comme $\lambda > 0$, ceci entraine que $\widehat{t_N}(k) - \frac{\lambda}{N} \leq \beta$ pour tout $k \in [\![1, N]\!]$. On distingue deux cas.

Supposons tout d'abord que $\widehat{t_N}(k) - \frac{\lambda}{N}$ appartient à $[\alpha, \beta]$. Comme F_Y est strictement croissante sur $[\alpha, \beta]$, la fonction F_Y^{-1} est strictement croissante sur $[0, 1]$ d'après le théorème de la bijection. Mais comme $F_Y(\widehat{t_N}(k) - \frac{\lambda}{N}) < \frac{k}{N}$ d'après la question précédente, il s'ensuit que :

$$\widehat{t_N}(k) - \frac{\lambda}{N} < F_Y^{-1}\left(\frac{k}{N}\right).$$

Supposons maintenant que $\widehat{t_N}(k) - \frac{\lambda}{N} < \alpha$. Comme F_Y^{-1} est à valeurs dans $[\alpha, \beta]$, on voit que :

$$\widehat{t_N}(k) - \frac{\lambda}{N} < \alpha \leq F_Y^{-1}\left(\frac{k}{N}\right).$$

Dans tous les cas, on vient de montrer que, pour tout $k \in [\![1, N]\!]$:

$$\boxed{\widehat{t_N}(k) < F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}.}$$

Comme $\frac{k}{N} \geq 0$ pour tout $k \in [\![1, N]\!]$, il s'ensuit par sommation que :

$$\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} \widehat{t_N}(k) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right). \quad (**)$$

Dès lors, on obtient en associant les inégalités $(*)$ et $(**)$ que :

$$\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} \widehat{t_N}(k) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right).$$

Par conséquent, on en déduit que :

$$\boxed{\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right).}$$

(c) Montrons l'inégalité $E(Ug(U)) \leq E(UF_Y^{-1}(U))$. D'après la question précédente, on sait que :

$$\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right).$$

Par linéarité de la somme, ceci nous donne que :

$$\frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} g\left(\frac{k}{N}\right) \leq \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N} F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N^3}\sum_{k=1}^{N} k,$$

ce que l'on peut réécrire sous la forme :

$$\frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}g\left(\frac{k}{N}\right)\leq\frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}F_Y^{-1}\left(\frac{k}{N}\right)+\frac{\lambda}{N^3}\frac{N(N+1)}{2}.\quad(*)$$

Comme les fonctions $x\longmapsto xg(x)$ et $x\longmapsto xF_Y^{-1}(x)$ sont continues sur $[0,1]$, le théorème sur les sommes de Riemann entraine que :

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}g\left(\frac{k}{N}\right)=\int_0^1 xg(x)dx\quad\text{et}\quad\lim_{N\to+\infty}\frac{1}{N}\sum_{k=1}^{N}\frac{k}{N}F_Y^{-1}\left(\frac{k}{N}\right)=\int_0^1 xF_Y^{-1}(x)dx.$$

Comme $\frac{\lambda}{N^3}\frac{N(N+1)}{2}$ tend vers 0 quand N tend vers $+\infty$, il s'ensuit par passage à la limite quand N tend vers $+\infty$ dans l'inégalité $(*)$ que :

$$\int_0^1 xg(x)dx\leq\int_0^1 xF_Y^{-1}(x)dx.\quad(**)$$

Comme U suit la loi uniforme sur $[0,1]$, on obtient avec le théorème de transfert que :

$$E(Ug(U))=\int_{-\infty}^{+\infty}xg(x)f_U(x)dx=\int_0^1 xg(x)f_U(x)dx=\int_0^1 xg(x)dx.$$

De la même façon, on voit par transfert que $E\left(UF_Y^{-1}(U)\right)=\int_0^1 xF_Y^{-1}(x)dx$. Par conséquent, on en déduit en retraduisant l'inégalité $(**)$ en termes d'espérances que :

$$\boxed{E(Ug(U))\leq E(UF_Y^{-1}(U)).}$$

(3) (a) Parmi les fonctions de transport de classe C^1 de U vers la loi de Y, trouvons une fonction de transport T^* de coût minimal. Pour cela, on conserve les notations de la question (2)(b) de la partie III, et on désigne par G la restriction de F_Y à l'intervalle $[\alpha,\beta]$. D'après la question (2)(b) de la partie III, on sait que G est une bijection de $[\alpha,\beta]$ sur $[0,1]$. Comme G est de classe $\mathcal{C}^1$ sur $[\alpha,\beta]$ et que $G'(x)=f_Y(x)>0$ pour tout $x\in[\alpha,\beta]$, sa bijection réciproque G^{-1} est de classe $\mathcal{C}^1$ sur $[0,1]$. On pose alors :

$$\boxed{V=G^{-1}(U).}$$

Comme G est une bijection de $[\alpha,\beta]$ sur $[0,1]$, la fonction G^{-1} est une bijection de $[0,1]$ sur $[\alpha,\beta]$. En particulier, elle est à valeurs dans $[\alpha,\beta]$, ce qui entraine que $V(\Omega)=G^{-1}(U)(\Omega)\subset[\alpha,\beta]$, et donc $F_V(x)=0$ si $x<\alpha$ et $F_V(x)=1$ si $x>\beta$. Comme F_Y est nulle sur $]-\infty,\alpha]$ et égale à 1 sur $[\beta,+\infty[$ d'après les arguments de la question (2)(b) de la partie III, on obtient que $F_V(x)=0=F_Y(x)$ si $x<\alpha$ et $F_V(x)=1=F_Y(x)$ si $x>\beta$ $(*)$. De plus, comme G est strictement croissante sur $[\alpha,\beta]$, on a pour tout $x\in[\alpha,\beta]$:

$$F_V(x)=P\left([G^{-1}(U)\leq x]\right)=P\left([U\leq G(x)]\right).$$

Comme G est est une bijection de $[\alpha,\beta]$ sur $[0,1]$, on voit que $G(x)$ appartient à $[0,1]$ pour tout $x\in[\alpha,\beta]$. Mais comme U suit la loi uniforme sur $[0,1]$, il s'ensuit que, pour tout $x\in[\alpha,\beta]$:

$$F_V(x)=P\left([U\leq G(x)]\right)=F_U(G(x))=G(x)=F_Y(x).\quad(**)$$

D'après les égalités $(*)$ et $(**)$, il s'ensuit que $F_V(x)=F_Y(x)$ pour tout $x\in\mathbb{R}$, et donc $V=G^{-1}(U)$ suit la même loi que Y. Par conséquent, on en déduit que :

$$\boxed{G^{-1}\text{ est une fonction de transport de classe }\mathcal{C}^1\text{ de }U\text{ vers la loi de }Y.}$$

Reste à montrer que G^{-1} est une fonction de transport de coût minimal. Par définition du coût de transport et par linéarité de l'espérance, on voit que :

$$C\left(G^{-1}\right)=E\left(\left(U-G^{-1}(U)\right)^2\right)=E\left(U^2\right)-2E\left(UG^{-1}(U)\right)+E\left(\left(G^{-1}(U)\right)^2\right).$$

Comme $G^{-1}(U)$ et Y suivent la même loi d'après ce qui précède, leurs moments d'ordre 2 sont égaux, ce qui entraine que :

$$C\left(G^{-1}\right)=E\left(U^2\right)-2E\left(UG^{-1}(U)\right)+E\left(Y^2\right).$$

Considérons une fonction de transport g de classe $\mathcal{C}^1$ de la variable aléatoire U vers la loi de Y. Comme $g(U)$ et Y suivent la même loi par définition d'une fonction de transport, on trouve comme précédemment que :

$$C\left(g\right)=E\left(U^2\right)-2E\left(Ug\left(U\right)\right)+E\left(Y^2\right).$$

Par différence, il s'ensuit que :

$$C(g) - C\left(G^{-1}\right) = 2\left[E\left(UG^{-1}(U)\right) - E\left(Ug(U)\right)\right].$$

Comme $E\left(UG^{-1}(U)\right) = E(UF_Y^{-1}(U)) \geq E\left(Ug(U)\right)$ d'après la question (2)(c) de la partie III, ceci entraine que $C(g) - C\left(G^{-1}\right) \geq 0$, et donc $C(g) \geqslant C\left(G^{-1}\right)$. Par conséquent, on en déduit que :

$\boxed{G^{-1} \text{ est une fonction de transport de coût minimal parmi les fonctions de transport de classe } \mathcal{C}^1 \text{ de } U \text{ vers la loi de } Y.}$

(b) On suppose que $Y = |4U-2|$. Déterminons T^* et $C(T^*)$. Pour ce faire, on commence par déterminer la fonction de répartition F_Y. Comme U suit la loi uniforme sur $[0,1]$, on voit que $0 \leq U \leq 1$, ce qui entraine que $-2 \leq 4U-2 \leq 2$, et donc $0 \leq Y = |4U-2| \leq 2$. En particulier, on a $Y(\Omega) \subset [0,2]$, et donc $F_Y(x) = 0$ si $x < 0$ et $F_Y(x) = 1$ si $x > 2$. De plus, pour tout $x \in [0,2]$, on trouve par des calculs simples que :

$$F_Y(x) = P\left([|4U-2| \leq x]\right) = P\left([-x \leq 4U-2 \leq x]\right) = P\left(\left[\frac{1}{2} - \frac{x}{4} \leq U \leq \frac{1}{2} + \frac{x}{4}\right]\right).$$

Comme x appartient à $[0,2]$, on voit que $\frac{x}{4}$ appartient à $\left[0, \frac{1}{2}\right]$, et donc $\frac{1}{2} + \frac{x}{4}$ et $\frac{1}{2} - \frac{x}{4}$ appartiennent à $[0,1]$. Dès lors, comme U suit la loi uniforme sur $[0,1]$, il s'ensuit que, pour tout $x \in [0,2]$:

$$F_Y(x) = F_U\left(\frac{1}{2} + \frac{x}{4}\right) - F_U\left(\frac{1}{2} - \frac{x}{4}\right) = \frac{1}{2} + \frac{x}{4} - \left(\frac{1}{2} - \frac{x}{4}\right) = \frac{x}{2}.$$

Par conséquent, on en déduit que :

$$\boxed{F_Y(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \dfrac{x}{2} & \text{si} \quad x \in [0,2] \\ 1 & \text{si} \quad x > 2 \end{cases}.}$$

En particulier, on voit que Y est une variable à densité, de densité f_Y égale à $\frac{1}{2}$ sur $[0,2]$ et nulle partout ailleurs, et donc f_Y est continue et strictement positive sur $[0,2]$ et nulle en dehors de $[0,2]$. De plus, avec les notations des questions précédentes, on a pour tout $x \in [0,2]$:

$$G(x) = F_Y(x) = \frac{x}{2}.$$

Dès lors, il s'ensuit que $G^{-1}(x) = 2x$ pour tout $x \in [0,1]$. Par conséquent, on en déduit avec la question précédente que, pour tout $x \in [0,1]$:

$$\boxed{T^*(x) = G^{-1}(x) = 2x.}$$

Reste à calculer $C(T^*)$. D'après le résultat ci-dessus, on voit que $T^*(U) = 2U$, et donc :

$$C(T^*) = E\left((U - T^*(U))^2\right) = E\left((U-2U)^2\right) = E\left(U^2\right).$$

Comme U suit la loi uniforme sur $[0,1]$, la formule de Koenig-Huygens nous donne que :

$$C(T^*) = E(U^2) = V(U) + E(U)^2 = \frac{1}{12} + \frac{1}{4} = \frac{1}{3}.$$

Par conséquent, on en déduit que :

$$\boxed{C(T^*) = \frac{1}{3}.}$$