Sujet 2 : type mixte EDHEC - ESSEC - HEC

Corrigé *Total sur 59 points - dont rédaction/présentation/clarté : 3 points*

Dans tout le sujet, concernant les codes Python, on supposera les importations suivantes faites :

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

Exercice

25 points

1. On pose $A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & \frac{3}{2} \\ 2 & 3 & -\frac{2}{3} & 1 \end{pmatrix}$

a. Déterminer le rang de A $- 3I_4$ où I_4 est la matrice identité d'ordre 4 *4 points*
En déduire une valeur propre λ_1 pour A et le sous-espace propre associé.

$$A - 3I_4 = \begin{pmatrix} -3 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 3/2 \\ 2 & 3 & -2/3 & -2 \end{pmatrix}$$

on note C_1, C_2, C_3, C_4 (dans cet ordre) les colonnes de $A - 3I_4$ rg $(A - 3I_4) = \dim \operatorname{Vect}(C_1, C_2, C_3, C_4)$
comme $C_1 = -3C_3$, $\operatorname{Vect}(C_1, C_2, C_3, C_4) = \operatorname{Vect}(C_2, C_3, C_4)$
soit $\lambda_2, \lambda_3, \lambda_4$ trois réels tels que $\lambda_2 C_2 + \lambda_3 C_3 + \lambda_4 C_4 = 0$
la première ligne donne $\lambda_3 = 0$, la troisième donne $\lambda_4 = 0$ et donc finalement $\lambda_2 = \lambda_3 = \lambda_4 = 0$
donc (C_2, C_3, C_4) est libre, et génératrice par définition du Vect
c'est donc une base de $\operatorname{Vect}(C_2, C_3, C_4) = \operatorname{Vect}(C_1, C_2, C_3, C_4)$
donc rg $(A - 3I_4) = 3$ donc $A - 3I_4$ n'est pas inversible et donc 3 est une valeur propre de A
en résolvant $(A - 3I_4)\, X$ d'inconnue $X = {}^{t}(x_1, x_2, x_3, x_4)$: $E_3(A) = \operatorname{Vect}\left({}^{t}(1, 0, 3, 0)\right)$ de dimension 1.

b. Calculer $A \begin{pmatrix} 3 \\ -2 \\ 6 \\ -4 \end{pmatrix}$

En déduire une valeur propre λ_2 pour A et déterminer le sous-espace propre associé. *2 points*

Si $V = {}^{t}(3, -2, 6, 4), AV = 2V$ donc $2 \in \operatorname{Sp}(A)$
en résolvant le système $AX = 2X$ d'inconnue $X = {}^{t}(x_1, x_2, x_3, x_4)$
on trouve $X = x_1 {}^{t}(1, -2/3, 2, -4/3) = \dfrac{x_1}{3} V$ donc $E_2(A) = \operatorname{Vect}(V)$

c. Soit f l'endomorphisme de $\mathbb{R}^4$ représenté par la matrice A suivant la base canonique (e_1, e_2, e_3, e_4) de $\mathbb{R}^4$

i. Déterminer une base de l'image de f *2 points*

$\mathrm{Im}(f) = \mathrm{Vect}\left(f(e_1), f(e_2), f(e_3), f(e_4)\right) = \mathrm{Vect}\left(f(e_2), f(e_3), f(e_4)\right)$ car $f(e_2) = \frac{2}{3} f(e_1)$

Si $\lambda_2, \lambda_3, \lambda_4$ sont trois réels tels que $\sum_{k-2}^{4} \lambda_k f(e_k) = 0$, la première ligne donne $\lambda_3 = 0$, la deuxième ligne donne $\lambda_4 = 0$ et donc $\lambda_2 = \lambda_3 = \lambda_4 = 0$

$\left(f(e_2), f(e_3), f(e_4)\right)$ est donc libre, et génératrice par définition du Vect, c'est une base de $\mathrm{Im}(f)$

ii. Calculer $f(3e_1 - 2e_2)$ *0,5 point*

$f(3e_1 - 2e_2) = 3 \times 2e_4 - 2 \times e_4 = 0$

iii. En déduire le noyau de f, sans résoudre de système linéaire. *1 point*

D'après le théorème du rang : $\dim \mathrm{Ker}(f) = 4 - \dim \mathrm{Im}(f) = 1$ et $3e_1 - 2e_2 \in \mathrm{Ker}(f)$
donc $(3e_1 - 2e_2)$ est une base de $\mathrm{Ker}(f)$ (famille libre à 1 élément dans un espace de dimension 1)

iv. En déduire une valeur propre λ_3 de A et le sous-espace propre associé. *0,5 point*

$\lambda_3 = 0$ est valeur propre de A et $E_0(A) = \mathrm{Vect}\left({}^t(3, -2, 0, 0)\right) = \mathrm{Vect}((1, -2/3, 0, 0))$

d. Montrer que -1 est une valeur propre de A et déterminer le sous-espace propre associé. *2 points*

On résout $(A + I_4)X = 0_{1,4}$ et on trouve $E_{-1}(A) = \mathrm{Vect}\left({}^t(1, -8/3, -1, 8/3)\right)$

e. Montrer qu'il existe une matrice inversible P de première ligne ne contenant que des 1 , et telle que :

$$A = PDP^{-1} \quad \text{avec } D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$ *3 points*

Soit $P = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2/3 & -2/3 & -8/3 \\ 3 & 2 & 0 & -1 \\ 0 & -4/3 & 0 & 8/3 \end{pmatrix}$

Les colonnes de P sont la concaténation de familles libres de sous-espaces propres différents, donc forment une famille libre et donc une base (4 élements dans une espace de dimension 4) donc P est inversible.

et avec $D = \mathrm{diag}(3, 2, 0, -1)$ par calcul : $AP = PD = \begin{pmatrix} 3 & 2 & 0 & -1 \\ 0 & -4/3 & 0 & 8/3 \\ 9 & 4 & 0 & 1 \\ 0 & -8/3 & 0 & -8/3 \end{pmatrix}$ et donc $A = PDP^{-1}$

On considère le système différentiel d'ordre 2 :

$$S : \begin{cases} 2x'' & +3y'' & -4x' & -6y' & -6x & -9y & = 0 \\ -6x'' & -6y'' & +14x' & +15y' & +12x & +18y & = 0 \end{cases}$$

d'inconnues x, y, fonctions de classe $\mathscr{C}^2$ sur $\mathbb{R}$

et, pour tout $t \in \mathbb{R}$: $X(t) = \begin{pmatrix} x(t) \\ y(t) \\ x'(t) \\ y'(t) \end{pmatrix}$ et $X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ x''(t) \\ y''(t) \end{pmatrix}$

2. En utilisant des opérations sur S, montrer que S est équivalent à l'équation : *2 points*

$$S_D : \forall t \in \mathbb{R}, \quad X'(t) = AX(t)$$

En faisant les opérations $L_2 \leftarrow L_1 + \frac{1}{2}L_2$, $\quad S \Leftrightarrow \begin{cases} 2x'' + 3y'' &= 4x' + 6y' + 6x + 9y \\ x'' &= 3x' + \frac{3}{2}y' \end{cases}$

puis en faisant $L_1 \leftarrow L_1 - 2L_2$: $\quad S \Leftrightarrow \begin{cases} 3y'' &= 6x + 9y - 2x' + 3y' \\ x'' &= 3x' + \frac{3}{2}y' \end{cases}$

donc $S \Leftrightarrow \begin{cases} x'' &= 3x' + \frac{3}{2}y' \\ y'' &= 2x + 3y - \frac{2}{3}x' + y' \end{cases}$ et donc $S \Leftrightarrow X' = AX$

3. Résoudre le système S_D *2,5 points*

En posant $Y(t) = P^{-1}X(t), Y'(t) = P^{-1}X'(t)$, pour $t \in \mathbb{R}, X' = AX \Leftrightarrow Y' = DY$ (cf. cours), qui se résout facilement (système diagonal)

en posant $Y(t) = {}^{t}(y1(t) \quad y2(t) \quad y3(t) \quad y4(t))$ pour $t \in \mathbb{R}, Y' = DY$ si et seulement si il existe quatre réels a, b, c, d tels que $\forall t \in \mathbb{R}, y_1(t) = ae^{3t}, \quad y_2(t) = be^{2t}, \quad y_3(t) = c, \quad y_4(t) = de^{-t}$

et donc X est solution de $X' = AX$ si, et seulement si il existe quatre réels a, b, c, d tels que :

$\forall t \in \mathbb{R}, X(t) = PY(t) = \begin{pmatrix} ae^{3t} + be^{2t} + c + de^{-t} \\ -\frac{2}{3}be^{2t} - \frac{2}{3}c - \frac{8}{3}de^{-t} \\ 3ae^{3t} + 2ce^{2t} - de^{-t} \\ -\frac{4}{3}be^{2t} + \frac{8}{3}de^{-t} \end{pmatrix}$ soit $\begin{cases} x(t) &= ae^{3t} + be^{2t} + c + de^{-t} \\ y(t) &= -\frac{2}{3}be^{2t} - \frac{2}{3}c - \frac{8}{3}de^{-t} \end{cases}$

4. Déterminer l'unique couple solution (x_0, y_0) de S vérifiant : *2,5 points*

$$x_0(0) = 0, \quad y_0(0) = 2, \quad x_0'(0) = 1, \quad y_0'(0) = -\frac{8}{3}$$

Quelles sont les limites de x_0, x_0', y_0 et y_0' en $+\infty$?

On résout le système $\begin{cases} a + b + c + d &= 0 \\ 3a + 2b - d &= 0 \\ -\frac{1}{3}(2b + 2c + 8d) &= 0 \\ -\frac{1}{3}(4b + 8d) &= 0 \end{cases}$ et on trouve $\forall t \in \mathbb{R}, x_0(t) = 1 - e^{-t}$ et $y_0(t) = -\frac{2}{3} + \frac{8}{3}e^{-t}$

et donc $\lim\limits_{+\infty} x = 1, \quad \lim\limits_{+\infty} y = -\frac{2}{3}, \quad \lim\limits_{+\infty} x' = \lim\limits_{+\infty} y' = 0$

5. Soit X une solution de S_D vérifiant $x(0) = a, y(0) = b, x'(0) = c$ et $y'(0) = d$ *3 points*

Montrer que x et y admettent une limite finie en $+\infty$ si et seulement si :

$$8c + 3d = 0 \quad \text{et} \quad 2a + 3b + 2c + 3d = 0$$

Attention aux notations ici, les a, b, c et d ne sont pas les mêmes que lors des questions précédentes. On considère une solution qui d'après la question **3.** s'écrit :

$\forall t \in \mathbb{R}, \begin{cases} x(t) &= \alpha e^{3t} + \beta e^{2t} + \gamma + \delta e^{-t} \\ y(t) &= -\frac{2}{3}\beta e^{2t} - \frac{2}{3}\gamma - \frac{8}{3}\delta e^{-t} \end{cases}$

alors x et y admettent une limite en $+\infty \Leftrightarrow \alpha = \beta = 0$

et donc avec les conditions initiales, on écrit les liens avec les coefficients $\alpha + \beta + \gamma + \delta = a \ldots$ puis on résout pour déterminer α et β et on obtient le résultat en écrivant $\alpha = \beta = 0$

Problème - une propriété limite des lois de Pareto

31 points

Question préliminaire

Soit g une fonction continue sur un intervalle I, à valeurs réelles.

1. **a.** Montrer que pour tout α et β dans I tels que $\alpha < \beta$, *1 point*

$$\frac{1}{\beta-\alpha}\int_\alpha^\beta g(t)\,\mathrm{d}t = \int_0^1 g\big(\alpha+(\beta-\alpha)x\big)\,\mathrm{d}x$$

Le changement de variable $t = \alpha + (\beta-\alpha)x$ donne directement le résultat :
$\displaystyle\frac{1}{\beta-\alpha}\int_\alpha^\beta g(t)dt = \int_0^1 g(\alpha+(\beta-\alpha)x)dx$

b. Soit a, b, c, d dans I tels que $a < c < d < b$ *3 points*

On suppose g décroissante sur I, établir l'encadrement :

$$\frac{1}{b-c}\int_c^b g(t)\,\mathrm{d}t \leqslant \frac{1}{d-c}\int_c^d g(t)\,\mathrm{d}t \leqslant \frac{1}{d-a}\int_a^d g(t)\,\mathrm{d}t$$

Avec l'égalité ci-dessus, l'encadrement cherché est équivalent à :

$$\int_0^1 g(c+(b-c)x)dx \leqslant \int_0^1 g(c+(d-c)x)dx \leqslant \int_0^1 g(a+(d-a)x)dx$$

Or on a $a < c < d < b$ donc pour tout $x \in [0;1] (x \geqslant 0)$:

$$d < b \Longrightarrow d-c \leqslant b-c \Longrightarrow (d-c)x \leqslant (b-c)x \Longrightarrow c+(d-c)x \leqslant c+(b-c)x$$

et par décroissance de $g, g(c+(b-c)x) \leqslant g(c+(d-c)x)$
en intégrant l'inégalité sur $[0;1]$ on obtient l'inégalité de gauche.
d'autre part on a $[c+(d-c)x]-[a+(d-a)x] = (c-a)+(a-c)x = (c-a)(1-x) \geqslant 0$
sur $[0;1]$, car $1-x \geqslant 0$ et $c-a \geqslant 0$
d'où $c+(d-c)x \geqslant a+(d-a)x$ et par décroissance de g, $g(c+(d-c)x) \leqslant g(a+(d-a)x$
en intégrant l'inégalité sur $[0;1]$, on obtient l'inégalité de droite.
l'encadrement $\displaystyle\int_0^1 g(c+(b-c)x)dx \leqslant \int_0^1 g(c+(d-c)x)dx \leqslant \int_0^1 g(a+(d-a)x)dx$
est donc vrai et par équivalence des deux résultats, on a bien :
$\displaystyle\frac{1}{b-c}\int_c^b g(t)dt \leqslant \frac{1}{d-c}\int_c^d g(t)dt \leqslant \frac{1}{d-a}\int_a^d g(t)dt$

Partie I - Partie fractionnaire d'une variable à densité

Pour tout réel x positif ou nul :

- on note $[x]$ la *partie entière* de x. On rappelle qu'il s'agit de l'unique entier naturel n qui vérifie l'encadrement : $n \leqslant x < n+1$
- on note $\{x\} = x - [x]$, que l'on appelle la *partie fractionnaire* de x

Par exemple, si $x = 12,34$, alors $[x] = 12$ et $\{x\} = 0,34$

Dans cette partie, X désigne une variable aléatoire à valeurs réelles admettant une densité f qui vérifie les propriétés :

- f est nulle sur $]-\infty, 0[$
- la restriction de f à $[0, +\infty[$ est continue et décroissante.

On pose $M = f(0)$, c'est le maximum de f sur $\mathbb{R}$

Soit $Y = \big\{X\big\} = X - \big[X\big]$, la variable aléatoire égale à la partie fractionnaire de X
On note F_Y la fonction de répartition de Y

2. Que vaut $F_Y(y)$ lorsque $y<0$? Que vaut $F_Y(y)$ lorsque $y \geqslant 1$? *1,5 points*
On justifiera les réponses.

Pour tout $x \in \mathbb{R}$ on a $\lfloor x \rfloor \leqslant x$ donc $\{x\} = \lfloor x \rfloor - x \geqslant 0$
d'où pour tout $y<0$, $F_Y(y) = \mathrm{P}(Y \leqslant y) = \mathrm{P}(\{X\} \leqslant y) = 0$ car c'est un événement impossible
d'autre part on a $x < [x]+1$ donc $\{x\} = x - [x] < 1$
pour tout $x \in \mathbb{R}$ d'où pour tout $y \geqslant 1$, $F_Y(y) = \mathrm{P}(Y \leqslant y) = 1$ car c'est un événement certain.

3. Justifier l'égalité entre les événements : $(Y=0) = \displaystyle\bigcup_{n\in\mathbb{N}} (X=n)$ *2 points*

En déduire : $F_Y(0) = 0$

On transforme l'événement par équivalence :
$Y=0 \Longleftrightarrow \{X\}=0 \Longleftrightarrow X-[X]=0 \Longleftrightarrow X=[X] \Longleftrightarrow X \in \mathbb{Z}$
de plus on sait que X est à valeurs dans $\mathbb{R}^+$donc ne peut être négatif, ce qui donne : $Y=0 \Longleftrightarrow X \in \mathbb{N}$
On en déduit bien que $(Y=0) = \displaystyle\bigcup_{n\in\mathbb{N}} (X=n)$

la réunion est incompatible, on en déduit (X est à densité) : $\mathrm{P}(Y=0) = \displaystyle\sum_{n=0}^{+\infty} \mathrm{P}(X=n) = \sum_{n=0}^{+\infty} 0 = 0$

d'où le résultat car $F_Y(0) = \mathrm{P}(Y \leqslant 0) = \mathrm{P}(Y<0) + \mathrm{P}(Y=0) = 0+0$

4. Soit y un réel de l'intervalle $]0,1[$

a. Montrer l'égalité : $F_Y(y) = \displaystyle\sum_{n=0}^{+\infty} \int_n^{n+y} f(t)\,\mathrm{d}t$ *3 points*

Soit y un réel de l'intervalle $]0,1[$.
avec le système complet d'événements $(\lfloor X \rfloor = n)_{n\in\mathbb{N}}$, les probabilités totales donnent :

$$F_Y(y) = \mathrm{P}(Y \leqslant y) = \mathrm{P}(X - \lfloor X \rfloor \leqslant y) = \sum_{n=0}^{+\infty} \mathrm{P}[(X - \lfloor X \rfloor \leqslant y) \cap (\lfloor X \rfloor = n)]$$

$$= \sum_{n=0}^{+\infty} \mathrm{P}[(X \leqslant n+y) \cap (\lfloor X \rfloor = n)] = \sum_{n=0}^{+\infty} \mathrm{P}[(X \leqslant n+y) \cap (n \leqslant X < n+1)]$$

or on a supposé que $y<1$, donc $n+y<n+1$ enfin : $(X \leqslant n+y) \cap (X < n+1) = (X \leqslant n+y)$

ce qui donne (X est à densité et sa densité est f) : $F_Y(y) = \displaystyle\sum_{n=0}^{+\infty} \mathrm{P}(n \leqslant X \leqslant n+y) = \sum_{n=0}^{+\infty} \int_n^{n+y} f(t)dt$

b. Montrer, en utilisant la question préliminaire, les inégalités : *2 points*

- Pour tout n entier naturel, $\displaystyle\int_n^{n+y} f(t)\,\mathrm{d}t \geqslant y \int_n^{n+1} f(t)\,\mathrm{d}t$
- Pour tout n entier naturel non nul, $\displaystyle\int_n^{n+y} f(t)\,\mathrm{d}t \leqslant y \int_{n-1+y}^{n+y} f(t)\,\mathrm{d}t$

Pour tout $n \in \mathbb{N}$ on a $n < n+y < n+1$ et f est décroissante et continue sur $\mathbb{R}^+$donc en appliquant le résultat préliminaire avec $c=n, b=n+1$ et $d=n+y$:

$$\frac{1}{n+1-n}\int_n^{n+1} f(t)dt \leqslant \frac{1}{n+y-n}\int_n^{n+y} f(t)dt \text{ et donc } y\int_n^{n+1} f(t)dt \leqslant \int_n^{n+y} f(t)dt$$

puis pour tout $n \geqslant 1$ on a $n-1+y < n < n+y$ car $0<y<1$ et f est décroissante et continue sur $\mathbb{R}^+$donc en appliquant le préliminaire avec $c=n, a=n-1+y$ et $d=n+y$:

$$\frac{1}{n+y-n}\int_n^{n+y} f(t)dt \leqslant \frac{1}{n+y-(n-1+y)}\int_{n-1+y}^{n+y} f(t)dt \text{ et donc } \int_n^{n+y} f(t)dt \leqslant y\int_{n-1+y}^{n+y} f(t)dt$$

c. En déduire : $y\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t \leqslant F_Y(y) \leqslant \int_0^y f(t)\,\mathrm{d}t + y\int_y^{+\infty} f(t)\,\mathrm{d}t$, puis l'encadrement *3,5 points*

$$y \leqslant F_Y(y) \leqslant y + M$$

On somme de 0 à $+\infty$ la première inégalité du **b.** : $y\displaystyle\int_0^{+\infty} f(t)dt \leqslant \sum_{k=0}^{+\infty}\int_n^{n+y} f(t)dt = F_Y(y)$

pour l'inégalité de droite, on somme la deuxième inégalité du **b.** pour k allant de 1 à $+\infty$:

$y\displaystyle\int_y^{+\infty} f(t)dt \geqslant \sum_{k=1}^{+\infty}\int_n^{n+y} f(t)dt$

enfin on ajoute $\displaystyle\int_0^y f(t)dt$ des deux côtés et on obtient : $\displaystyle\int_0^y f(t)dt + y\int_y^{+\infty} f(t)dt \geqslant F_Y(y)$

ce qui donne bien : $y\displaystyle\int_0^{+\infty} f(t)dt \leqslant F_Y(y) \leqslant \int_0^y f(t)dt + y\int_y^{+\infty} f(t)dt$

de plus X est à valeurs positives donc :

$\displaystyle\int_0^{+\infty} f(t)dt = 1 = \mathrm{P}(X \geqslant 0) = 1$ et l'inégalité de gauche donne $y \leqslant F_Y(y)$

de plus on a $\displaystyle\int_y^{+\infty} f(t)dt = \int_0^{+\infty} f(t)dt - \int_0^y f(t)dt = 1 - \int_0^y f(t)dt \leqslant 1$ par positivité de l'intégrale

en effet $f(t) \geqslant 0$ et les bornes sont dans l'ordre croissant donc $\displaystyle\int_0^y f(t)dt \geqslant 0$

on en déduit que $F_Y(y) \leqslant \displaystyle\int_0^y f(t)dt + y$

enfin on pour tout $t \in [0; y]$, par décroissance de f, $f(t) \leqslant f(0) = M$ donc en intégrant avec des bornes dans l'ordre croissant : $\displaystyle\int_0^y f(t)dt \leqslant \int_0^y Mdt = My \leqslant M$ car $y \leqslant 1$

cela donne bien $F_Y(y) \leqslant y + M$, puis : $y \leqslant F_Y(y) \leqslant y + M$

Partie II - Premier chiffre significatif d'une variable de Pareto

Pour tout réel λ strictement positif, on définit la fonction g_λ sur $\mathbb{R}$ par $g_\lambda : x \mapsto \begin{cases} \dfrac{\lambda}{x^{\lambda+1}} & \text{si } x \geqslant 1 \\ 0 & \text{sinon} \end{cases}$

5. Montrer que pour tout réel λ strictement positif, g_λ est une densité de probabilité sur $\mathbb{R}$ (loi dite *de Pareto*). *1,5 points*

g_λ est une fonction à valeurs positives, continue sur $]-\infty; 1[$ et sur $]1; +\infty[$, il reste à vérifier que $\displaystyle\int_{-\infty}^{+\infty} g_\lambda(t)dt$ converge et vaut 1 .

par définition, $\displaystyle\int_{-\infty}^{+\infty} g_\lambda(t)dt = \int_{-\infty}^{1} 0dt + \int_1^{+\infty}\frac{\lambda}{t^{\lambda+1}}dt$ car $\displaystyle\int_{-\infty}^{1} g_\lambda(t)dt$ converge et vaut 0 et $\displaystyle\int_{-\infty}^{1} g_\lambda(t)dt$ converge en tant que combinaison linéaire d'une intégrale de Riemann convergente ($\lambda + 1 > 1$)

et pour $x \geqslant 1$, $\displaystyle\int_1^x g_\lambda(t)dt = \left[\frac{\lambda t^{-\lambda}}{-\lambda}\right]_1^x = -\frac{1}{x^\lambda} + 1 \xrightarrow[x\to+\infty]{} 1$ car $\lambda > 0$, donc $\displaystyle\lim_{x\to+\infty} x^\lambda = +\infty$ et $\displaystyle\lim_{x\to+\infty}\frac{1}{x^\lambda} = 0$

On en déduit que $\displaystyle\int_{-\infty}^{+\infty} g_\lambda(t)dt$ converge et vaut 1, et g_λ est une densité de probabilité.

Dans toute la suite, on note Z_λ une variable aléatoire admettant g_λ pour densité.

6. Déterminer la fonction de répartition G_λ de Z_λ *1,5 points*

f est une densité (et admet des limites finies à droite et à gauche en tout point) donc

pour tout $x \in \mathbb{R}$, $G_\lambda(x) = \displaystyle\int_{-\infty}^{x} g_\lambda(t)dt$

pour tout $x < 1$, $G_\lambda(x) = \mathrm{P}(Z \leqslant x) = \displaystyle\int_{-\infty}^{x} 0dt = 0$

pour tout $x \geqslant 1$, $G_\lambda(x) = \mathrm{P}(Z \leqslant x) = \displaystyle\int_{-\infty}^{1} 0dt + \int_1^x \frac{\lambda}{t^{\lambda+1}}dt = 0 + \left[\frac{\lambda t^{-\lambda}}{-\lambda}\right]_1^x = -\frac{1}{x^\lambda} + 1$

on rassemble : $G_\lambda(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 - \dfrac{1}{x^\lambda} & \text{si } x \geqslant 1 \end{cases}$

7. On note ln la fonction *logarithme népérien*, et log la fonction *logarithme décimal*.

Cette fonction est définie sur $]0, +\infty[$ par : $\log(x) = \dfrac{\ln(x)}{\ln(10)}$ pour tout réel x strictement positif.

On pose $X_\lambda = \log\big(Z_\lambda\big)$, et on note F_λ la fonction de répartition de X_λ

a. Etablir, pour tout réel x, l'égalité : $F_\lambda(x) = G_\lambda\big(10^x\big)$ *1 point*

Pour tout $x \in \mathbb{R}, F_\lambda(x) = \mathrm{P}\left(X_\lambda \leqslant x\right) = \mathrm{P}\left(\log\left(Z_\lambda\right) \leqslant x\right) = \mathrm{P}\left(\dfrac{\ln\left(Z_\lambda\right)}{\ln 10} \leqslant x\right)$

$$= \mathrm{P}\left(\ln\left(Z_\lambda\right) \leqslant x\ln 10\right) = \mathrm{P}\left(Z_\lambda \leqslant e^{x\ln 10}\right) = \mathrm{P}\left(Z_\lambda \leqslant 10^x\right)$$
$$= G_\lambda\left(10^x\right)$$

b. En déduire que X_λ suit une loi exponentielle dont on précisera le paramètre en fonction de λ *2 points*

On en déduit que : $F_\lambda(x) = \begin{cases} 0 & \text{si } 10^x < 1 \\ 1 - \dfrac{1}{(10^x)^\lambda} & \text{si } 10^x \geqslant 1 \end{cases}$

On résout les inéquations (avec $\ln(10) > 0$ car $10 > 1$) :
$10^x < 1 \iff x\ln(10) < \ln(1) \iff x < 0$ donc $10^x \geqslant 1 \iff x \geqslant 0$

enfin on calcule $(10^x)^\lambda = 10^{\lambda x} = e^{\lambda x\ln(10)}$ et on obtient : $F_\lambda(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-[\lambda\ln(10)]x} & \text{si } x \geqslant 0 \end{cases}$

et X_λ suit donc une loi exponentielle de paramètre $\lambda\ln(10)$

8. On pose $Y_\lambda = \big\{X_\lambda\big\}$, la partie fractionnaire de X_λ *3 points*
Montrer, en utilisant les résultats de la **partie I**, que pour tout réel y de l'intervalle $]0,1[$:
$$\lim_{\lambda\to 0^+} P\big(Y_\lambda \leqslant y\big) = y$$
En déduire que, pour tout y, lorsque λ tend vers 0^+, $P\big(Y_\lambda \leqslant y\big)$ tend vers $F_T(y)$ où F_T est la fonction de répartition d'une variable aléatoire T suivant une loi usuelle, loi que l'on précisera.
On dit alors que Y_λ converge en loi vers la variable aléatoire T

D'après le cours, la fonction, $f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \lambda\ln(10)e^{-\lambda\ln(10)x} & \text{si } x \geqslant 0 \end{cases}$ est une densité de X_λ, qui de plus est nulle sur $\mathbb{R}^{-*}$, continue, décroissante et strictement positive sur $\mathbb{R}^+$donc les résultats de la **partie I** s'appliquent.
On en déduit que pour tout $y \in]0;1[$, $y \leqslant F_{Y_\lambda}(y) \leqslant y + f(0) = y + \lambda\ln(10)$
or $\lim\limits_{\lambda\to 0^+} \lambda\ln(10) = 0$ donc par encadrement, $\lim\limits_{\lambda\to 0^+} F_{Y_\lambda}(y) = \lim\limits_{\lambda\to 0^+} \mathrm{P}\left(Y_\lambda \leqslant y\right) = y$

Il reste à prouver que : $\forall y \in]-\infty;0]$, $\lim\limits_{\lambda\to 0^+} \mathrm{P}\left(Y_\lambda \leqslant y\right) = 0$ et $\forall y \in \Big[1;+\infty\Big[$, $\lim\limits_{\lambda\to 0^+} \mathrm{P}\left(Y_\lambda \leqslant y\right) = 1$
or la **partie I** montre que la fonction de répartition de Y_λ vérifie exactement ces deux propriétés (nulle sur $\mathbb{R}^-$et égale à 1 sur $[1;+\infty[$)
En passant à la limite pour $\lambda \to 0$ on obtient le résultat.
D'où en posant une variable aléatoire Y suivant la loi uniforme sur $[0;1]$ on a $\forall y \in \mathbb{R}$, $\lim\limits_{\lambda\to 0^+} F_{Y_\lambda}(y) = F_Y(y)$
et Y_λ converge en loi vers Y quand λ tend vers 0 par valeurs positives.

9. Pour tout réel x supérieur ou égal à 1, on note $\alpha(x)$ le premier chiffre dans l'écriture décimale de x. C'est un entier de l'intervalle $[\![1,9]\!]$

Par exemple, $\alpha(50) = 5$ et $\alpha(213,43) = 2$

a. Pour tout $k \in [\![1,9]\!]$, montrer l'équivalence : $\alpha(x) = k \quad \Leftrightarrow \quad \big\{\log(x)\big\} \in \big[\log k, \log(k+1)\big[$ *3 points*

Question difficile.
$\alpha(x) = k$ signifie qu'il existe $n \in \mathbb{N}$ tel que
$k \times 10^n \leqslant x < (k+1) \times 10^n$
(où $n = 0$ si $1 \leqslant x < 10, n = 1$ si $10 \leqslant x < 100, n = 2$ si $100 \leqslant x < 1000$, etc...

$$\alpha(x) = k \iff \exists n \in \mathbb{N},\quad \ln k + n\ln 10 \leqslant \ln x < \ln(k+1) + n\ln 10$$

d'où
$$\iff \exists n \in \mathbb{N},\quad \frac{\ln k}{\ln 10} + n \leqslant \frac{\ln x}{\ln 10} < \frac{\ln(k+1)}{\ln 10} + n$$
$$\iff \exists n \in \mathbb{N},\quad \log(k) + n \leqslant \log x < \log(k+1) + n$$

montrons alors que n est la partie entière de $\log(x)$:

Pour tout $k \in [1;9], \log 1 = 0 \leqslant \log k \leqslant \log 9 = \dfrac{\ln 9}{\ln 10} < 1$ donc on a $\log x \geqslant n$ car $\ln k \geqslant 0$

d'autre part pour tout $k \in [1;9], \log(k+1) \leqslant \log 10 = 1$ donc $\log x < n+1$

on obtient que $n \leqslant \log x < n+1$ et $n \in \mathbb{N}$ donc $n = \lfloor \log x \rfloor$

en remplaçant on obtient :

$$\begin{aligned}\alpha(x) = k &\iff \log k + \lfloor \log x \rfloor \leqslant \log x < \log(k+1) + \lfloor \log x \rfloor \iff \log k \leqslant \log x - \lfloor \log x \rfloor < \log(k+1)\\ &\iff \log k \leqslant \{\log x\} < \log(k+1) \iff \{\log(x)\} \in [\log k, \log(k+1)[.\end{aligned}$$

b. On note $C_\lambda = \alpha\big(Z_\lambda\big)$ la variable aléatoire prenant comme valeur le premier chiffre de Z_λ *3 points*

Montrer, pour tout $k \in [\![1,9]\!]$: $\lim\limits_{\lambda \to 0^+} P(C_\lambda = k) = \log\left(1 + \dfrac{1}{k}\right)$

Cette loi limite obtenue pour le premier chiffre de Z_λ est appelée *loi de Benford.*

Pour tout $\lambda > 0$ et tout $k \in [\![1;9]\!]$ on a :

$$\mathrm{P}\left(C_\lambda = k\right) = \mathrm{P}\left(\alpha\left(Z_\lambda\right) = k\right) = \mathrm{P}\left(\log k \leqslant \{\log Z_\lambda\} < \log(k+1)\right) = \mathrm{P}\left(\log k \leqslant Y_\lambda < \log(k+1)\right)$$

Or Y_λ converge en loi vers une variable aléatoire Y suivant la loi uniforme sur $[0;1]$ lorsque λ tend vers 0^+donc

$$\lim_{\lambda \to 0^+} \mathrm{P}\left(\log k \leqslant Y_\lambda < \log(k+1)\right) = \mathrm{P}(\log k \leqslant Y < \log(k+1))$$

comme $k \in [\![1;9]\!]$, on a vu que $0 \leqslant \log(k) \leqslant 1$

donc $F_Y(\log(k)) = \log(k)$ et $F_Y(\log(k+1)) = \log(k+1)$

et enfin :

$$\mathrm{P}\left(C_\lambda = k\right) = \log(k+1) - \log(k) = \frac{\ln(k+1) - \ln k}{\ln(10)} = \frac{\ln\left(\frac{k+1}{k}\right)}{\ln(10)} = \log\left(\frac{k+1}{k}\right) = \log\left(1 + \frac{1}{k}\right).$$