MATHEMATIQUES
Algèbre linéaire et Variables aléatoires à densité

Le sujet se compose de ***deux problèmes****. On prendra soin de lire l'ensemble des énoncés avant de commencer à composer.*
Il sera tenu compte de la présentation et en particulier de $\boxed{\textit{l'encadrement}}$ *des résultats.*
*L'usage de la calculatrice n'****est pas*** *autorisé au cours de l'épreuve.*

Problème 1 :

I. Préliminaires

Dans cette partie **I.**, λ désigne un réel strictement positif.

① Soit X une variable aléatoire suivant la loi exponentielle de paramètre λ.

a) Déterminer la fonction : $x \mapsto \mathrm{P}(X > x)$ (appelée *fonction de survie de* X).

b) Pour tous nombres réels strictement positifs x et y, calculer la probabilité conditionnelle $\mathrm{P}_{(X>x)}(X > x+y)$; justifier alors que, si X modélise la durée de vie d'un phénomène, on dise de ce dernier qu'il est « sans vieillissement ».

② Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, de même loi exponentielle de paramètre λ. Pour tout entier naturel n non nul, on pose :

$$S_n = \sum_{k=1}^{n} X_k.$$

a) Déterminer l'espérance et la variance de S_n.

b) Démontrer par récurrence que, pour tout n de $\mathbb{N}^*$, la variable aléatoire S_n admet pour densité la fonction f_n :

$$t \longmapsto \begin{cases} 0 & \text{si } t < 0\,; \\ \dfrac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} & \text{si } t \geq 0. \end{cases}$$

Pour cela, on admettra que, si U et V sont des variables aléatoires indépendantes admettant respectivement pour densité les fonctions f_U et f_V, alors la variable aléatoire $U+V$ admet pour densité la fonction f_{U+V} définie sur $\mathbb{R}$ par :

$$f_{U+V}(t) = \int_{-\infty}^{+\infty} f_U(x) f_V(t-x)\, dx.$$

II. Loi de Pareto (Vilfredo Pareto (1848-1923), sociologue et économiste italien)

Soient a et b des réels strictement positifs. Par définition, on dit d'une variable aléatoire qu'elle suit la loi de Pareto de paramètres a et b si elle admet pour densité la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} 0 & \text{si } t < b\,; \\ a\dfrac{b^a}{t^{a+1}} & \text{si } t \geq b. \end{cases}$$

Soit alors X une variable aléatoire de loi de Pareto de paramètres a et b.

① Vérifier que f est bien une densité de probabilité ; calculer l'espérance et la variance de X, en précisant à quelles conditions chacune de ces quantités existe.

② Déterminer la fonction de répartition F_X de X. Préciser la fonction de survie : $x \mapsto \mathrm{P}(X > x)$.

③ Montrer que F_X est une bijection de $]b, +\infty[$ sur $]0, 1[$.
En considérant pour cette question que $X(\Omega) =]b, +\infty[$ plutôt que $[b, +\infty[$ et en supposant que U est une variable aléatoire qui suit une loi uniforme sur $]0, 1[$, montrer que $F_X^{-1}(U)$ suit une loi de Pareto de paramètres a et b.
En déduire une fonction Python `estimEspPareto(a, b, m = 1000)` permettant d'estimer l'espérance d'une telle loi.

④ Démontrer que, pour tout réel y positif ou nul, la probabilité conditionnelle $\mathrm{P}_{(X>x)}(X > x + y)$ tend vers 1 quand x tend vers $+\infty$. De façon analogue à la question **I.1.b)**, que peut-on dire d'un phénomène dont la durée de vie est modélisée par X ?

⑤ On pose dans cette question : $Y = \ln \dfrac{X}{b}$.
Démontrer que Y suit une loi exponentielle dont on précisera le paramètre.

III. Estimation du paramètre a d'une loi de Pareto

Les instants aléatoires des arrivées de paquets (symboles binaires représentant de l'information de type audio, vidéo, données, ...) dans un canal de communication sont modélisés par une variable aléatoire X suivant une loi de Pareto de paramètres α et β ($\alpha > 0$, $\beta > 0$).
On suppose ici que le paramètre β fait partie des caractéristiques connues du canal de communication ;
On se propose d'estimer α par une méthode dite « du maximum de vraisemblance ».
Précisons son principe : Soit n un entier naturel non nul et x_1, ..., x_n un échantillon de n réalisations de la loi de Pareto dont les paramètres sont β connu et α, qu'on cherche à déterminer. ✐ Notons que ce sont des réels supérieurs ou égaux à β qu'on supposera non tous égaux à β.
Connaissant ces nombres, on cherche à estimer la valeur de α la plus vraisemblable pour que chacun de ces x_i puisse être considéré comme la réalisation de n variables indépendantes $X_1, \cdots, X_n$ qui suivent une même loi de Pareto de paramètres α et β.
On introduit pour ça la fonction $\mathcal{L}$, à valeurs dans $\mathbb{R}$ et définie sur $]0, +\infty[$ par :

$$\mathcal{L}(a) = f_a(x_1) \times \cdots \times f_a(x_n) = \prod_{k=1}^{n} f_a\left(x_k\right),$$

où f_a est la fonction définie sur $\mathbb{R}$ par :

$$f_a(t) = \begin{cases} 0 & \text{si } t < \beta\,; \\ a\dfrac{\beta^a}{t^{a+1}} & \text{si } t \geq \beta. \end{cases}$$

On admettra que la valeur de a qui maximise $\mathcal{L}$ est la valeur la plus vraisemblable pour α.

① Exprimer $\mathcal{L}(a)$, puis $\ln\left(\mathcal{L}(a)\right)$ en fonction des x_k, de n et des paramètres a et β.

② On considère la fonction φ de $]0, +\infty[$ dans $\mathbb{R}$:

$$a \mapsto n\ln(a) + na\ln(\beta) - (a+1)\sum_{k=1}^{n} \ln\left(x_k\right).$$

a) Démontrer que la fonction φ admet un maximum, atteint en un seul réel que l'on notera w.

b) Exprimer w en fonction de x_1, ..., x_n, n et β.

c) Que peut-on dire de w pour la fonction $\mathcal{L}$? Conclure sur la valeur de α la plus vraisemblable au regard de notre échantillon $(x_1, \cdots, x_n)$.

③ On pose dorénavant, pour tout n de $\mathbb{N}^*$:

$$W_n = \frac{n}{\displaystyle\sum_{k=1}^{n} \ln \frac{X_k}{\beta}}.$$

(La suite $(W_n)_{n\geq 1}$ est appelée *estimateur du maximum de vraisemblance.*)

a) Justifier que la variable aléatoire $\displaystyle\sum_{k=1}^{n} \ln \frac{X_k}{\beta}$ admet pour densité la fonction f_n définie dans **I.2.b)** en prenant $\lambda = \alpha$.

b) À l'aide du théorème de transfert, en déduire que W_n admet pour espérance $\frac{n\alpha}{n-1}$ lorsque $n \geq 2$, puis proposer une variable aléatoire W_n' telle que $\mathbb{E}(W_n') = \alpha$.

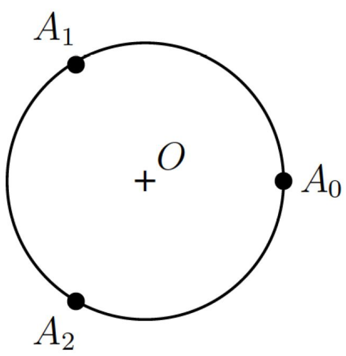
④ On propose les fonctions Python suivantes :

```
def mystere(a, b):
    return b/rdm.random()**(1/a)

def calcul(Lx, b):
    ''' Lx est une liste de n réels supérieurs ou égaux à b '''
    S = sum([np.log(x/b) for x in Lx])
    return (len(Lx)-1)/S

def mystere2(b, n = 30, m = 10000):
    LW = []
    a = rdm.randint(1, 6)
    for k in range(m):
        Lx = [simulPareto(a, b) for k in range(n)]
        LW.append(calculW(Lx))
    return a, np.mean(LW)
```

Commentez une à une chacune de ces fonctions en expliquant leur rôle en insistant sur le sens des variables `a` et `np.mean(LW)` retournées par `mystere2`.

Problème 2 :

Dans la suite p est un entier naturel supérieur ou égal à 3.
On considère un repère orthonormal du plan $\left(O, \vec{i}, \vec{j}\right)$ et le cercle unité $\mathcal{C}$ sur lequel on place dans le sens trigonométrique p points équidistants $A_0, \dots, A_{p-1}$ tels que A_0 soit d'affixe 1.
Ainsi pour tout $k \in [\![0; p-1]\!]$, A_k est le point d'affixe $z_k = e^{\frac{2ik\pi}{p}}$.
Pour $p = 3$, on a la représentation suivante :



1 Résultats préliminaires

① On considère la matrice $D_p = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \in \mathcal{M}_p(\mathbb{C})$.

a) Inverser les matrices D_3 et D_4.

b) Prouver que la matrice D_p est inversible et donner son inverse.
On pourra utiliser les résultats de la question précédente pour conjecturer l'inverse de D_p.

② Soit $f : \mathcal{M}_{p,1}(\mathbb{C}) \to \mathbb{C}$, $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \mapsto \sum\limits_{k=1}^{p} x_k$.

a) Prouver que f est une application $\mathbb{C}$-linéaire.

b) La fonction f est-elle injective ?

c) La fonction f est-elle surjective ?

d) Déterminer la dimension du noyau de f.

③ Calculez $\frac{1}{p} \sum\limits_{k=0}^{p-1} z_k$. Interprétez géométriquement le résultat obtenu.

④ Soit z un complexe non nul. Prouver que : $z^p = 1 \Leftrightarrow \exists k \in \mathbb{Z} \ : \ z = e^{\frac{2ik\pi}{p}}$.
On pourra mettre z sous forme exponentielle ou trigonométrique.

On admettra que l'équation $z^p = 1$ possède p solutions distinctes : $z_0, z_1, \dots, z_{p-1}$.

2 Étude d'un modèle de diffusion sur le cercle

On considère l'expérience suivante : une particule est libre de se déplacer parmi les p points $A_0, A_1, \ldots, A_{p-1}$. Initialement la particule se situe sur le point A_0 et, à chaque étape, on choisit de façon équiprobable de la déplacer vers l'un de ses deux plus proches voisins.
Pour tout $n \in \mathbb{N}$, on note U_n la variable aléatoire réelle à valeurs dans $[\![0; p-1]\!]$ et telle que l'emplacement occupé à l'étape n soit A_{U_n}. La variable aléatoire U_0 est donc constante égale à 0 et la variable aléatoire U_1 est égale à 1 avec une probabilité $\frac{1}{2}$ et à $p-1$ avec une probabilité $\frac{1}{2}$. La variable U_2 est à valeurs dans $\{2; 0; p-2\}$ et $\mathbb{P}(U_2 = 2) = \mathbb{P}(U_2 = p-2) = \frac{1}{4}$.

Pour tout $n \in \mathbb{N}$, on note $X_n = \begin{pmatrix} \mathbb{P}(U_n = 0) \\ \mathbb{P}(U_n = 1) \\ \vdots \\ \mathbb{P}(U_n = p-1) \end{pmatrix}$.

5. Déterminer une matrice $M_p \in \mathcal{M}_p(\mathbb{R})$ telle que, pour tout entier n, on ait
$$X_{n+1} = M_p X_n.$$
Indication : une récurrence n'est pas nécessaire.
6. Soit n un entier. Donner sans justifications l'expression de X_n en fonction de la matrice M_p et de n.
7. Vérifier que $M_p = \frac{1}{2}\left(D_p + D_p^{-1}\right)$.
8. On suppose ici que $p = 3$ et on admet que $M_3 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.
   a) Montrer qu'il existe deux valeurs de λ pour lesquelles $M_3 - \lambda I_3$ n'est pas inversible.
   b) Déterminer pour chacune de ces valeurs de λ une base de l'espace vectoriel $E_\lambda(M_3)$ défini par $E_\lambda(M_3) = \ker(M_3 - \lambda I_3)$. Vérifier que pour tout $X \in E_\lambda(M_3)$, on a : $M_3 X = \lambda X$.
   c) Si on suppose que M_3 est la matrice d'un endomorphisme f_3 de $\mathbb{R}^3$ dans la base canonique $\mathcal{B}$ de $\mathbb{R}^3$ et $X = \mathcal{M}_{\mathcal{B}}(u)$ où u est un vecteur de $\mathbb{R}^3$, alors $M_3 X = \lambda X \Leftrightarrow f_3(u) = \lambda u$.
   En déduire que M_3 et $D = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ sont semblables et donner une matrice P, élément de $\mathcal{M}_3(\mathbb{R})$ telle que $M_3 = PDP^{-1}$.
   d) Déterminer P^{-1}.
   e) Déterminer $\underset{n\to\infty}{lim} D^n$ et en déduire, pour tout $k \in [\![0; 2]\!]$, la limite de $\mathbb{P}(U_n = k)$ lorsque n tend vers $+\infty$. Interprétez le résultat obtenu.
9. Soit $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{C})$ et soit $\lambda \in \mathbb{C}$. Montrer que si $\exists X \neq 0 / D_p X = \lambda X$ alors $\lambda^p = 1$.

   En déduire p valeurs distinctes $\lambda_0, \lambda_1, \cdots, \lambda_{p-1}$ pour lesquelles $D_p X_k = \lambda_k X_k$, avec $X_k = \begin{pmatrix} 1 \\ z_k \\ z_k^2 \\ \vdots \\ z_k^{p-1} \end{pmatrix}$ pour tout $k \in [\![0, p-1]\!]$.

10. Soit $Q = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & z_1 & z_2 & \cdots & z_{p-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & z_1^{p-1} & z_2^{p-1} & \cdots & z_{p-1}^{p-1} \end{pmatrix}$. On cherche à montrer que Q est inversible :

    Pour ça on pose $R = {}^tQ$ et on considère le système homogène (S) : $RX = 0$ où $X = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{pmatrix}$.

    — Montrer que 1, $z_1, \cdots z_{p-1}$ sont racines du polynôme $P(X) = a_0 + a_1X + \cdots a_{p-1}X^{p-1}$.
    — Conclure que la seule solution de (S) est la solution nulle.
    — En déduire l'inversibilité de Q.

11. Posons $X_k = \mathcal{M}_{\mathcal{B}_2}(v_k)$ ou $\mathcal{B}_2$ désigne la base canonique de $\mathbb{C}^p$. Montrer que la famille $(v_0, v_1, \cdots, v_{p-1})$ est une base de $\mathbb{C}^p$.

12. Soit $D_p = \mathcal{M}_{\mathcal{B}_2}(g_p)$ où g_p est un endomorphisme de $\mathbb{C}^p$ et $\mathcal{B}_2$ est sa base canonique. Montrer que :

$$D_p = Q\Delta_pQ^{-1} \text{ où } \Delta_p = \begin{pmatrix} z_0 & 0 & \cdots & \cdots & 0 \\ 0 & z_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & z_{p-1} \end{pmatrix}$$

13. Justifier l'inversibilité de D_p et exprimer D_p^{-1} en fonction de Q et Δ_p.
    En déduire, en utilisant la question 7, qu'il existe une base de $\mathbb{C}^p$ dans laquelle la matrice de l'endomorphisme f_p dont M_p est la matrice dans la base canonique, est diagonale. Donner cette matrice diagonale.

14. Montrer que M_p est semblable à $T_p = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \cos(2\pi/p) & 0 & \cdots & 0 \\ 0 & 0 & \cos(4\pi/p) & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \cos\left(\frac{2(p-1)\pi}{p}\right) \end{pmatrix}$

15. On suppose dans cette question que p est impair.
    a) Soit $k \in [\![1; p-1]\!]$. Déterminer $\lim\limits_{n \to +\infty} \cos\left(\frac{2k\pi}{p}\right)^n$.
    b) En déduire $\underset{n\to\infty}{lim} T_p^n$
    c) En déduire enfin, pour tout $l \in [\![0; p-1]\!]$, la limite de $\mathbb{P}(U_n = l)$ lorsque n tend vers $+\infty$.
    d) Interprétez le résultat obtenu.

- FIN -