CB2 - sujet 1 type Edhec/Ecricome - Lundi 10 mars 2025

Ce devoir est composé de deux exercices et d'un problème.
Vous soignerez votre rédaction et vous encadrerez vos résultats.

Exercice 1 : suites, fonctions, intégrales

On considère les fonctions f et g définies sur $]0,+\infty[$ par :

$$\forall x > 0, \quad f(x) = 2x + 2\sqrt{x(x+1)} \text{ et } g(x) = \ln\left(\frac{\sqrt{x+1}-1}{\sqrt{x+1}+1}\right).$$

On note $(u_n)_{n\in\mathbb{N}}$ et $(I_n)_{n\in\mathbb{N}}$ les suites définies par :

$$\begin{cases} u_0 = 1 \\ u_{n+1} = 2u_n + 2\sqrt{u_n\,(u_n+1)} \quad \text{si } n \in \mathbb{N} \end{cases} \text{ et } I_n = \int_{u_n}^{+\infty} \frac{1}{t\sqrt{t+1}}dt$$

1. (a) Montrer que, pour tout entier naturel n, u_n existe et $u_n > 0$.
   (b) Etudier la monotonie de $(u_n)_{n\in\mathbb{N}}$ et montrer que : $\lim_{n\to+\infty} u_n = +\infty$.
2. Python : écrire un programme affichant le plus petit entier n_0 tel que $u_{n_0} \geq 1000$.
3. Montrer que pour tout $n \in \mathbb{N}$, I_n est une intégrale convergente.
4. (a) Prouver la dérivabilité de g sur $]0,+\infty[$.
   (b) Calculer $g'(x)$ pour tout $x > 0$.
   (c) Montrer que $I_0 = 2\ln(\sqrt{2}+1)$.
5. (a) Montrer que: $\forall x > 0, f(x) + 1 = (\sqrt{x} + \sqrt{x+1})^2$
   (b) En déduire que
$$\forall x > 0, \frac{f'(x)}{f(x)\sqrt{f(x)+1}} = \frac{1}{2x\sqrt{x+1}}$$
6. (a) Montrer, à l'aide du changement de variable $t = f(x)$ que $(I_n)_{n\in\mathbb{N}}$ est géométrique. (vous justifierez bien sûr que ce changement de variable est autorisé)
   (b) Exprimer I_n en fonction de n pour tout $n \in \mathbb{N}$.
7. (a) Montrer, à l'aide de g, que :
$$\forall x > 0, \int_x^{+\infty} \frac{1}{t\sqrt{t+1}}dt = \ln\left(1 + \frac{1}{\sqrt{x+1}}\right) - \ln\left(1 - \frac{1}{\sqrt{x+1}}\right)$$
   (b) En déduire que : $\displaystyle\int_x^{+\infty} \frac{1}{t\sqrt{t+1}}dt \underset{x\to+\infty}{\sim} \frac{2}{\sqrt{x}}$.
   (c) Déterminer enfin un équivalent de u_n lorsque n tend vers $+\infty$.

Exercice 2 : les matrices de Gram

On note n un entier supérieur ou égal à 2 . On considère l'espace vectoriel $\mathbb{R}^n$ muni de sa base canonique $\mathscr{B} = (e_1, e_2, \dots, e_n)$ et de son produit scalaire canonique noté $\langle ., . \rangle$. On note $\|.\|$ la norme associée.
Soit p un entier vérifiant $1 \leq p \leq n$. À toute famille $(u_1, u_2, \dots, u_p)$ de vecteurs de $\mathbb{R}^n$, on associe la matrice $G = (g_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq p}}$ de $\mathscr{M}_p(\mathbb{R})$, appelée matrice de Gram de $(u_1, u_2, \dots, u_p)$, telle que, pour tout $(i, j) \in [\![1, p]\!]^2$, on a $g_{i,j} = \langle u_i \mid u_j \rangle$. Ainsi on a :

$$G = \begin{pmatrix} \langle u_1 \mid u_1 \rangle & \langle u_1 \mid u_2 \rangle & \dots & \langle u_1 \mid u_p \rangle \\ \langle u_2 \mid u_1 \rangle & \langle u_2 \mid u_2 \rangle & \dots & \langle u_2 \mid u_p \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle u_p \mid u_1 \rangle & \langle u_p \mid u_2 \rangle & \dots & \langle u_p \mid u_p \rangle \end{pmatrix}$$

1. Dans cette question, on suppose que $n = p = 3$ et on considère les trois vecteurs:

$$u_1 = (1, -1, 0), u_2 = (1, 0, -1), u_3 = (1, 1, 1)$$

   (a) Déterminer la matrice de Gram G de (u_1, u_2, u_3). Justifier que G est diagonalisable.

   (b) Exprimer G^2 en fonction de G et de I_3, où I_3 est la matrice identité de $\mathscr{M}_3(\mathbb{R})$.

   (c) Déterminer le spectre de G et préciser la dimension des sous-espaces propres associés.

   Dans toute la suite de l'exercice, on revient au cas général où n est un entier supérieur ou égal à 2.

2. Soit p un entier appartenant à $[\![1, n]\!]$ et $(u_1, u_2, \dots, u_p)$ une famille de vecteurs de $\mathbb{R}^n$. On note G la matrice de Gram de $(u_1, u_2, \dots, u_p)$.

   (a) Justifier que la matrice G est diagonalisable.

   (b) On considère la forme quadratique q_G associée à G.
   Montrer que pour tout $(x_1, ..., x_p) \in \mathbb{R}^p$, $q_G(x_1, ..., x_p) \geq 0$.
   Que peut-on en déduire quand au spectre de G ?

3. Soit encore p un entier appartenant à $[\![1, n]\!]$ et $(u_1, u_2, \dots, u_p)$ une famille de vecteurs de $\mathbb{R}^n$. On pose F le sous-espace vectoriel de $\mathbb{R}^n$ engendré par $(u_1, u_2, \dots, u_p)$ et on note G la matrice de Gram de $(u_1, u_2, \dots, u_p)$.

   On considère des réels $\alpha_1, \alpha_2, \dots, \alpha_p$ et on pose $X = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix} \in \mathscr{M}_{p,1}(\mathbb{R})$.

   (a) Montrer que si $\sum_{j=1}^{p} \alpha_j u_j = 0$, alors $GX = 0$.

   (b) On suppose réciproquement que $GX = 0$. On note $F^\perp$ l'orthogonal de F.
   Montrer que $\sum_{j=1}^{p} \alpha_j u_j$ appartient à $F^\perp$ et en déduire que $\sum_{j=1}^{p} \alpha_j u_j = 0$.

   (c) Déduire de ce qui précède que la famille $(u_1, u_2, \dots, u_p)$ est libre si et seulement si la matrice G est inversible.

4. On considère dans cette question une famille $(v_1, v_2, \dots, v_n)$ de vecteurs de $\mathbb{R}^n$ telle que :

$$\forall i \in [\![1, n]\!], \quad \|v_i\| = 1 \text{ et } \quad \forall (i, j) \in [\![1, n]\!]^2, \quad i \neq j, \quad \|v_i - v_j\| = 1$$

   (a) Pour tous a et b dans $\mathbb{R}^n$, exprimer $\|a - b\|^2$ à l'aide de $\|a\|^2$, $\|b\|^2$ et $\langle a, b \rangle$.

(b) En déduire la matrice de Gram G de $(v_1, v_2, \dots, v_n)$.

(c) On pose $A = 2G$ et on note J la matrice de $\mathscr{M}_n(\mathbb{R})$ dont tous les coefficients valent 1 et enfin I_n la matrice identité de $\mathscr{M}_n(\mathbb{R})$.

i. Exprimer A à l'aide de I_n et J, puis A^2 en fonction de n, I_n et J et enfin A^2 en fonction de n, A et I_n.

ii. Prouver que A est inversible.

iii. En déduire que $(v_1, v_2, \dots, v_n)$ est une base de $\mathbb{R}^n$.

5. Soit $\mathscr{B}_1 = (w_1, w_2, \dots, w_n)$ une base quelconque de $\mathbb{R}^n$ et soit G la matrice de Gram de $(w_1, w_2, \dots, w_n)$.
Soit $x = \sum_{j=1}^{n} \alpha_j w_j$ un vecteur de $\mathbb{R}^n$.
On pose $X = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ et $Z = \begin{pmatrix} \langle x \mid w_1 \rangle \\ \langle x \mid w_2 \rangle \\ \vdots \\ \langle x \mid w_n \rangle \end{pmatrix}$ deux éléments de $\mathscr{M}_{n,1}(\mathbb{R})$.

(a) Justifier que G est inversible et montrer que $X = G^{-1}Z$.

(b) Montrer qu'il existe une unique famille $\mathscr{B}_1^* = (w_1^*, w_2^*, \dots, w_n^*)$ de vecteurs de $\mathbb{R}^n$ telle que :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad \langle w_i^* \mid w_j \rangle = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

(c) Montrer que $\mathscr{B}_1^*$ est une base de $\mathbb{R}^n$.

(d) Donner, à l'aide des w_i^*, les coordonnées d'un vecteur x de $\mathbb{R}^n$ dans la base $\mathscr{B}_1$.

Problème : autour de la loi de Pareto *(Vilfredo Pareto, économiste italien 1848-1923)*

Partie I

Soit a un réel fixé, avec $a > 0$. On considère la fonction f_a définie sur $\mathbb{R}$ par :

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

1. (a) Justifier que f_a est une densité de probabilité d'une variable aléatoire X.
On dit que X suit la loi de Pareto de paramètre a et on note $X \hookrightarrow \mathcal{VP}(a)$.

(b) Déterminer la fonction de répartition F_a de la variable aléatoire X.

(c) Déterminer pour tout $x > 1$, $P(X > x)$.
Justifier que pour tout $y > 0$, $\lim_{x \to +\infty} P_{[X>x]}(X > x + y) = 1$.

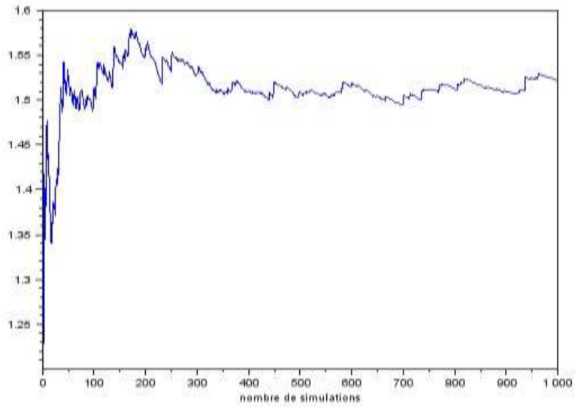
2. (a) Préciser les valeurs de a telles que $E(X)$ existe. Calculer $E(X)$ lorsque cette espérance existe.

(b) Préciser les valeurs de a telles que $V(X)$ existe. Calculer $V(X)$ lorsque cette variance existe.

3. Soit U une variable aléatoire de loi uniforme sur $]0,1]$ et $R = U^{-\frac{1}{a}}$.
Justifier que R a la même loi que X.

4. On considère le script suivant et une de ses réalisations pour $a = 3$ et $n = 1000$

```
n=1000
a=3
U=rd.random(n)
X=U**(-1/a)
C=np.cumsum(X)
for k in range(n):
   C[k]=C[k]/k
plt.plot(np.arange(0,n),C)
plt.show()
```



Expliquer ce que fait ce programme pas à pas. Si le vecteur X a pour valeur $(x_1, \cdots, x_n)$, quel est le contenu du vecteur C à l'issue du programme ?
La figure illustre un résultat de convergence pour une suite de variables aléatoires.
Lequel ? Justifiez-le en introduisant des variables adaptées.

5. Soit Y et X deux variables aléatoires où $X = \exp(Y)$. Justifier que : Y suit une loi exponentielle de paramètre a si et seulement si X suit une loi de Pareto de paramètre a.

   Autrement dit, si X suit une loi de Pareto de paramètre a, la variable $Y = \ln(X)$ suit une loi exponentielle de paramètre a.

6. Soient X et X' deux variables aléatoires indépendantes où $X \hookrightarrow \mathcal{VP}(a)$ et $X' \hookrightarrow \mathcal{VP}(a')$, de densité f_a et $f_{a'}$.
On note $Y = ln(X)$, $Y' = ln(X')$, et on a donc $Y \hookrightarrow \mathcal{E}(a)$ et $Y' \hookrightarrow \mathcal{E}(a')$.

   (a) Préciser une densité de $-Y$.

   (b) Justifier que la VAR $D = Y' - Y$ est à densité et qu'une densité de D est la fonction f_D telle que
$$\forall x \in [0; +\infty[, \quad f_D(x) = \frac{a.a'}{a + a'}.e^{-a'x}$$
Déterminer f_D sur $]-\infty, 0[$.

   (c) En déduire que $P(X \leq X') = \dfrac{a}{a + a'}$. Que vaut $P(X = X')$?

Partie II : autour du minimum

On considère une suite de variables aléatoires indépendantes $(X_k)_{k\in\mathbb{N}^*}$, où pour tout $k \in \mathbb{N}^*$, X_k suit la loi de Pareto de paramètre a_k strictement positif. On note pour tout $n \in \mathbb{N}^*$, $A_n = min(X_1, \cdots, X_n)$.

1. Déterminer la fonction de répartition de A_n. Vérifier que A_n suit une loi de Pareto et préciser son paramètre.

2. Soit $n \geq 2$.
A l'aide du I.6.(c) appliqué à A_{n-1}, justifier que $P(X_n = min(X_1, \cdots, X_n)) = \dfrac{a_n}{a_1 + \cdots + a_n}$.

3. On considère le programme suivant :

```
n=int(input("n="))
C=0
for k in range(1,10000):
    X=np.zeros(n)
    for i in range(0,n):
        X[i]=rd.random()**(-1/(i+1)) #X_i suit la loi de Pareto de paramètre (i+1)
    if X[n-1]==np.min(X):
        C=C+1
print(C/10000)
```

Quel est l'objectif de ce programme ?
Si $n = 100$, le programme renvoie 0.0206.
Ceci est-il cohérent avec les résultats précédents ?

Partie III : estimations du paramètre a

Soit a un réel fixé, où $a > 0$.
Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant toutes la loi de Pareto de paramètre a. On note pour tout $k \in \mathbb{N}^*$, $Y_k = \ln(X_k)$ et pour tout $n \in \mathbb{N}$ où $n \geq 3$,

$$S_n = \sum_{k=1}^{n} \ln(X_k), \qquad \overline{Y_n} = \frac{1}{n}\sum_{k=1}^{n} Y_k, \qquad T_n = \frac{n-1}{S_n} = \frac{a.(n-1)}{a.S_n}$$

1. Justifier que la suite $(\overline{Y_n})$ converge en probabilité vers une variable certaine à préciser. (on pourra utiliser le I.5.)
2. (a) Justifier que la variable $a.S_n$ suit la loi $\gamma(n)$ (loi petit gamma de paramètre n) Préciser une densité de $a.S_n$.

   (b) Justifier que la suite de VAR $(\sqrt{n}.(a.\overline{Y_n}-1))$ converge en loi vers une variable aléatoire M suivant la loi normale centrée réduite.
3. (a) Prouver que l'estimateur T_n est un estimateur sans biais de a.

   (b) Prouver que T_n admet un moment d'ordre 2 égal à $E(T_n^2) = \dfrac{(n-1).a^2}{n-2}$ et que T_n est convergent.
4. On suppose dans cette question et uniquement dans cette question que $1 < a < 2$.
   A l'aide de l'inégalité de Bienaymé-Tchebychev, déterminer un entier $n \geq 3$ tel que l'intervalle $[T_n - \frac{1}{10}, T_n + \frac{1}{10}]$ soit un intervalle de confiance de a au niveau de confiance 0.95.
5. On note, pour tout $n \in \mathbb{N}$ où $n \geq 3$, $Z_n = \frac{n}{S_n} = \frac{1}{\overline{Y_n}}$ et $\Delta_n = \frac{\sqrt{n}.(Z_n-a)}{a}$.

   On admet pour la fin de l'exercice que la suite (Δ_n) converge en loi vers une variable N où N suit une loi normale centrée réduite.

   (a) Soit Φ la fonction de répartition de N, où $N \hookrightarrow \mathcal{N}(0,1)$.
   Soit $\alpha \in]0,1[$. Justifier qu'il existe un unique réel t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

   (b) En déduire que $[\dfrac{\sqrt{n}}{\sqrt{n}+t_\alpha}.\dfrac{1}{\overline{Y_n}}, \dfrac{\sqrt{n}}{\sqrt{n}-t_\alpha}.\dfrac{1}{\overline{Y_n}}]$ est un intervalle de confiance asymptotique de a au niveau de risque α.