

DS9 (version B)

I. Exercice : suites et calcul matriciel (ESSEC III - 2007)

Soient (u_n) , (v_n) et (w_n) les trois suites définies sur \mathbb{N} par leur premier terme :

$$u_0 = 1, \quad v_0 = 0, \quad w_0 = 0$$

et les relations de récurrence :

$$\begin{cases} u_{n+1} = 3u_n - v_n + w_n \\ v_{n+1} = u_n + 2v_n \\ w_{n+1} = v_n + w_n \end{cases}$$

Pour tout entier naturel n , on pose : $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$, et on note A la matrice $\begin{pmatrix} 3 & -1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

1. a) Reconnaître, pour tout entier naturel n , le produit AX_n .

Démonstration.

$$\forall n \in \mathbb{N}, A X_n = \begin{pmatrix} 3 & -1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = \begin{pmatrix} 3u_n - v_n + w_n \\ u_n + 2v_n \\ v_n + w_n \end{pmatrix} = X_{n+1}$$

□

b) En déduire l'expression de X_n en fonction des matrices A , X_0 et de l'entier naturel n .

Démonstration.

Démontrons par récurrence que : $\forall n \in \mathbb{N}, \mathcal{P}(n)$ où $\mathcal{P}(n) : X_n = A^n X_0$.

► **Initialisation :**

$$A^0 X_0 = I X_0 = X_0$$

D'où $\mathcal{P}(0)$.

► **Hérédité :** soit $n \in \mathbb{N}$.

Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (i.e. $X_{n+1} = A^{n+1} X_0$).

$$\begin{aligned} X_{n+1} &= A X_n && \text{(d'après la question précédente)} \\ &= A (A^n X_0) && \text{(par hypothèse de récurrence)} \\ &= A^{n+1} X_0 \end{aligned}$$

D'où $\mathcal{P}(n+1)$.

Par principe de récurrence : $\forall n \in \mathbb{N}, X_n = A^n X_0$.

□

2. a) Démontrer que A admet une seule valeur propre.

Démonstration.

Soit $\lambda \in \mathbb{R}$.

Cherchons les valeurs λ pour lesquelles $A - \lambda I$ est non inversible.

$$\begin{aligned}
 \operatorname{rg}(A - \lambda I) &= \operatorname{rg} \left(\begin{pmatrix} 3 - \lambda & -1 & 1 \\ 1 & 2 - \lambda & 0 \\ 0 & 1 & 1 - \lambda \end{pmatrix} \right) \\
 &\stackrel{L_1 \leftrightarrow L_2}{=} \operatorname{rg} \left(\begin{pmatrix} 1 & 2 - \lambda & 0 \\ 3 - \lambda & -1 & 1 \\ 0 & 1 & 1 - \lambda \end{pmatrix} \right) \\
 &\stackrel{L_2 \leftarrow L_2 - (3 - \lambda)L_1}{=} \operatorname{rg} \left(\begin{pmatrix} 1 & 2 - \lambda & 0 \\ 0 & -1 - (2 - \lambda)(3 - \lambda) & 1 \\ 0 & 1 & 1 - \lambda \end{pmatrix} \right) \\
 &\stackrel{L_3 \leftrightarrow L_2}{=} \operatorname{rg} \left(\begin{pmatrix} 1 & 2 - \lambda & 0 \\ 0 & 1 & 1 - \lambda \\ 0 & -1 - (2 - \lambda)(3 - \lambda) & 1 \end{pmatrix} \right) \\
 &\stackrel{L_3 \leftrightarrow L_3 + q(\lambda)L_2}{=} \operatorname{rg} \left(\begin{pmatrix} 1 & 2 - \lambda & 0 \\ 0 & 1 & 1 - \lambda \\ 0 & 0 & 1 + (1 - \lambda)q(\lambda) \end{pmatrix} \right)
 \end{aligned}$$

où : $q(\lambda) = 1 + (2 - \lambda)(3 - \lambda)$. Ainsi :

$$\begin{aligned}
 1 + (1 - \lambda)q(\lambda) &= 1 + (1 - \lambda) \left(1 + (2 - \lambda)(3 - \lambda) \right) \\
 &= 1 + (1 - \lambda) + (1 - \lambda)(2 - \lambda)(3 - \lambda) \\
 &= (2 - \lambda) + (1 - \lambda)(2 - \lambda)(3 - \lambda) \\
 &= (2 - \lambda) \left(1 + (1 - \lambda)(3 - \lambda) \right) \\
 &= (2 - \lambda) \left(1 + (3 - 4\lambda + \lambda^2) \right) \\
 &= (2 - \lambda) (4 - 4\lambda + \lambda^2) \\
 &= (2 - \lambda) (2 - \lambda)^2 = (2 - \lambda)^3
 \end{aligned}$$

La réduite obtenue est triangulaire (supérieure).

Elle est donc non inversible si et seulement si l'un de ses coefficients diagonaux est nul.

On en déduit :

$$\operatorname{rg}(A - \lambda \cdot I_3) < 3 \Leftrightarrow (2 - \lambda)^3 = 0 \Leftrightarrow \lambda = 2$$

On en déduit : $\operatorname{Sp}(A) = \{2\}$.

Commentaire

- Il faut faire attention aux opérations que l'on effectue lors du pivot de Gauss. Il ne faut **JAMAIS** prendre comme pivot une quantité dépendant de λ . C'est d'ailleurs pour cela que l'on commence par échanger les lignes L_1 et L_2 . Sinon, on serait obligé de faire l'opération :

$$L_2 \leftarrow (3 - \lambda) L_2 - L_1$$

Cette opération n'est autorisée que si $\lambda \neq 3$. En effet, si $\lambda = 3$, on effectue l'opération : $L_2 \leftarrow -L_1$ et on tue ainsi la ligne L_2 . Si l'on souhaite faire ce type d'opération, on est alors obligé d'effectuer une disjonction de cas (soit $\lambda = 3$, soit $\lambda \neq 3$).

- L'opération $L_3 \leftarrow L_3 + q(\lambda) L_2$ est quant à elle bien autorisée. Au pire, $q(\lambda) = 0$ et l'opération devient alors : $L_3 \leftarrow L_3$ (pas de perte d'information).
- Grâce à l'introduction de la quantité $q(\lambda)$, on a repoussé le calcul en fin de question. Il est conseillé, lorsque c'est possible, d'agir ainsi. Cela évite de commettre des erreurs.
- Si on regarde attentivement la suite de l'énoncé, il est évident que la seule valeur propre possible de A est 2. En effet, on sait que la matrice représentative d'un endomorphisme f a même spectre que f . Ainsi :

$$\text{Sp}(A) = \text{Sp}(f) = \text{Sp}(T)$$

et comme T est triangulaire supérieure et que sa diagonale ne contient que des 2, on obtient directement que $\text{Sp}(T) = \{2\}$.

- Cela ne signifie pas que l'on peut rédiger la **2.a)** avec comme argument : « d'après la question **2.a)** ... ». Par contre, si on ne parvient pas à déterminer la seule valeur propre en **2.a)**, on peut le signifier au correcteur et admettre que cette valeur propre est 2 (ce qui est validé par la **3.a)**). On peut même alors démontrer que 2 est valeur propre dans la **2.a)** en démontrant que $A - 2I$ est non inversible.

□

- b) Déterminer le sous-espace vectoriel propre de A associé à l'unique valeur propre.
 La matrice A est-elle diagonalisable ?

Démonstration.

- Déterminons $E_2(A)$, le sous-espace propre de A associé à la valeur propre 2.

Soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$.

$$\begin{aligned} X \in E_2(A) &\iff (A - 2I) X = 0 \\ &\iff \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff \begin{cases} x - y + z = 0 \\ x & & = 0 \\ & y - z = 0 \end{cases} \\ &\stackrel{L_1 \leftrightarrow L_2}{\iff} \begin{cases} x & & = 0 \\ x - y + z = 0 \\ & y - z = 0 \end{cases} \\ &\iff \begin{cases} x & & = 0 \\ & y - z = 0 \end{cases} \\ &\iff \begin{cases} x & = 0 \\ & y = z \end{cases} \end{aligned}$$

- On en déduit que :

$$\begin{aligned} E_2(A) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R}) \mid AX = 2X \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x = 0 \text{ et } y = z \right\} = \left\{ \begin{pmatrix} 0 \\ z \\ z \end{pmatrix} \mid z \in \mathbb{R} \right\} \\ &= \left\{ z \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mid z \in \mathbb{R} \right\} = \text{Vect} \left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \end{aligned}$$

- La famille $\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$ est :

- × génératrice de $E_2(A)$.
- × libre car constituée uniquement d'un vecteur non nul.

C'est donc une base de $E_2(A)$. Ainsi : $\dim(E_2(A)) = \text{Card} \left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) = 1$.

$\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \text{ est une base de } E_2(A).$

On obtient :

- × $A \in \mathcal{M}_3(\mathbb{R})$,
- × $\dim(E_2(A)) = 1 \neq 3$.

On en déduit que A n'est pas diagonalisable.

Commentaire

- La question de la diagonalisabilité de A peut être traitée même si l'on n'a pas déterminé $\text{Sp}(A)$. L'important ici est que la matrice A ne possède qu'une valeur propre (donné dans l'énoncé). Même si on ne la connaît pas, on la note λ . On effectue alors le raisonnement classique suivant.
- Raisonnons par l'absurde. On suppose que la matrice A est diagonalisable. Alors il existe alors :
 - × $P \in \mathcal{M}_3(\mathbb{R})$ inversible,
 - × $D \in \mathcal{M}_3(\mathbb{R})$ diagonale, dont les coefficients diagonaux sont les valeurs propres de A , telles que : $A = PDP^{-1}$. Or $\text{Sp}(A) = \{\lambda\}$. Donc :

$$A = P \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} P^{-1} = \lambda \cdot P I_2 P^{-1} = \lambda \cdot PP^{-1} = \lambda \cdot I_2$$

ce qui est absurde.

- On peut « lire » directement sur la matrice $A - 2I$ que $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ est vecteur propre de A .

On rappelle :

$$A - 2I = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

La matrice $A - 2I$ n'est pas inversible car ses colonnes vérifient : $C_2 = -C_3$.
De manière générale :

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x C_1 + y C_2 + z C_3$$

En prenant $x = 0$ (on laisse la 1ère colonne de côté) et $y = z = 1$, on obtient la somme des colonnes C_2 et C_3 , qui n'est autre que le vecteur colonne nul.

□

3. On note f l'endomorphisme de \mathbb{R}^3 canoniquement associé à A , c'est-à-dire tel que A soit la matrice de f dans la base canonique \mathcal{B} de \mathbb{R}^3 .

a) Déterminer une base (e'_1, e'_2, e'_3) de \mathbb{R}^3 telle que la matrice T de f dans cette base vérifie :

$$T = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

et que les vecteurs e'_1, e'_2, e'_3 aient respectivement pour troisième composante 1, -1 et 2. On notera dorénavant \mathcal{B}' la base (e'_1, e'_2, e'_3) .

Démonstration.

• Il s'agit ici de déterminer les coordonnées de e'_1, e'_2 et e'_3 . Dans la suite, on note :

$$E'_1 = \text{Mat}_{\mathcal{B}}(e'_1) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad E'_2 = \text{Mat}_{\mathcal{B}}(e'_2) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad E'_3 = \text{Mat}_{\mathcal{B}}(e'_3) = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

L'énoncé fixe les contraintes suivantes : $x_3 = 1, y_3 = -1$ et $z_3 = 2$.

• D'après l'énoncé, $T = \text{Mat}_{(e'_1, e'_2, e'_3)}(f)$. Cela signifie :

$$\times \text{Mat}_{(e'_1, e'_2, e'_3)}(f(e'_1)) = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \text{ et donc : } f(e'_1) = 2 \cdot e'_1 + 0 \cdot e'_2 + 0 \cdot e'_3.$$

Cette dernière égalité se traduit dans la base \mathcal{B} :

$$\begin{aligned} \text{Mat}_{\mathcal{B}}(f(e'_1)) &= 2 \text{Mat}_{\mathcal{B}}(e'_1) \\ \Leftrightarrow AE'_1 &= 2E'_1 \end{aligned}$$

Autrement dit, E'_1 est un vecteur propre de A associé à la valeur propre 2.

Il suffit alors de choisir : $E'_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ ce qui correspond à $e'_1 = (0, 1, 1)$.

$$\times \text{Mat}_{(e'_1, e'_2, e'_3)}(f(e'_2)) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \text{ et donc : } f(e'_2) = 1 \cdot e'_1 + 2 \cdot e'_2 + 0 \cdot e'_3.$$

Cette dernière égalité se traduit dans la base \mathcal{B} :

$$\begin{aligned} \text{Mat}_{\mathcal{B}}(f(e'_2)) &= \text{Mat}_{\mathcal{B}}(e'_1) + 2 \cdot \text{Mat}_{\mathcal{B}}(e'_2) \\ \Leftrightarrow AE'_2 &= E'_1 + 2 \cdot E'_2 \\ \Leftrightarrow \begin{pmatrix} 3y_1 - y_2 + y_3 \\ y_1 + 2y_2 \\ y_2 + y_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ \Leftrightarrow \begin{pmatrix} y_1 - y_2 + y_3 \\ y_1 \\ y_2 - y_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\ \Leftrightarrow \begin{pmatrix} y_1 - y_2 \\ y_1 \\ y_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad (\text{car } y_3 = -1) \end{aligned}$$

On obtient le système suivant :

$$\begin{cases} y_1 - y_2 = 1 \\ y_1 = 1 \\ y_2 = 0 \end{cases} \Leftrightarrow \begin{cases} y_1 = 1 \\ y_2 = 0 \end{cases}$$

Ce qui correspond à : $E'_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ et donc à $e'_2 = (1, 0, -1)$.

$$\times \text{Mat}_{(e'_1, e'_2, e'_3)}(f(e'_3)) = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ et donc : } f(e'_2) = 0 \cdot e'_1 + 1 \cdot e'_2 + 2 \cdot e'_3.$$

Cette dernière égalité se traduit dans la base \mathcal{B} :

$$\begin{aligned} \text{Mat}_{\mathcal{B}}(f(e'_3)) &= \text{Mat}_{\mathcal{B}}(e'_2) + 2 \cdot \text{Mat}_{\mathcal{B}}(e'_3) \\ \Leftrightarrow AE'_3 &= E'_2 + 2 \cdot E'_3 \\ \Leftrightarrow \begin{pmatrix} 3z_1 - z_2 + z_3 \\ z_1 + 2z_2 \\ z_2 + z_3 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 2 \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \\ \Leftrightarrow \begin{pmatrix} z_1 - z_2 + z_3 \\ z_1 \\ z_2 - z_3 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ \Leftrightarrow \begin{pmatrix} z_1 - z_2 \\ z_1 \\ z_2 \end{pmatrix} &= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad (\text{car } z_3 = 2) \\ \Leftrightarrow \begin{cases} z_1 - z_2 = -1 \\ z_1 = 0 \\ z_2 = 1 \end{cases} &\Leftrightarrow \begin{cases} z_1 = 0 \\ z_2 = 1 \end{cases} \end{aligned}$$

Ce qui correspond à : $E'_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ et donc à $e'_3 = (0, 1, 2)$.

Commentaire

- Comme on l'a vu dans la question précédente, la matrice A et donc l'endomorphisme f n'est pas diagonalisable. Il n'existe donc pas de base dans laquelle la matrice représentant f est diagonale.
- Dans ce cas, on se rabat sur une propriété plus faible : existe-t-il une base dans laquelle la représentation matricielle de f serait triangulaire supérieure? Cette propriété est beaucoup plus simple à obtenir notamment si l'on accepte d'utiliser des matrices dont les coefficients sont complexes (hors de notre portée en ECE).
 On parle alors de *trigonaliser* (on dit aussi *triangulariser*) la matrice A .
- Si un endomorphisme $f : E \rightarrow E$ est triangularisable, comment le triangularise-t-on? Notons $\lambda_1, \dots, \lambda_m$ les valeurs propres de f . On cherche alors une base de chaque espace propre E_{λ_i} et on considère la famille obtenue en concaténant toutes ces bases.
 Cette famille N'EST PAS une base de E . Si tel était le cas, on aurait formé une base de vecteurs propres et donc E serait diagonalisable.
 Par contre, cette famille est libre. On peut alors la compléter en une base de E .
 Sans entrer dans les détails, on peut faire en sorte (en choisissant correctement les vecteurs qu'on ajoute) que la matrice représentative de f dans la base \mathcal{B}' soit triangulaire supérieure.
- C'est la méthode développée dans cette question. Ici, f n'a qu'une valeur propre. L'espace propre $E_2(f)$ a pour base la famille $((0, 1, 1))$. On complète alors cette famille en ajoutant e'_2 et e'_3 choisis de sorte à ce que la matrice représentant f dans la base \mathcal{B}' ainsi formée soit triangulaire.

b) À l'aide de la formule du binôme de Newton et de la décomposition suivante T :

$$T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

déterminer l'expression de la matrice T^n en fonction de l'entier naturel n .

Démonstration.

Notons : $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2I$ et $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

- Remarquons tout d'abord :

$$N^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$N^3 = N^2 \times N = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Ainsi, $N^3 = 0_{\mathcal{M}_3(\mathbb{R})}$ et, par une récurrence immédiate : $\forall k \geq 3, N^k = 0$.

- De plus, comme la matrice identité I commute avec toute matrice :

$$DN = 2 \cdot IN = 2 \cdot NI = N(2 \cdot I) = ND$$

Ainsi D et N commutent.

On peut donc appliquer la formule du binôme de Newton, pour $n \geq 2$:

$$\begin{aligned}
 T^n &= (D + N)^n \\
 &= \sum_{k=0}^n \binom{n}{k} N^k D^{n-k} \\
 &= \sum_{k=0}^2 \binom{n}{k} N^k D^{n-k} + \sum_{k=3}^n \binom{n}{k} N^k D^{n-k} && \text{(ce découpage est valable car } n \geq 2) \\
 &= \sum_{k=0}^2 \binom{n}{k} N^k D^{n-k} && \text{(car on a montré : } \forall k \geq 3, N^k = 0) \\
 &= \binom{n}{0} N^0 D^n + \binom{n}{1} N^1 D^{n-1} + \binom{n}{2} N^2 D^{n-2} \\
 &= (2I)^n + n N (2I)^{n-1} + \frac{n(n-1)}{2} N^2 (2I)^{n-2} \\
 &= 2^n I + 2^{n-1} n N + 2^{n-2} \frac{n(n-1)}{2} N^2 && \text{(car } (2I)^m = 2^m I^m = 2^m I) \\
 &= 2^{n-3} \left(2^3 I + 2^2 n N + \frac{n(n-1)}{2} N^2 \right)
 \end{aligned}$$

• D'où :

$$\begin{aligned}
 &2^3 I + 2^2 n N + n(n-1) N^2 \\
 &= 2^3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 2^2 n \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + n(n-1) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2^3 & 2^2 n & n(n-1) \\ 0 & 2^3 & 2^2 n \\ 0 & 0 & 2^3 \end{pmatrix}
 \end{aligned}$$

• Enfin :

- × si $n = 0$, $T^0 = I$.
- × si $n = 1$, $T^1 = T$.

Ainsi : $T^0 = I$, $T^1 = T$ et pour tout $n \geq 2$, $T^n = 2^{n-3} \begin{pmatrix} 2^3 & 2^2 n & n(n-1) \\ 0 & 2^3 & 2^2 n \\ 0 & 0 & 2^3 \end{pmatrix}$.

Commentaire

- Comme noté dans la démonstration, l'hypothèse $n \geq 2$ est essentielle pour pouvoir découper la somme. Les cas $n = 0$ et $n = 1$ doivent donc être traités à part.
- Ici, la matrice N vérifie : $\forall k \geq 3, N^k = 0$. Elle est dite nilpotente d'ordre 3 (ce terme n'est pas au programme et il est préférable de ne pas l'utiliser dans une copie).

4. Soit P la matrice de passage de la base canonique \mathcal{B} à la base \mathcal{B}' .

a) Exprimer A en fonction de T , P et P^{-1} , puis A^n en fonction des mêmes matrices et de l'entier naturel n .

Démonstration.

• Tout d'abord :

$$\begin{aligned}
 \text{Mat}_{\mathcal{B}}(f) &= P_{\mathcal{B},\mathcal{B}'} \times \text{Mat}'_{\mathcal{B}}(f) \times P_{\mathcal{B},\mathcal{B}'} \\
 \parallel & \qquad \qquad \parallel & \qquad \qquad \parallel & \qquad \qquad \parallel \\
 A &= P \times T \times P^{-1}
 \end{aligned}$$

$$A = PTP^{-1}$$

- Démontrons alors par récurrence que : $\forall n \in \mathbb{N}, \mathcal{P}(n)$ où $\mathcal{P}(n) : A^n = PT^n P^{-1}$.

► **Initialisation :**

- D'une part : $A^0 = I$.
- D'autre part : $PT^0 P^{-1} = PP^{-1} = I$.

D'où $\mathcal{P}(0)$.

► **Hérédité :** soit $n \in \mathbb{N}$.

Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (i.e. $A^{n+1} = PT^{n+1} P^{-1}$).

$$\begin{aligned}
 A^{n+1} &= A \times A^n \\
 &= A \times PT^n P^{-1} && \text{(par hypothèse de récurrence)} \\
 &= PTP^{-1} \times PT^n P^{-1} && \text{(d'après la question précédente)} \\
 &= PT \times P^{-1}P \times T^n P^{-1} \\
 &= PT \times T^n P^{-1} \\
 &= PT^{n+1} P^{-1}
 \end{aligned}$$

D'où $\mathcal{P}(n+1)$.

Par principe de récurrence : $\forall n \in \mathbb{N}, A^n = PT^n P^{-1}$.

Commentaire

- On a développé la démonstration du résultat : $\forall n \in \mathbb{N}, A^n = PT^n P^{-1}$.
- Cependant ce résultat n'étant pas directement donné dans l'énoncé, on peut penser qu'invoquer « une récurrence immédiate » est suffisant. □

- b) Calculer P^{-1} (les calculs devront figurer sur la copie).

Démonstration.

- Par définition : $P = P_{\mathcal{B}, \mathcal{B}'}$.

Cette matrice s'obtient en concaténant les vecteurs de la base \mathcal{B}' exprimés dans la base \mathcal{B} .

Ainsi : $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$.

- On utilise la méthode de Gauss-Jordan.

$$\left(\begin{array}{ccc|ccc}
 0 & 1 & 0 & 1 & 0 & 0 \\
 1 & 0 & 1 & 0 & 1 & 0 \\
 1 & -1 & 2 & 0 & 0 & 1
 \end{array} \right)$$

On effectue l'opération $L_1 \leftrightarrow L_2$. On obtient :

$$\left(\begin{array}{ccc|ccc}
 1 & 0 & 1 & 0 & 1 & 0 \\
 0 & 1 & 0 & 1 & 0 & 0 \\
 1 & -1 & 2 & 0 & 0 & 1
 \end{array} \right)$$

On effectue l'opération $L_3 \leftarrow L_3 - L_1$. On obtient :

$$\left(\begin{array}{ccc|ccc}
 1 & 0 & 1 & 0 & 1 & 0 \\
 0 & 1 & 0 & 1 & 0 & 0 \\
 0 & -1 & 1 & 0 & -1 & 1
 \end{array} \right)$$

On effectue l'opération $L_3 \leftrightarrow L_3 + L_2$. On obtient :

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array} \right)$$

On effectue l'opération $L_1 \leftarrow L_1 - L_3$. On obtient :

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 2 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array} \right)$$

Ainsi P est inversible d'inverse $P^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

□

c) Déterminer les expressions de u_n, v_n, w_n en fonction de l'entier naturel n .

Démonstration.

Soit $n \in \mathbb{N}$.

- D'après la question 1. et 4.a) : $\begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = X_n = A^n X_0 = P T^n P^{-1} X_0 = P T^n P^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

- On en déduit, grâce à la question 3.b) et à la précédente :

$$\begin{aligned} \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} &= 2^{n-3} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 8 & 4n & n(n-1) \\ 0 & 8 & 4n \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} -1 & 2 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= 2^{n-3} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 8 & 4n & n(n-1) \\ 0 & 8 & 4n \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \\ &= 2^{n-3} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} -8 + 4n + n(n-1) \\ 8 + 4n \\ 8 \end{pmatrix} \\ &= 2^{n-3} \begin{pmatrix} 8 + 4n \\ 4n + n(n-1) \\ n(n-1) \end{pmatrix} = 2^{n-3} \begin{pmatrix} 4(n+2) \\ n(n+3) \\ n(n-1) \end{pmatrix} \end{aligned}$$

Ainsi, pour tout $n \in \mathbb{N}$: $u_n = 2^{n-1} (n+2)$, $v_n = 2^{n-3} n(n+3)$ et $w_n = 2^{n-3} n(n-1)$.

Commentaire

- L'énoncé ne demande à aucun moment le calcul de A^n . Il ne faut pas le faire : ce serait une perte de temps inutile. Ici, il y a un bon ordre pour faire les calculs :
 - × si on va de gauche à droite, on fait 2 produits successifs de matrices 3×3 . Puis un produit d'une 3×3 par un vecteur 3×1 .
 - × si on va de droite à gauche, on ne fait que des produits successifs d'une matrice 3×3 par un vecteur 3×1 .
- Évidemment, si l'énoncé demande de déterminer A^n , on ne peut y couper ...

□

Prenons un peu de recul . . .

- On avait affaire ici à un énoncé de facture classique. L'endomorphisme $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ est présenté sous sa forme matricielle A initialement. On détermine alors les valeurs propres et espaces propres associés de cette matrice A . Elle n'est pas diagonalisable mais on démontre qu'on peut la trigonaliser (triangulariser).
- Pour autant, cet énoncé n'est pas simple simple :
 - × la première difficulté est de déterminer les valeurs propres. Il faut appliquer méticuleusement l'algorithme du pivot de Gauss sur la matrice paramétrée $A - \lambda I$.
 - × l'obtention de la base (e'_1, e'_2, e'_3) n'est pas du tout guidée. Il faut avoir les idées claires sur la notion de changement de base ainsi que sur l'isomorphisme de représentation (passerelle endomorphisme / matrice).
- Si on ne parvient pas à déterminer $\text{Sp}(A)$ (question **2.a**), il paraît compliqué de faire entièrement les questions **2.b**, **3.a**, **4.b** et **4.c**. Cependant, cela n'empêche pas de faire TOUTES les autres !
- Enfin, comme précisé dans une précédente remarque, par simple lecture de la matrice T , on obtient : $\text{Sp}(T) = \{2\}$ et donc $\text{Sp}(A) = \{2\}$. Ceci permet de traiter la question **2.b**. Une lecture précise de l'énoncé peut permettre de se tirer d'un mauvais pas . . .

Problème (ESSEC I 2007)

Dans tout l'exercice, les variables aléatoires sont définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

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 : $x \mapsto \mathbb{P}([X > x])$ (appelée *fonction de survie* de X).

Démonstration.

- Comme $X \hookrightarrow \mathcal{E}(\lambda)$, sa fonction de répartition est donnée par :

$$F_X : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

- Soit $x \in \mathbb{R}$. Deux cas se présentent alors :

× si $x < 0$:

$$\mathbb{P}([X > x]) = 1 - \mathbb{P}([X \leq x]) = 1 - F_X(x) = 1 - 0 = 1$$

× si $x \geq 0$:

$$\mathbb{P}([X > x]) = 1 - F_X(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

Finalement, la fonction de survie est : $x \mapsto \begin{cases} 1 & \text{si } x < 0 \\ e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$.

□

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

Démonstration.

Soient $x > 0$ et $y > 0$.

- D'après la question précédente, $\mathbb{P}([X > x]) > 0$.
- On peut donc calculer la probabilité conditionnelle suivante :

$$\begin{aligned} \mathbb{P}_{[X > x]}([X > x + y]) &= \frac{\mathbb{P}([X > x] \cap [X > x + y])}{\mathbb{P}([X > x])} \\ &= \frac{\mathbb{P}([X > x + y])}{\mathbb{P}([X > x])} && \text{(car, comme } y > 0, \\ & && [X > x + y] \subset [X > x]) \\ &= \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} && \text{(car } x + y \geq 0 \text{ et } x \geq 0) \\ &= \frac{\cancel{e^{-\lambda x}} e^{-\lambda y}}{\cancel{e^{-\lambda x}}} \\ &= e^{-\lambda y} = \mathbb{P}([X > y]) \end{aligned}$$

$\mathbb{P}_{[X > x]}([X > x + y]) = \mathbb{P}([X > y])$

La probabilité que le phénomène ait encore lieu après $x + y$ « heures » sachant qu'il a déjà eut lieu durant x heures ne dépend que de la durée supplémentaire $x + y - x = y$ ajoutée.
 Il est donc sans vieillissement.

Commentaire

Soit $(x, y) \in]0, +\infty[^2$. Démontrons l'inclusion : $[X > x + y] \subset [X > x]$.
 Soit $\omega \in [X > x + y]$, alors : $X(\omega) > x + y$.
 Par transitivité, on obtient :

$$X(\omega) > x + y > x \quad (\text{car } y > 0)$$

Ainsi : $X(\omega) > x$. Ou encore : $\omega \in [X > x]$. □

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 .

Démonstration.

Soit $n \in \mathbb{N}^*$.

- La v.a.r. S_n admet une espérance en tant que somme de variables aléatoires qui admettent toutes une espérance.

$$\begin{aligned} \mathbb{E}(S_n) &= \mathbb{E}\left(\sum_{k=1}^n X_k\right) \\ &= \sum_{k=1}^n \mathbb{E}(X_k) && (\text{par linéarité de l'espérance}) \\ &= \sum_{k=1}^n \frac{1}{\lambda} && (\text{car } X_1, \dots, X_n \text{ ont} \\ &&& \text{même loi } \mathcal{E}(\lambda)) \\ &= n \frac{1}{\lambda} \end{aligned}$$

$$\mathbb{E}(S_n) = \frac{n}{\lambda}$$

- La v.a.r. S_n admet une variance en tant que somme de variables aléatoires qui admettent toutes une variance.

$$\begin{aligned} \mathbb{V}(S_n) &= \mathbb{V}\left(\sum_{k=1}^n X_k\right) \\ &= \sum_{k=1}^n \mathbb{V}(X_k) && (\text{car } X_1, \dots, X_n \text{ sont} \\ &&& \text{indépendantes}) \\ &= \sum_{k=1}^n \frac{1}{\lambda^2} && (\text{car } X_1, \dots, X_n \text{ ont} \\ &&& \text{même loi } \mathcal{E}(\lambda)) \\ &= n \frac{1}{\lambda^2} \end{aligned}$$

$$\mathbb{V}(S_n) = \frac{n}{\lambda^2}$$

□

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

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

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 :

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

Démonstration.

Démontrons par récurrence que : $\forall n \in \mathbb{N}^*$, $\mathcal{P}(n)$

où $\mathcal{P}(n)$: S_n est une variable à densité, de densité f_n .

► **Initialisation** :

- D'une part, par définition : $f_1 : t \mapsto \begin{cases} 0 & \text{si } t < 0; \\ \frac{\lambda}{0!} e^{-\lambda t} t^0 & \text{si } t \geq 0. \end{cases}$

Or, pour tout $t \geq 0$:

$$\frac{\lambda}{0!} e^{-\lambda t} t^0 = \lambda e^{-\lambda t}$$

$$\text{Ainsi : } f_1 : t \mapsto \begin{cases} 0 & \text{si } t < 0; \\ \lambda e^{-\lambda t} & \text{si } t \geq 0. \end{cases} .$$

On reconnaît une densité d'une variable aléatoire de loi $\mathcal{E}(\lambda)$.

- D'autre part : $S_1 = X_1$. Donc : $S_1 \hookrightarrow \mathcal{E}(\lambda)$.

D'où $\mathcal{P}(1)$.

► **Hérédité** : soit $n \in \mathbb{N}^*$.

Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$

(i.e. S_{n+1} est une variable à densité, de densité f_{n+1}).

- Remarquons tout d'abord : $S_{n+1}(\Omega) \subset [0, +\infty[$.
 En effet, pour tout $i \in \llbracket 1, n+1 \rrbracket$, $X_i \in [0, +\infty[$.

- De plus :

$$S_{n+1} = \sum_{k=1}^{n+1} X_k = \left(\sum_{k=1}^n X_k \right) + X_{n+1} = S_n + X_{n+1}$$

On est dans le cadre d'utilisation du théorème fourni par l'énoncé :

- × par hypothèse de récurrence, S_n est une variable à densité, de densité f_n .
- × d'après l'énoncé, X_{n+1} est une variable à densité de densité f_1 , car $X_{n+1} \hookrightarrow \mathcal{E}(\lambda)$.
- × d'après le lemme des coalitions, les v.a.r. S_n et X_{n+1} sont indépendantes, car les v.a.r. X_1, \dots, X_{n+1} sont indépendantes.

- On en déduit que S_{n+1} est une v.a.r. à densité et, pour tout $t \in \mathbb{R}$:

$$f_{S_{n+1}}(t) = f_{S_n + X_{n+1}}(t) = \int_{-\infty}^{+\infty} f_n(x) f_1(t-x) dx$$

Soit $t \in \mathbb{R}$. Deux cas se présentent :

× si $t < 0$, alors, comme $S_{n+1}(\Omega) \subset [0, +\infty[: F_{S_{n+1}}(t) = 0$. On en déduit :

$$f_{S_{n+1}}(t) = F'_{S_{n+1}}(t) = 0$$

× si $t \geq 0$. Cherchons d'abord à savoir sur quel ensemble la fonction $x \mapsto f_n(x) f_1(t-x)$ ne s'annule pas afin de préciser l'intervalle d'intégration.

- Comme, par hypothèse de récurrence, en particulier, f_n est nulle en dehors de $[0, +\infty[:$

$$f_{S_{n+1}}(t) = \int_0^{+\infty} f_n(x) f_1(t-x) dx$$

- De plus, pour tout $x \in \mathbb{R} :$

$$\begin{cases} x \in [0, +\infty[\\ f_1(t-x) \neq 0 \end{cases} \Leftrightarrow \begin{cases} x \in [0, +\infty[\\ (t-x) \in [0, +\infty[\end{cases} \Leftrightarrow \begin{cases} x \geq 0 \\ t-x \geq 0 \end{cases} \Leftrightarrow \begin{cases} x \geq 0 \\ x \leq t \end{cases} \Leftrightarrow 0 \leq x \leq t$$

Ainsi :

$$f_{n+1}(t) = \int_0^t f_n(x) f_1(t-x) dx$$

On obtient :

$$\begin{aligned} f_{S_{n+1}}(t) &= \int_0^t \frac{\lambda^n}{(n-1)!} e^{-\lambda x} x^{n-1} \lambda e^{-\lambda(t-x)} dx \\ &= \frac{\lambda^{n+1}}{(n-1)!} \int_0^t \cancel{e^{-\lambda x}} x^{n-1} e^{-\lambda t} \cancel{e^{\lambda x}} dx \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda t} \int_0^t x^{n-1} dx \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda t} \left[\frac{x^n}{n} \right]_0^t \\ &= \frac{\lambda^{n+1}}{n!} e^{-\lambda t} t^n \end{aligned}$$

$$\text{Finalement : } f_{S_{n+1}} : t \mapsto \begin{cases} 0 & \text{si } t < 0 \\ \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} & \text{si } t \geq 0 \end{cases} .$$

D'où $\mathcal{P}(n+1)$.

Par principe de récurrence, pour tout $n \in \mathbb{N}^*$, S_n est une variable à densité de densité f_n .

Commentaire

Pour bien comprendre la présentation du cas « $t \geq 0$ », remarquons qu'on cherche, comme annoncé, l'ensemble sur lequel produit $f_n(x) f(t-x)$ est non nul. Soit $x \in \mathbb{R}$.

$$\begin{aligned} f_n(x) f_1(t-x) \neq 0 &\Leftrightarrow \{ f_n(x) \neq 0 \text{ et } f_1(t-x) \neq 0 \\ &\Leftrightarrow \begin{cases} x \in [0, +\infty[\\ (t-x) \in [0, +\infty[\end{cases} \Leftrightarrow \begin{cases} x \geq 0 \\ t-x \geq 0 \end{cases} \\ &\Leftrightarrow \begin{cases} x \geq 0 \\ x \leq t \end{cases} \Leftrightarrow 0 \leq x \leq t \end{aligned}$$

Cette présentation est d'ailleurs tout aussi acceptable, mais elle sort un peu des présentations usuelles. C'est pourquoi on lui a préféré la précédente. □

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 :

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

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

3. Vérifier que l'égalité : $\int_{-\infty}^{+\infty} f(t) dt = 1$ est bien satisfaite; calculer l'espérance et la variance de X , en précisant à quelles conditions chacune de ces quantités existe.

Démonstration.

- Tout d'abord, comme f est nulle en dehors de $[b, +\infty[$:

$$\int_{-\infty}^{+\infty} f(t) dt = \int_b^{+\infty} f(t) dt$$

- La fonction f est continue par morceaux sur $[b, +\infty[$.
- Soit $A \in [b, +\infty[$.

$$\begin{aligned} \int_b^A f(t) dt &= \int_b^A a \frac{b^a}{t^{a+1}} dt = a b^a \int_b^A t^{-a-1} dt \\ &= a b^a \left[\frac{t^{-a}}{-a} \right]_b^A = -b^a \left[\frac{1}{t^a} \right]_b^A \quad (\text{car } a \neq 0) \\ &= -b^a \left(\frac{1}{A^a} - \frac{1}{b^a} \right) \\ &\xrightarrow{A \rightarrow +\infty} -b^a \left(0 - \frac{1}{b^a} \right) = 1 \quad (\text{car } a > 0) \end{aligned}$$

Ainsi, l'intégrale impropre $\int_{-\infty}^{+\infty} f(t) dt$ est convergente et vaut 1.

Commentaire

- La question demande de démontrer la convergence d'une intégrale impropre mais exige aussi la valeur de cette intégrale. Dans ce cas, comme rédigé ici, la convergence est une conséquence du résultat obtenu par calcul.
- La démonstration de la convergence était simple puisque $\int_b^{+\infty} \frac{1}{t^{a+1}} dt$ est une intégrale de Riemann, impropre en $+\infty$ ($b > 0$) et d'exposant $a + 1 > 1$. Elle est donc convergente.

- La v.a.r. X admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour des calculs de moments du type $\int_{-\infty}^{+\infty} t^m f(t) dt$.

× Comme f est nulle en dehors de $[b, +\infty[$:

$$\int_{-\infty}^{+\infty} t f(t) dt = \int_b^{+\infty} t f(t) dt$$

× De plus, pour tout $t \in [b, +\infty[$:

$$t f(t) = t \frac{a b^a}{t^{a+1}} = a b^a \frac{1}{t^a}$$

- × Or $\int_b^{+\infty} \frac{1}{t^a} dt$ est une intégrale de Riemann, impropre en $+\infty$ ($b > 0$), d'exposant a . Elle est donc convergente si et seulement si $a > 1$.

On en déduit que la v.a.r. X admet une espérance si et seulement si $a > 1$.

- Supposons alors $a > 1$.
Soit $A \in [b, +\infty[$.

$$\begin{aligned} \int_b^A t f(t) dt &= a b^a \int_b^A t^{-a} dt \\ &= a b^a \left[\frac{t^{-a+1}}{-a+1} \right]_b^A \quad (\text{car } a \neq 1) \\ &= -\frac{a b^a}{a-1} \left[\frac{1}{t^{a-1}} \right]_b^A \\ &= -\frac{a b^a}{a-1} \left(\frac{1}{A^{a-1}} - \frac{1}{b^{a-1}} \right) \\ &\xrightarrow{A \rightarrow +\infty} -\frac{a b^a}{a-1} \left(0 - \frac{1}{b^{a-1}} \right) \quad (\text{car } a-1 > 0) \end{aligned}$$

Enfin :

$$-\frac{a b^a}{a-1} \left(-\frac{1}{b^{a-1}} \right) = \frac{a b^a}{(a-1)b^{a-1}} = \frac{a b}{a-1}$$

Ainsi, si $a > 1$: $\mathbb{E}(X) = \frac{a b}{a-1}$.

Commentaire

Insistons sur le fait que les limites :

$$\lim_{A \rightarrow +\infty} A^{a-1} = +\infty \quad \text{et donc} \quad \lim_{A \rightarrow +\infty} \frac{1}{A^{a-1}} = 0$$

ne sont valables que si $a - 1 > 0$.

- On procède de même pour la variance.

La v.a.r. X admet une variance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t^2 f(t) dt$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour des calculs de moments du type $\int_{-\infty}^{+\infty} t^m f(t) dt$.

× Comme f est nulle en dehors de $[b, +\infty[$:

$$\int_{-\infty}^{+\infty} t^2 f(t) dt = \int_b^{+\infty} t^2 f(t) dt$$

× De plus, pour tout $t \in [b, +\infty[$:

$$t^2 f(t) = t^2 \frac{a b^a}{t^{a+1}} = a b^a \frac{1}{t^{a-1}}$$

× Or $\int_b^{+\infty} \frac{1}{t^{a-1}} dt$ est une intégrale de Riemann, impropre en $+\infty$ ($b > 0$), d'exposant $a - 1$. Elle est donc convergente si et seulement si $a - 1 > 1$.

On en déduit que la v.a.r. X admet une variance si et seulement si $a > 2$.

- Supposons alors $a > 2$.
 Soit $A \in [b, +\infty[$.

$$\begin{aligned} \int_b^A t f(t) dt &= a b^a \int_b^A t^{-a+1} dt \\ &= a b^a \left[\frac{t^{-a+2}}{-a+2} \right]_b^A \quad (\text{car } a \neq 2) \\ &= -\frac{a b^a}{a-2} \left[\frac{1}{t^{a-2}} \right]_b^A \\ &= -\frac{a b^a}{a-2} \left(\frac{1}{A^{a-2}} - \frac{1}{b^{a-2}} \right) \\ &\xrightarrow{A \rightarrow +\infty} -\frac{a b^a}{a-2} \left(0 - \frac{1}{b^{a-2}} \right) \quad (\text{car } a-2 > 0) \end{aligned}$$

Enfin :

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

Ainsi, si $a > 2$: $\mathbb{E}(X^2) = \frac{a b^2}{a-2}$.

- Ainsi, par la formule de Koenig-Huygens :

$$\begin{aligned}
 \mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\
 &= \frac{ab^2}{a-2} - \left(\frac{ab}{a-1}\right)^2 \\
 &= ab^2 \left(\frac{1}{a-2} - \frac{a}{(a-1)^2}\right) \\
 &= ab^2 \left(\frac{(a-1)^2 - a(a-2)}{(a-2)(a-1)^2}\right) \\
 &= ab^2 \left(\frac{(\cancel{a^2} - \cancel{2a} + 1) - (\cancel{a^2} - \cancel{2a})}{(a-2)(a-1)^2}\right) \\
 &= ab^2 \frac{1}{(a-2)(a-1)^2}
 \end{aligned}$$

Ainsi, si $a > 2$: $\mathbb{V}(X) = \frac{ab^2}{(a-2)(a-1)^2}$.

□

4. Déterminer la fonction de répartition de X . Préciser la fonction de survie : $x \mapsto \mathbb{P}([X > x])$.

Démonstration.

Soit $x \in \mathbb{R}$. Deux cas se présentent :

- × si $x \in]-\infty, b[$. Comme f est nulle en dehors de $[b, +\infty[$:

$$F_X(x) = \mathbb{P}([X \leq x]) = \int_{-\infty}^x f(t) dt = 0$$

- × si $x \in [b, +\infty[$:

$$\begin{aligned}
 F_X(x) &= \int_{-\infty}^x f(t) dt \\
 &= \int_b^x f(t) dt && \text{(car } f \text{ est nulle en dehors de } [b, +\infty[) \\
 &= \int_b^x a \frac{b^a}{t^{a+1}} dt \\
 &= ab^a \int_b^x t^{-a-1} dt \\
 &= \cancel{a} b^a \left[\frac{t^{-a}}{-\cancel{a}} \right]_b^x && \text{(car } a \neq 0) \\
 &= -b^a \left(\frac{1}{x^a} - \frac{1}{b^a} \right) = 1 - \left(\frac{b}{x} \right)^a
 \end{aligned}$$

Finalement : $F_X : x \mapsto \begin{cases} 0 & \text{si } x < b \\ 1 - \left(\frac{b}{x}\right)^a & \text{si } x \geq b \end{cases}$

Déterminons la fonction de survie de X .

Soit $x \in \mathbb{R}$. Deux cas se présentent :

× si $x \in]-\infty, b[$:

$$\mathbb{P}([X > x]) = 1 - \mathbb{P}([X \leq x]) = 1 - F_X(x) = 1 - 0 = 1$$

× si $x \in [b, +\infty[$:

$$\mathbb{P}([X > x]) = 1 - F_X(x) = 1 - \left(1 - \left(\frac{b}{x}\right)^a\right) = \left(\frac{b}{x}\right)^a$$

Ainsi, la fonction de survie de X est :

$$x \mapsto \begin{cases} 1 & \text{si } x < b \\ \left(\frac{b}{x}\right)^a & \text{si } x \geq b \end{cases}$$

□

5. Démontrer que, pour tout réel y positif ou nul, la probabilité conditionnelle $\mathbb{P}_{[X > x]}([X > x + y])$ 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 ?

Démonstration.

Soit $y \geq 0$ et soit $x \in [b, +\infty[$.

- Comme $\mathbb{P}([X > x]) \neq 0$, on a (d'après **I.1.b**) :

$$\begin{aligned} \mathbb{P}_{[X > x]}([X > x + y]) &= \frac{\mathbb{P}([X > x + y] \cap [X > x])}{\mathbb{P}([X > x])} \\ &= \frac{\mathbb{P}([X > x + y])}{\mathbb{P}([X > x])} && \text{(car, comme } y \geq 0 : \\ & && [X > x + y] \subset [X > x]) \\ &= \frac{\left(\frac{b}{x+y}\right)^a}{\left(\frac{b}{x}\right)^a} && \text{(car } x \in [b, +\infty[\text{ et } x + y \in [b, +\infty[)} \\ &= \left(\frac{b}{x+y}\right)^a \left(\frac{x}{b}\right)^a \\ &= \frac{b^a}{(x+y)^a} \frac{x^a}{b^a} = \left(\frac{x}{x+y}\right)^a \end{aligned}$$

- Or : $\frac{x}{x+y} \underset{x \rightarrow +\infty}{\sim} \frac{x}{x} = 1$. On en déduit : $\lim_{x \rightarrow +\infty} \frac{x}{x+y} = 1$. Ainsi : $\lim_{x \rightarrow +\infty} \left(\frac{x}{x+y}\right)^a = 1$.

$$\text{D'où : } \lim_{x \rightarrow +\infty} \mathbb{P}_{[X > x]}([X > x + y]) = 1.$$

Si on considère que X modélise la durée de vie d'un phénomène, la propriété précédente signifie que plus le phénomène a duré longtemps (plus x est grand), plus la probabilité que le phénomène dure encore est grande. On parle alors de rajeunissement.

Commentaire

On cherche dans cette question à déterminer la limite de l'expression $\mathbb{P}_{[X > x]}([X > x + y])$ lorsque x tend vers $+\infty$. Il suffit donc de déterminer une expression de $\mathbb{P}_{[X > x]}([X > x + y])$ pour x assez grand. C'est pourquoi on choisit ici en début de preuve : $x \in [b, +\infty[$.

□

6. On pose dans cette question : $Y = \ln\left(\frac{X}{b}\right)$.

Démontrer que Y suit une loi exponentielle dont on précisera le paramètre.

Démonstration.

- Commençons par déterminer $Y(\Omega)$.

Notons $h : x \mapsto \ln\left(\frac{x}{b}\right)$, de telle sorte que $Y = h(X)$.

On considère ici $X(\Omega) \subset [b, +\infty[$. On en déduit :

$$\begin{aligned} Y(\Omega) &= (h(X))(\Omega) = h(X(\Omega)) \\ &= h([b, +\infty[) \\ &= [h(b), \lim_{x \rightarrow +\infty} h(x)[\quad (\text{car la fonction } h \text{ est continue et} \\ &\quad \text{strictement croissante sur } [b, +\infty[) \\ &= [0, +\infty[\end{aligned}$$

Et ainsi : $Y(\Omega) = [0, +\infty[$.

- Soit $x \in \mathbb{R}$. Deux cas se présentent :

× si $x < 0$, alors $[Y \leq x] = \emptyset$ car $Y(\Omega) = [0, +\infty[$. Donc :

$$F_Y(x) = \mathbb{P}([Y \leq x]) = \mathbb{P}(\emptyset) = 0$$

× si $x \geq 0$ alors :

$$\begin{aligned} F_Y(x) &= \mathbb{P}([Y \leq x]) \\ &= \mathbb{P}\left(\left[\ln\left(\frac{X}{b}\right) \leq x\right]\right) \\ &= \mathbb{P}\left(\left[\frac{X}{b} \leq e^x\right]\right) \quad (\text{par stricte croissance} \\ &\quad \text{de la fonction } \exp \text{ sur } \mathbb{R}) \\ &= \mathbb{P}([X \leq b e^x]) \quad (\text{car } b > 0) \\ &= F_X(b e^x) \\ &= 1 - \left(\frac{b}{b e^x}\right)^a \quad (\text{car, comme } x \geq 0, \\ &\quad b e^x \geq b) \\ &= 1 - (e^{-x})^a = 1 - e^{-ax} \end{aligned}$$

On obtient finalement : $F_Y : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-ax} & \text{si } x \geq 0 \end{cases}$.

- On reconnaît la fonction de répartition d'une v.a.r. qui suit une loi $\mathcal{E}(a)$. Or la fonction de répartition caractérise la loi.

On en déduit : $X \leftrightarrow \mathcal{E}(a)$

Commentaire

Cette question amène une remarque sur la notation $X(\Omega)$ lorsque X est une v.a.r. .

- Dans notre sujet, aucune information n'est donnée sur $X(\Omega)$. Ainsi, la v.a.r. $Y = \ln\left(\frac{X}{b}\right)$ est mal définie. Pour qu'elle le soit il faudrait savoir que la v.a.r. X est à valeurs strictement positives, *i.e.* : $X(\Omega) \subset]0, +\infty[$. Quitte à faire cette hypothèse supplémentaire (indispensable dans cette question), nous nous sommes permis de considérer : $X(\Omega) \subset [b, +\infty[$; ce qui est raisonnable au vu de la densité donnée par l'énoncé pour la loi de Pareto. On prendra cependant garde à ne pas confondre $X(\Omega)$, et l'ensemble d'annulation d'une densité de X . On détaille cette distinction ci-dessous.

- Rappelons qu'une v.a.r. X est une application $X : \Omega \rightarrow \mathbb{R}$. Comme la notation le suggère, $X(\Omega)$ est l'image de Ω par l'application X . Ainsi, $X(\Omega)$ n'est rien d'autre que l'ensemble des valeurs prises par la v.a.r. X :

$$\begin{aligned} X(\Omega) &= \{X(\omega) \mid \omega \in \Omega\} \\ &= \{x \in \mathbb{R} \mid \exists \omega \in \Omega, X(\omega) = x\} \end{aligned}$$

- Il faut bien noter que, dans la définition de $X(\Omega)$, aucune application probabilité \mathbb{P} n'apparaît. Même si cela ne correspond pas directement à la définition, il est d'usage courant de confondre, dans le cas des v.a.r. discrètes, l'ensemble de valeurs possibles de la v.a.r. X (*i.e.* l'ensemble $X(\Omega)$) et l'ensemble des valeurs que X prend avec probabilité non nulle (dans le cas où X est une v.a.r. discrète, cet ensemble est appelé support de X et est noté $\text{Supp}(X)$). □

III. Estimation des paramètres 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 ($\alpha > 0, \beta > 0$).

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes, toutes de même loi que X .

7. On suppose tout d'abord que le paramètre β fait partie des caractéristiques connues du canal de communication; on se propose de déterminer un estimateur de α par une méthode dite du « maximum de vraisemblance ». Pour cela, n désignant un entier naturel non nul et x_1, \dots, x_n des réels supérieurs ou égaux à β , on introduit la fonction \mathcal{L} , à valeurs dans \mathbb{R} et définie sur $]0, +\infty[$ par :

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

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

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

a) Exprimer $\mathcal{L}(a)$, puis $\ln(\mathcal{L}(a))$.

Démonstration.

- Soit $a \in]0, +\infty[$. Comme x_1, \dots, x_n sont supérieurs ou égaux à β :

$$\begin{aligned} \mathcal{L}(a) &= f_a(x_1) \times \dots \times f_a(x_n) \\ &= a \frac{\beta^a}{(x_1)^{a+1}} \times \dots \times a \frac{\beta^a}{(x_n)^{a+1}} \\ &= \frac{(a \beta^a)^n}{(x_1 \dots x_n)^{a+1}} \end{aligned}$$

$$\forall a \in]0, +\infty[, \mathcal{L}(a) = \frac{(a \beta^a)^n}{(x_1 \dots x_n)^{a+1}}$$

- Soit $a \in]0, +\infty[$. Comme $\mathcal{L}(a) > 0$, on peut déterminer $\ln(\mathcal{L}(a))$:

$$\begin{aligned} \ln(\mathcal{L}(a)) &= \ln\left(\frac{(a \beta^a)^n}{(x_1 \dots x_n)^{a+1}}\right) \\ &= \ln((a \beta^a)^n) - \ln((x_1 \dots x_n)^{a+1}) \\ &= n \ln(a \beta^a) - (a+1) \ln(x_1 \dots x_n) \\ &= n \ln(a) + n \ln(\beta^a) - (a+1) \sum_{k=1}^n \ln(x_k) \end{aligned}$$

$$\forall a \in]0, +\infty[, \ln(\mathcal{L}(a)) = n \ln(a) + n a \ln(\beta) - (a+1) \sum_{k=1}^n \ln(x_k)$$

□

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

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

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

Démonstration.

- La fonction φ est de classe \mathcal{C}^1 sur $]0, +\infty[$ comme somme des fonctions :

× $a \mapsto \ln(a)$ de classe \mathcal{C}^1 sur $]0, +\infty[$,

× $a \mapsto \left(n \ln(\beta) - \sum_{k=1}^n \ln(x_k)\right) a - \sum_{k=1}^n \ln(x_k)$ de classe \mathcal{C}^1 sur $]0, +\infty[$ car affine.

- Soit $a > 0$.

$$\varphi'(a) = \frac{n}{a} + n \ln(\beta) - \sum_{k=1}^n \ln(x_k) = \frac{n}{a} - \sum_{k=1}^n (\ln(x_k) - \ln(\beta)) = \frac{n}{a} - \sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)$$

- Déterminons le signe φ' .

$$\begin{aligned} \varphi'(a) \geq 0 &\Leftrightarrow \frac{n}{a} - \sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right) \geq 0 \\ &\Leftrightarrow \frac{n}{a} \geq \sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right) \\ &\Leftrightarrow \frac{a}{n} \leq \frac{1}{\sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)} \quad (\text{par stricte décroissance de la fonction inverse sur }]0, +\infty[) \\ &\Leftrightarrow a \leq \frac{n}{\sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)} \quad (\text{car } n > 0) \end{aligned}$$

- En notant $w = \frac{n}{\sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)}$, on en déduit le tableau de variations de φ .

| | | | |
|-------------------------|-----------|--------------|-----------|
| x | 0 | w | $+\infty$ |
| Signe de $\varphi'(x)$ | + | 0 | - |
| Variations de φ | $-\infty$ | $\varphi(w)$ | $-\infty$ |

On en déduit que φ admet un maximum atteint en un seul point $w = \frac{n}{\sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)} \in]0, +\infty[$.

Commentaire

- Il faudrait préciser que $(x_1, \dots, x_k) \neq (\beta, \dots, \beta)$. Si tel n'est pas le cas, $\ln\left(\frac{x_k}{\beta}\right) = 0$ et l'écriture de w précédente n'est pas valide.
- Il s'agit certainement d'un petit oubli du concepteur de sujet.
En effet, si $(x_1, \dots, x_k) = (\beta, \dots, \beta)$, alors $\varphi : a \mapsto n \ln\left(\frac{a}{\beta}\right)$ et dans ce cas, la fonction φ n'admet pas de maximum.

- (ii) Exprimer w en fonction de x_1, \dots, x_n .

Démonstration.

D'après la question précédente, $w = \frac{n}{\sum_{k=1}^n \ln\left(\frac{x_k}{\beta}\right)}$.

(iii) Que peut-on dire de w pour la fonction \mathcal{L} ?

Démonstration.

Soit $t \in]0, +\infty[$ tel que $t \neq w$.

$$\begin{aligned} \varphi(t) &< \varphi(w) && \text{(car } \varphi \text{ atteint son unique maximum en } w) \\ \Leftrightarrow \ln(\mathcal{L}(t)) &< \ln(\mathcal{L}(w)) && \text{(par définition de } \varphi) \\ \Leftrightarrow \mathcal{L}(t) &< \mathcal{L}(w) && \text{(par stricte croissance de la fonction exp sur } \mathbb{R}) \end{aligned}$$

Ainsi, w est l'unique point en lequel \mathcal{L} atteint un maximum. □

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

$$W_n = \frac{n}{\sum_{k=1}^n \ln\left(\frac{X_k}{\beta}\right)}$$

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

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

Démonstration.

- Par définition, pour tout $k \in \llbracket 1, n \rrbracket$, X_k suit la loi de Pareto de paramètres α et β .
- D'après la question **II.4**, la v.a.r. $\ln\left(\frac{X_k}{\beta}\right)$ suit la loi exponentielle de paramètre α .
- Les v.a.r. X_1, \dots, X_n sont indépendantes.
 Par lemme des coalitions, on en déduit que les v.a.r. $\ln\left(\frac{X_1}{\beta}\right), \dots, \ln\left(\frac{X_n}{\beta}\right)$ sont elles aussi indépendantes.

Ainsi, d'après la question **I.2.b**), $\sum_{k=1}^n \ln\left(\frac{X_k}{\beta}\right)$ admet pour densité f_n en prenant $\lambda = \alpha$. □

(ii) 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 un estimateur sans biais de α construit sur W_n .

Démonstration.

Soit $n \geq 2$.

- W_n s'écrit sous la forme $g(U_n)$ où :
 - × $U_n = \sum_{k=1}^n \ln\left(\frac{X_k}{\beta}\right)$ est une v.a.r. de densité f_n (en prenant $\lambda = \alpha$),
 - × $g : t \mapsto \frac{n}{t}$.

- D'après le théorème de transfert, W_n admet une espérance si et seulement si $\int_{-\infty}^{+\infty} g(t) f_n(t) dt$ est absolument convergente.

Or : $\forall t \in \mathbb{R}, g(t) f_n(t) \geq 0$. Donc, cela revient à montrer que l'intégrale $\int_{-\infty}^{+\infty} g(t) f_n(t) dt$ converge.

- La fonction f_n étant nulle en dehors de $[0, +\infty[$:

$$\int_{-\infty}^{+\infty} g(t) f_n(t) dt = \int_0^{+\infty} g(t) f_n(t) dt$$

- Or, pour $t \in]0, +\infty[$:

$$\begin{aligned} g(t) f_n(t) &= \frac{n}{t} \frac{\alpha^n}{(n-1)!} e^{-\alpha t} t^{n-1} \\ &= \frac{n \alpha \alpha^{n-1}}{(n-1)(n-2)!} e^{-\alpha t} t^{n-2} \\ &= \frac{n}{n-1} \alpha \frac{\alpha^{n-1}}{(n-2)!} e^{-\alpha t} t^{n-2} \\ &= \frac{n}{n-1} \alpha f_{n-1}(t) \end{aligned}$$

- On reconnaît une densité de probabilité, à multiplication par une constante près.

Ainsi, $\int_0^{+\infty} g(t) f_n(t) dt$ est convergente et :

$$\int_0^{+\infty} g(t) f_n(t) dt = \frac{n}{n-1} \alpha \underbrace{\int_0^{+\infty} f_{n-1}(t) dt}_1 = \frac{n}{n-1} \alpha$$

On en déduit : $\mathbb{E}(W_n) = \frac{n}{n-1} \alpha$.

- Or :

$$\mathbb{E}(W_n) = \frac{n}{n-1} \alpha \Leftrightarrow \frac{n-1}{n} \mathbb{E}(W_n) = \alpha \Leftrightarrow \mathbb{E}\left(\frac{n-1}{n} W_n\right) = \alpha \quad (\text{par linéarité de l'espérance})$$

- De plus, la v.a.r. $\frac{n-1}{n} W_n = \frac{n-1}{n} \frac{\sum_{k=1}^n X_k}{\sum_{k=1}^n \ln\left(\frac{X_k}{\beta}\right)}$ s'exprime :

× à l'aide d'un n -échantillon (X_1, \dots, X_n) de la v.a.r. X .

× sans mention du paramètre α .

$\frac{n-1}{n} W_n$ est un estimateur sans biais de α .

□

d) On pose, pour tout n de \mathbb{N}^* :

$$W'_n = \frac{n-1}{n} W_n.$$

(i) Soit n un entier supérieur ou égal à 3.

En admettant que le moment d'ordre 2 de W'_n est égal à $\frac{(n-1)\alpha^2}{n-2}$, calculer la variance de W'_n puis établir, à l'aide de l'inégalité de Bienaymé-Tchebychev, que pour tout réel ε strictement positif, on a l'inégalité :

$$\mathbb{P}([W'_n - \varepsilon < \alpha < W'_n + \varepsilon]) \geq 1 - \frac{\alpha^2}{\varepsilon^2(n-2)}.$$

Démonstration.

Soit $n \geq 3$.

- On suppose que W'_n admet un moment d'ordre 2. Par la formule de Kœnig-Huygens :

$$\begin{aligned} \mathbb{V}(W'_n) &= \mathbb{E}(W_n'^2) - (\mathbb{E}(W'_n))^2 \\ &= \frac{(n-1)\alpha^2}{n-2} - \alpha^2 \\ &= \frac{(n-1)\alpha^2 - (n-2)\alpha^2}{n-2} = \frac{\alpha^2}{n-2} \end{aligned}$$

$$\mathbb{V}(W'_n) = \frac{\alpha^2}{n-2}$$

- Soit $\varepsilon > 0$. Comme W'_n admet un moment d'ordre 2, l'inégalité de Bienaymé-Tchebychev permet d'affirmer :

$$\mathbb{P}(|W'_n - \mathbb{E}(W'_n)| \geq \varepsilon) \leq \frac{\alpha^2}{\varepsilon^2}$$

d'où $-\mathbb{P}(|W'_n - \alpha| \geq \varepsilon) \geq -\frac{\alpha^2}{(n-2)\varepsilon^2}$

et $1 - \mathbb{P}(|W'_n - \alpha| \geq \varepsilon) \geq 1 - \frac{\alpha^2}{(n-2)\varepsilon^2}$

ainsi $\mathbb{P}(|W'_n - \alpha| < \varepsilon) \geq 1 - \frac{\alpha^2}{(n-2)\varepsilon^2}$ *(probabilité de l'événement contraire)*

Enfin, pour tout $\omega \in \Omega$:

$$|W'_n(\omega) - \alpha| < \varepsilon \Leftrightarrow -\varepsilon < W'_n(\omega) - \alpha < \varepsilon \Leftrightarrow W'_n(\omega) - \varepsilon < \alpha < W'_n(\omega) - \varepsilon$$

Ainsi : $[|W'_n - \alpha| < \varepsilon] = [W'_n - \varepsilon < \alpha < W'_n - \varepsilon]$.

On en déduit : $\mathbb{P}([W'_n - \varepsilon < \alpha < W'_n + \varepsilon]) \geq 1 - \frac{\alpha^2}{\varepsilon^2(n-2)}$.

□

- (ii) On suppose dans cette question (et elle seule) que α est strictement compris entre 1 et 2. Déterminer un entier naturel N tel que, pour tout entier n supérieur ou égal à N , $\left[W'_n - \frac{1}{10}, W'_n + \frac{1}{10} \right]$ soit un intervalle de confiance du paramètre réel α au niveau de confiance 0,95.

Démonstration.

- Soit $\alpha \in \mathbb{R}$ tel que : $1 < \alpha < 2$. On considère $\varepsilon = \frac{1}{10}$. D'après la question précédente :

$$\mathbb{P} \left(\left[W'_n - \frac{1}{10} < \alpha < W'_n + \frac{1}{10} \right] \right) \geq 1 - \frac{\alpha^2}{\left(\frac{1}{10}\right)^2 (n-2)}$$

- Pour que $\left[W'_n - \frac{1}{10}, W'_n + \frac{1}{10} \right]$ soit un intervalle de confiance au niveau de confiance 0,95, il suffit de choisir n tel que :

$$\begin{aligned} 1 - \frac{\alpha^2}{\left(\frac{1}{10}\right)^2 (n-2)} &\geq 0,95 \\ \Leftrightarrow 0,05 &\geq \frac{\alpha^2}{\left(\frac{1}{10}\right)^2 (n-2)} \\ \Leftrightarrow 0,05 (n-2) &\geq 10^2 \alpha^2 \\ \Leftrightarrow n-2 &\geq \frac{10^2}{0,05} \alpha^2 \\ \Leftrightarrow n &\geq \frac{10^2}{0,05} \alpha^2 + 2 \end{aligned}$$

De plus, comme $\alpha < 2$:

$$\frac{10^2}{0,05} \alpha^2 + 2 = \frac{10^2}{\frac{5}{100}} \alpha^2 + 2 = \frac{10^4}{5} \alpha^2 + 2 < \frac{4 \times 10^4}{5} + 2 = 4 \times 2 \times 10^3 + 2 = 8002$$

Donc, par transitivité, si $n \geq 8002$, alors $n \geq \frac{10^2}{0,05} \alpha^2 + 2$.

D'où : $1 - \frac{\alpha^2}{\left(\frac{1}{10}\right)^2 (n-2)} \geq 0,95$

Pour tout $n \geq 8002$, $\left[W'_n - \frac{1}{10}, W'_n + \frac{1}{10} \right]$ est un intervalle de confiance au niveau de confiance 0,95. □

8. On suppose maintenant que seul le paramètre α est déjà identifié et qu'il vérifie : $\alpha > 2$.

a) Pour tout entier strictement positif n , on pose :

$$Y_n = c_n \sum_{k=1}^n X_k,$$

où le réel c_n est choisi de sorte que $(Y_n)_{n \geq 1}$ soit un estimateur sans biais de β .

(i) Calculer c_n .

Démonstration.

La v.a.r. Y_n admet une espérance en tant que combinaison linéaire de v.a.r. qui admettent une espérance. De plus, par linéarité de l'espérance :

$$\mathbb{E}(Y_n) = \mathbb{E}\left(c_n \sum_{k=1}^n X_k\right) = c_n \sum_{k=1}^n \mathbb{E}(X_k) = c_n \sum_{k=1}^n \mathbb{E}(X_1) = c_n n \mathbb{E}(X_1)$$

Enfin, d'après la question **II.1**, $\mathbb{E}(X_1) = \frac{\alpha\beta}{\alpha-1}$ et ainsi : $\mathbb{E}(Y_n) = c_n n \frac{\alpha}{\alpha-1} \beta$.

En choisissant $c_n = \frac{\alpha-1}{\alpha n}$, on obtient $\mathbb{E}(Y_n) = \beta$.

□

(ii) Quelle est la limite de la variance de Y_n quand n tend vers $+\infty$?
 (On dit que l'estimateur est convergent.)

Démonstration.

La v.a.r. Y_n admet une variance comme somme de v.a.r. indépendantes qui admettent des variances. De plus :

$$\begin{aligned} \mathbb{V}(Y_n) &= \mathbb{V}\left(c_n \sum_{k=1}^n X_k\right) = c_n^2 \mathbb{V}\left(\sum_{k=1}^n X_k\right) && \text{(par propriété de la variance)} \\ &= c_n^2 \sum_{k=1}^n \mathbb{V}(X_k) && \text{(par indépendance de } X_1, \dots, X_n) \\ &= c_n^2 \sum_{k=1}^n \mathbb{V}(X) = c_n^2 n \mathbb{V}(X) \\ &= \frac{(\alpha-1)^2}{\alpha^2 n^2} n \frac{\alpha \beta^2}{(\alpha-1)^2 (\alpha-2)} \\ &= \frac{\beta^2}{\alpha(\alpha-2)} \frac{1}{n} \\ &\xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

On en déduit que : $\lim_{n \rightarrow +\infty} \mathbb{V}(Y_n) = 0$.

□

b) Pour tout entier strictement positif n , on pose : $Z_n = \min(X_1, \dots, X_n)$.

(i) Déterminer la fonction de répartition de Z_n , puis reconnaître sa loi et préciser son espérance. Quelle est la limite de cette dernière quand n tend vers $+\infty$?

Démonstration.

- On a déjà considéré, pour tout $k \in \llbracket 1, n \rrbracket$: $X_k(\Omega) \subset [\beta, +\infty[$.
 On en déduit : $Z_n(\Omega) \subset [\beta, +\infty[$.
- Soit $x \in \mathbb{R}$. Deux cas se présentent :

× si $x < \beta$, alors $[Z_n \leq x] = \emptyset$ car $Z_n(\Omega) \subset [\beta, +\infty[$. Donc :

$$F_{Z_n}(x) = \mathbb{P}([Z_n \leq x]) = \mathbb{P}(\emptyset) = 0$$

× si $x \geq \beta$ alors :

$$\begin{aligned}
 F_{Z_n}(x) &= \mathbb{P}([Z_n \leq x]) \\
 &= \mathbb{P}([\min(X_1, \dots, X_n) \leq x]) \\
 &= 1 - \mathbb{P}([\min(X_1, \dots, X_n) > x]) \\
 &= 1 - \mathbb{P}([X_1 > x] \cap \dots \cap [X_n > x]) \\
 &= 1 - \mathbb{P}([X_1 > x]) \times \dots \times \mathbb{P}([X_n > x]) \quad (\text{par indépendance de } X_1, \dots, X_n) \\
 &= 1 - \mathbb{P}([X_1 > x]) \times \dots \times \mathbb{P}([X_1 > x]) \quad (\text{car } X_1, \dots, X_n \text{ ont} \\
 &\quad \text{toutes même loi que } X) \\
 &= 1 - (\mathbb{P}([X_1 > x]))^n \\
 &= 1 - \left(\left(\frac{\beta}{x}\right)^\alpha\right)^n = 1 - \left(\frac{\beta}{x}\right)^{\alpha n} \quad (\text{d'après II.2. et car } x \geq \beta)
 \end{aligned}$$

On reconnaît (d'après **II.2.**) la fonction de répartition d'une v.a.r. qui suit la loi de Pareto de paramètres αn et β .

Ainsi, Z_n admet une espérance et $\mathbb{E}(Z_n) = \frac{\alpha n \beta}{\alpha n - 1}$ (d'après **II.1.**)

• Enfin, comme $\alpha > 0$ et $\beta > 0$:

$$\mathbb{E}(Z_n) = \frac{\alpha n \beta}{\alpha n - 1} \underset{n \rightarrow +\infty}{\sim} \frac{\cancel{\alpha n} \beta}{\cancel{\alpha n}} = \beta \underset{n \rightarrow +\infty}{\rightarrow} \beta$$

En conclusion : $\mathbb{E}(Z_n) \underset{n \rightarrow +\infty}{\rightarrow} \beta$.
 L'estimateur Z_n est donc asymptotiquement sans biais. □

(ii) Pour tout entier strictement positif n , on pose : $Z'_n = d_n Z_n$, où le réel d_n est choisi de telle sorte que $(Z'_n)_{n \geq 1}$ soit un estimateur sans biais de β .
 Quelle est la limite de la variance de Z'_n quand n tend vers $+\infty$?

Démonstration.

• D'après la question précédente : $\mathbb{E}(Z_n) = \frac{\alpha n \beta}{\alpha n - 1}$. Or :

$$\mathbb{E}(Z_n) = \frac{\alpha n \beta}{\alpha n - 1} \Leftrightarrow \frac{\alpha n - 1}{\alpha n} \mathbb{E}(Z_n) = \beta \Leftrightarrow \mathbb{E}\left(\frac{\alpha n - 1}{\alpha n} Z_n\right) = \beta \quad (\text{par linéarité de l'espérance})$$

Donc $Z'_n = \frac{\alpha n - 1}{\alpha n} Z_n$ est un estimateur sans biais de β ($d_n = \frac{\alpha n - 1}{\alpha n}$).

• La v.a.r. $Z'_n = d_n Z_n$ admet une variance en tant que transformée affine de Z_n , et :

$$\begin{aligned}
 \mathbb{V}(Z'_n) &= \mathbb{V}(d_n Z_n) = d_n^2 \mathbb{V}(Z_n) = \frac{(\cancel{\alpha n - 1})^2}{(\alpha n)^2} \frac{\cancel{\alpha n} \beta^2}{(\alpha n - 2) (\cancel{\alpha n - 1})^2} \\
 &= \frac{1}{\alpha n (\alpha n - 2)} \beta^2 \underset{n \rightarrow +\infty}{\rightarrow} 0
 \end{aligned}$$

$\mathbb{V}(Z'_n) \underset{n \rightarrow +\infty}{\rightarrow} 0$ □

- (iii) Démontrer que l'estimateur $(Z'_n)_{n \geq 1}$ est plus efficace que l'estimateur $(Y_n)_{n \geq 1}$, c'est-à-dire, qu'à partir d'un certain rang, la variance de Z'_n est inférieure à celle de Y_n .

Démonstration.

Raisonnons par équivalence.

$$\begin{aligned} \mathbb{V}(Z'_n) &\leq \mathbb{V}(Y_n) \\ \Leftrightarrow \frac{\beta^2}{n\alpha(n\alpha-2)} &\leq \frac{\beta^2}{\alpha(\alpha-2)} \frac{1}{n} \\ \Leftrightarrow \frac{1}{n\alpha-2} &\leq \frac{1}{\alpha-2} \quad (\text{car } \frac{\beta^2}{n\alpha} > 0) \\ \Leftrightarrow n\alpha-2 &\geq \alpha-2 \quad (\text{par stricte décroissance de la} \\ &\quad \text{fonction inverse sur }]0, +\infty[) \\ \Leftrightarrow n\alpha &\geq \alpha \\ \Leftrightarrow n &\geq 1 \quad (\text{car } \alpha > 0) \end{aligned}$$

À partir du rang $n = 1$, $\mathbb{V}(Z'_n) \leq \mathbb{V}(Y_n)$ et l'estimateur Z'_n est donc plus efficace que l'estimateur Y_n . □

Commentaire

[Prenons un peu de recul ...]

- La loi présentée dans la question **II.2.b)** est appelée loi d'Erlang (hors programme) de paramètres n et λ . Cette loi est liée à des lois de probabilité classiques :
 - × lorsque $n = 1$ (cf initialisation de la récurrence du **II.2.b)**), on reconnaît la loi exponentielle de paramètre λ .
 - × de manière générale, la loi d'Erlang est un cas spécial de la loi Gamma (qui admet deux paramètres notés généralement α et β) dont une densité est donnée par :

$$t \mapsto \begin{cases} 0 & \text{si } t < 0; \\ \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta t} t^{\alpha-1} & \text{si } t \geq 0. \end{cases}$$

En prenant $\alpha = n$, on reconnaît la densité f_n de la question de l'énoncé.

- Cette loi Gamma se définit à l'aide de la fonction :

$$\Gamma : x \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt$$

que l'on peut rencontrer par exemple dans ESSEC II 2005.

Pour rappel, la fonction Γ est définie sur \mathbb{R}^{+*} et on peut démontrer (penser à une IPP) :

$$\begin{cases} \Gamma(1) = 1 \\ \forall x > 0, \Gamma(x+1) = x\Gamma(x) \end{cases}$$

de sorte que : $\forall n \in \mathbb{N}, \Gamma(n+1) = n!$

(on peut voir la fonction Γ comme un prolongement de la fonction factorielle)

- La loi de Pareto est elle aussi très classique aux concours. Elle est elle aussi reliée à la loi exponentielle (c'était l'objet de la question **II.4.**).

Commentaire

Enfin, la partie estimation portait sur l'estimateur du maximum de vraisemblance.

Détaillons ce point.

- On dispose d'un échantillon (x_1, \dots, x_n) d'observations.
- On suppose que ces observations proviennent d'un n -échantillon (X_1, \dots, X_n) d'une v.a.r. X dont la loi dépend d'un paramètre α , a priori inconnu et qu'on cherche à déterminer. Pour ce faire, une idée naturelle consiste à considérer que la valeur de α qui a permis de générer les observations est celle qui avait la plus grande probabilité de les générer. C'est cette idée qui guide la méthode dite du maximum de vraisemblance.
- Le réel w est précisément la valeur du paramètre α maximisant la réalisation des observations initiales.
- La méthode du maximum de vraisemblance conduit à considérer la variable aléatoire construite à l'aide de ce maximum : W_n .
- Plaçons-nous dans le cas où X est une v.a.r. discrète (la méthode est plus simple à appréhender dans ce cas). L'idée est de choisir comme estimation de α le réel w tel que la **vraisemblance** d'avoir obtenu l'échantillon utilisé soit maximisée.

Autrement dit, le réel w tel que la probabilité :

$$\begin{aligned} \mathcal{L}(a) &= \mathbb{P}_a([X_1 = x_1] \cap \dots \cap [X_n = x_n]) \\ &= \mathbb{P}_a([X_1 = x_1]) \times \dots \times \mathbb{P}([X_n = x_n]) \quad (\text{par indépendance de } X_1, \dots, X_n) \\ &= \prod_{i=1}^n \mathbb{P}_a([X_i = x_i]) \end{aligned}$$

soit maximale. L'énoncé portait sur le cas de v.a.r. à densité, que l'on comprend aisément par analogie avec le cas des v.a.r. discrètes.