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 $\mathscr{B}$ 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 :

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

et que les vecteurs e'_1, e'_2, e'_3 aient respectivement pour troisième composante 1, -1 et 2.

On notera dorénavant $\mathscr{B}'$ la base (e'_1, e'_2, e'_3).

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

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

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

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

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.

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

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

Problème : probabilités

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

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).

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 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 :

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

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 :

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

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

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

3. 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 ?

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

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.

1. 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 $\mathscr{L}$, à valeurs dans $\mathbb{R}$ et définie sur $]0,+\infty[$ par :

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

où f_a est la fonction définie sur $\mathbb{R}$ par : $f_a(t)=\begin{cases} 0 \text{ si } t<\beta, \\ a\dfrac{\beta^a}{t^{a+1}} \text{ si } t\geq \beta. \end{cases}$

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

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

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

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

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

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

c) On pose dorénavant, pour tout n de $\mathbb{N}^*$: $W_n = \dfrac{n}{\displaystyle\sum_{k=1}^{n} \ln \frac{X_k}{\beta}}$.

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

i) 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 $\lambda = \alpha$.

ii) À l'aide du théorème de transfert, en déduire que W_n admet pour espérance $\dfrac{n\alpha}{n-1}$ lorsque $n \geq 2$, puis proposer un estimateur sans biais de α construit sur W_n.

d) On pose, pour tout n de $\mathbb{N}^*$: $W'_n = \dfrac{n-1}{n} W_n$.

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

En admettant que le moment d'ordre 2 de W'_n est égal à $\dfrac{(n-1)\alpha^2}{n-2}$, 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 : $\mathbb{P}(W'_n - \varepsilon < \alpha < W'_n + \varepsilon) \geq 1 - \dfrac{\alpha^2}{\varepsilon^2 (n-2)}$.

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, $\left[W'_n - \dfrac{1}{10}, W'_n + \dfrac{1}{10} \right]$ soit un intervalle de confiance du paramètre α au niveau de confiance 0,95.

2. 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 : $Y_n = c_n \displaystyle\sum_{k=1}^{n} X_k$, 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.

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

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 $+\infty$?

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 $+\infty$?

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.

*****