Devoir surveillé n° 5

Durée de l'épreuve : 3h

Le devoir comporte un exercice et un problème.
La calculatrice n'est pas autorisée. Il sera tenu compte de la rigueur, du soin et de la rédaction dans la notation. Les résultats seront soulignés ou encadrés ***à l'aide d'une règle****.*

On suppose, pour toutes les questions en langage Python, que les bibliothèques habituelles sont déjà importées sous leurs raccourcis habituels.

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

Exercice

Dans tout l'exercice, on considère un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

Soit f la fonction définie sur $\mathbb{R}$ par : $\forall x \in \mathbb{R},\ f(x) = \dfrac{1}{\pi(1+x^2)}$.

1. Montrer que f peut être considérée comme une densité de probabilité.
2. Déterminer la nature de $\displaystyle\int_1^{+\infty} \frac{1}{x}\mathrm{d}x$ et de $\displaystyle\int_1^{+\infty} \frac{1}{x^2}\mathrm{d}x$ et donner leur valeur en cas de convergence.

On note X une variable aléatoire de densité f. On dit que X suit la loi de Cauchy. On note F la fonction de répartition de X.

3. Montrer que X n'admet ni espérance, ni variance.
4. (a) Donner, pour tout $x \in \mathbb{R}$, l'expression de $F(x)$.
   (b) Montrer que F réalise une bijection de $\mathbb{R}$ sur $]0;1[$ et préciser, pour tout $y \in]0;1[$, l'expression de $F^{-1}(y)$.
5. (a) Soit U une variable aléatoire de loi uniforme sur $]0;1[$. Montrer que $Y = F^{-1}(U)$ suit la même loi que X.
   (b) **Informatique.** Déduire de la question précédente l'écriture d'une fonction Python d'en-tête `def cauchy() :` qui renvoie une simulation de X.

On note maintenant $Z = \sqrt{|X|}$. On **admet** que Z est une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbb{P})$.

6. Montrer que Z est une variable aléatoire à densité et expliciter une densité f_Z de Z.
7. Justifier que Z admet une espérance, mais pas de variance.
8. Le but de cette question est de calculer explicitement $E(Z)$.
   (a) Déterminer deux réels α et β tels que
   $$\forall x \geqslant 0, \qquad \frac{x^2}{(x^2-\sqrt{2}\,x+1)(x^2+\sqrt{2}\,x+1)} = \frac{\alpha x}{x^2-\sqrt{2}\,x+1} + \frac{\beta x}{x^2+\sqrt{2}\,x+1}.$$
   (b) Justifier que les intégrales $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^2-\sqrt{2}\,x+1}$ et $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^2+\sqrt{2}\,x+1}$ convergent.
   À l'aide du changement de variable $u = \sqrt{2}x+1$ montrer que :
   $$\int_0^{+\infty} \frac{\mathrm{d}x}{x^2+\sqrt{2}\,x+1} = \frac{\pi}{2\sqrt{2}}.$$
   On **admet** qu'on peut obtenir de la même manière $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^2-\sqrt{2}\,x+1} = \frac{3\pi}{2\sqrt{2}}$.
   (c) En observant que, pour tout $x \geqslant 0$,
   $$\frac{\alpha x}{x^2-\sqrt{2}\,x+1} + \frac{\beta x}{x^2+\sqrt{2}\,x+1} = \frac{\alpha}{2}\left(\frac{2x-\sqrt{2}}{x^2-\sqrt{2}\,x+1} + \frac{\sqrt{2}}{x^2-\sqrt{2}\,x+1}\right) + \frac{\beta}{2}\left(\frac{2x+\sqrt{2}}{x^2+\sqrt{2}\,x+1} - \frac{\sqrt{2}}{x^2+\sqrt{2}\,x+1}\right),$$
   en déduire que $E(Z) = \sqrt{2}$.

(d) **Informatique.** On suppose écrite correctement la fonction de la question 4.b. On dispose du programme ci-dessous dont l'exécution produit, après un certain temps, l'affichage ci-après.

```
def mystere(eps, n):
    L=np.zeros(1000)
    for k in range(1000):
        ech=np.zeros(n)
        for i in range(n):
            ech[i]=np.sqrt(np.abs(cauchy()))
        if np.abs(np.sum(ech)/n - np.sqrt(2)) <= eps :
            L[k]=1
    return np.sum(L)/1000

M=np.zeros([4,7])
eps=np.array([1,0.5,0.1,0.05])
n=np.array([100, 500, 1000, 1500, 3000, 5000, 50000])
for i in range(4):
    for j in range(7):
        M[i,j]=mystere(eps[i],n[j])

print(M)
```

Affichage Python

```
> > >
[[0.99 1. 0.999 1. 1. 1. 1. ]
[0.964 0.995 0.995 0.998 0.999 0.999 1. ]
[0.411 0.725 0.846 0.903 0.965 0.988 0.997]
[0.21 0.428 0.548 0.599 0.742 0.834 0.992]]
```

i. La fonction `mystere` renvoie une valeur approchée de la probabilité d'un certain événement. De quelle probabilité s'agit-il ?
Indication : on pourra considérer n VAR $Z_1, \ldots, Z_n$ indépendantes et suivant toutes la même loi que Z.

ii. À l'aide de l'affichage, quel résultat peut-on conjecturer ?

Problème

Ce problème présente deux modélisations mathématiques, l'une discrète et l'autre continue, du phénomène de sélection en biologie. La première partie consiste à démontrer des résultats qui seront utilisés dans la suite du sujet. La partie II est consacrée à l'étude probabiliste du devenir d'une stratégie évolutive en temps discret. Dans la partie III, on étudie une modélisation continue utilisant des équations différentielles. Les deux dernières parties sont indépendantes entre elles. On pourra utiliser les résultats de la première partie dans les suivantes. Plus généralement, on pourra admettre le résultat d'une question pour traiter les suivantes.

Dans tout le problème, λ est un réel strictement positif et, pour tout réel x, on note

$$S(x) = \sum_{k=0}^{+\infty} \frac{\lambda^k x^k}{k!} \mathrm{e}^{-\lambda}.$$

Partie I : Quelques résultats utiles

1. Justifier l'existence de $S(x)$ pour tout réel x, et en donner une expression simple.

Étude du nombre de points fixes d'une fonction

On rappelle que λ est un réel strictement positif ; on considère la fonction d'une variable réelle $f : x \mapsto \mathrm{e}^{\lambda(x-1)}$ et on s'intéresse aux solutions de l'équation $f(x) = x$ sur $[0\,;\,1]$.

2. Déterminer le signe sur $\mathbb{R}^+$de la fonction $g : x \mapsto x\mathrm{e}^{-x} - 1$.
3. Montrer que, si $\lambda \leqslant 1$, alors l'équation $f(x) = x$ admet une unique solution sur $[0\,;\,1]$.
Indication : on pourra dériver deux fois la fonction $\phi : x \mapsto f(x) - x$.
4. Montrer que, si $\lambda > 1$, alors l'équation $f(x) = x$ a exactement deux solutions sur $[0\,;\,1]$.
Indication : on pourra prouver que la dérivée de la fonction $\phi : x \mapsto f(x) - x$ s'annule en un seul point α sur $[0\,;\,]1$ dont on ne cherchera pas l'expression.

Dans la suite du sujet, on notera x_λ la plus petite de ces deux solutions.

Autour de la loi de Poisson

Soit $(t_n)_{n\in\mathbb{N}}$ une suite de réels strictement positifs et $(T_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes telle que, pour tout entier n, T_n suive une loi de Poisson de paramètre t_n.

5. Donner sans justifications l'espérance et la variance de T_1.
6. (a) On note $S = T_1 + T_2$, compléter l'égalité suivante :

$$\forall i \in \mathbb{N}, \qquad [S = i] = \bigcup_{j=\dots}^{\dots} [T_1 = j] \cap [T_2 = \dots].$$

   (b) En déduire une expression de $\mathbb{P}(T_1 + T_2 = i)$.
   (c) Quelle est la loi suivie par $T_1 + T_2$?
7. Montrer par récurrence que, pour tout entier n non nul, la variable aléatoire $S_n = \sum_{k=1}^{n} T_k$ suit une loi de Poisson de paramètre $\sum_{k=1}^{n} t_k$.

Résultats sur les équations différentielles

Soit a une fonction continue sur un intervalle I de $\mathbb{R}$. On considère l'équation différentielle suivante :

$$(E_1) : y'(t) = a(t)y(t)$$

8. Donner, sans justification, les solutions de l'équation différentielle (E_1).
9. En déduire que, si f est une solution de (E_1) s'annulant sur I, alors f est la fonction nulle sur l'intervalle I.

Soit b une fonction continue sur $\mathbb{R}, C$ une constante réelle et g une solution sur un intervalle I de l'équation différentielle :

$$(E_2) : y'(t) = b(y(t))[y(t) - C]$$

10. Montrer que $g - C$ est solution sur I de :

$$(E_3) : y'(t) = b(g(t))y(t)$$

11. En déduire que, s'il existe $t_0 \in I$ tel que $g(t_0) = C$, alors g est constante sur I.

Partie II : Probabilité d'extinction d'une stratégie

Dans cette partie, on s'intéresse à l'évolution en temps discret de la population qui suit une certaine stratégie de reproduction. On se donne un réel $\lambda > 0$ et on introduit une suite de variables aléatoires réelles $(Z_n)_{n\in\mathbb{N}}$ qui représente le nombre d'individus qui suivent la stratégie étudiée à chaque génération. On construit la suite $(Z_n)_{n\in\mathbb{N}}$ récursivement de la façon suivante :

— Z_0 est la variable aléatoire constante à 1 (initialement, un seul individu suit la nouvelle stratégie).
— Z_1 suit une loi de Poisson de paramètre λ.
— Si l'on note z_1 la réalisation de Z_1, alors Z_2 est la somme de z_1 variables aléatoires suivant une loi de Poisson de paramètre λ indépendantes entre elles et indépendantes de Z_1. Ces variables aléatoires représentent les descendants de la génération précédente.
— Plus généralement, pour tout entier naturel n, si z_n est la réalisation de Z_n, alors Z_{n+1} est la somme de z_n variables aléatoires suivant une loi de Poisson de paramètre λ indépendantes entre elles et indépendantes de Z_n, on notera

$$Z_{n+1} = \sum_{k=1}^{z_n} X_{n,k}$$

avec $(X_{n,k})_{k\in\mathbb{N}}$ une suite de variables aléatoires suivant une loi de Poisson de paramètre λ indépendantes entre elles et indépendantes de Z_n. De nouveau, ces variables aléatoires représentent les descendants de la génération précédente.

On rappelle que, par convention, toute somme vide est nulle. Par exemple, $\sum_{k=1}^{0} u_k$ est nulle.

Pour tout entier n, on introduit l'évènement $A_n = (Z_n = 0)$ et on note $p_n = \mathbb{P}(A_n)$.

12. Exprimer p_0, p_1 et p_2 en fonction de λ.
Indication : on pourra utiliser le système complet d'événements $([Z_1 = k])_{k\in\mathbb{N}}$ *pour le calcul de* p_2.
13. Montrer que la suite $(p_n)_{n\in\mathbb{N}}$ est monotone et en déduire qu'elle converge.

Pour tout couple d'entiers naturels (k, n), on admet que, sachant que $(Z_1 = k)$ est réalisé, Z_{n+1} est la somme de k variables aléatoires indépendantes qui suivent la même loi que Z_n. Ainsi, $\mathbb{P}(Z_{n+1} = 0 \mid Z_1 = k) = p_n^k$. (*On rappelle que* $\mathbb{P}(A|B)$ *désigne la probabilité de* A *sachant* B)

14. En déduire que, pour tout entier n, on a $p_{n+1} = S(p_n)$.
15. En utilisant les résultats et les notations de la première partie, prouver que :
   — si $\lambda \leqslant 1$, alors $\lim_{n\to+\infty} p_n = 1$;
   — si $\lambda > 1$, alors $\lim_{n\to+\infty} p_n = x_\lambda$.
16. Donner une interprétation du résultat obtenu.

Partie III : Étude de deux exemples

On s'intéresse ici à l'évolution d'une population dont les individus peuvent suivre différentes stratégies mais en utilisant une modélisation en temps continu. Sa compréhension n'est pas nécessaire.

On se donne une matrice $M \in \mathscr{M}_n(\mathbb{R})$ dont on note $m_{i,j}$ le terme situé à la i-ième ligne et la j-ième colonne. On cherche à déterminer n fonctions dérivables sur $\mathbb{R}^+$, notées $x_1, \dots, x_n$ telles que, pour tout entier i entre 1 et n, on ait :

$$\forall t \in \mathbb{R}^+, \quad x_i'(t) = x_i(t)\left[\sum_{j=1}^{n} m_{i,j} x_j(t) - \sum_{j=1}^{n}\sum_{k=1}^{n} m_{j,k} x_j(t) x_k(t).\right] \tag{1}$$

On étudie d'abord la dynamique du système (1) lorsque $M = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$.

Ce choix conduit à s'intéresser aux triplets de fonctions $(\ x_1, x_2, x_3\ )$ qui sont solutions sur $\mathbb{R}^+$ du système suivant :

$$\left\{\begin{array}{l} x_1' = x_1(x_2 - x_3) \\ x_2' = x_2(x_3 - x_1) \\ x_3' = x_3(x_1 - x_2) \\ 1 = x_1(0) + x_2(0) + x_3(0). \end{array}\right. \tag{2}$$

17. Déterminer le (ou les) triplets de réels $(\ C_1, C_2, C_3\ )$ tel(s) qu'il existe une solution de (2) constante à $(\ C_1, C_2, C_3\ )$.

On s'intéresse à une petite perturbation autour d'une solution constante identifiée à la question précédente, ce qui conduit à étudier les triplets de fonctions $(\ h_1, h_2, h_3\ )$ vérifiant le système suivant :

$$\left\{\begin{array}{l} h_1' = \dfrac{1}{3}(h_2 - h_3) \\ h_2' = \dfrac{1}{3}(h_3 - h_1) \\ h_3' = \dfrac{1}{3}(h_1 - h_2) \\ 1 = h_1(0) + h_2(0) + h_3(0). \end{array}\right. \tag{3}$$

On admet que la diagonalisation de la matrice M permet sa résolution.

18. Déterminer les valeurs propres complexes de M et les espaces propres associés.
19. Justifier que M est diagonalisable et déterminer une matrice diagonale $D \in \mathscr{M}_3(\mathbb{C})$ et une matrice $P \in GL_3(\mathbb{C})$ (c'est-à-dire que P est une matrice inversible de $\mathscr{M}_3(\mathbb{C})$) telles que :

$$M = PDP^{-1}$$

On fera en sorte que la première ligne de P ne soit constituée que de 1 .

On étudie maintenant la dynamique du système (1) lorsque $M = \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}$, c'est-à-dire que l'on considère que les individus ne peuvent adopter que deux stratégies. Dans ce cas, on peut montrer que la proportion x_1 d'individus suivant la stratégie 1 est solution de l'équation différentielle suivante :

$$y' = y(y-1)(4y-3). \tag{4}$$

20. Déterminer trois nombres réels A, B et C tels que :

$$\forall x \in \mathbb{R}\backslash\{0, 3/4, 1\}, \quad \frac{1}{x(x-1)(4x-3)} = \frac{A}{x} + \frac{B}{x-1} + \frac{C}{4x-3}.$$

21. En déduire l'ensemble des primitives de la fonction $h : x \mapsto \dfrac{1}{x(x-1)(4x-3)}$ sur l'intervalle $]0, 3/4[$.
22. Prouver que si $x_1(0) \in]\,0, 3/4\,[$, alors pour tout $t \in \mathbb{R}^+, x_1(t) \in]\,0, 3/4[$.
Indication : on pourra utiliser les résultats de la partie I.

On suppose dans la suite que $x_1(0) \in]0\,;\,3/4[$.

23. Montrer qu'il existe une constante $D \in \mathbb{R}$ telle que, pour tout $t \in \mathbb{R}^+$, on ait :

$$t + D = \frac{1}{3}\ln\left(\frac{x_1(t)\,[1 - x_1(t)]^3}{[x_1(t) - 3/4]^4}\right).$$

24. En déduire que $\lim\limits_{t \to +\infty} x_1(t) = 3/4$.

Interprétation du système (1)

Dans cette équation, la fonction x_i s'interprète comme la proportion des individus d'une population qui adoptent la stratégie i et m_{ij} représente l'avantage sélectif qu'un individu qui adopte la stratégie i retire d'une interaction avec un autre individu qui adopte la stratégie j. Ainsi, le taux de croissance de la proportion x_i à un temps t s'écrit comme la différence entre l'avantage sélectif tiré de son interaction avec les autres $\sum_{j=1}^{n} m_{i,j}x_j$ et l'avantage moyen des individus de la population représenté par le terme $\sum_{j=1}^{n}\sum_{k=1}^{n} m_{j,k}x_jx_k$.