ENS CACHAN S. Nat.

DURÉE : 4 heures

Le but de ce problème est l'étude d'un processus de croissance démographique et plus particulièrement de la probabilité d'extinction d'une lignée d'individus. Dans une première partie on établira certains résultats d'analyse qui nous seront utiles pour la suite. Dans la seconde partie nous étudierons le problème annoncé. Enfin dans la troisième partie nous étudierons plus particulièrement quelques exemples.

PARTIE I

Dans toute cette partie, on considère une suite $(p_k)_{k \in \mathbb{N}}$ de nombres réels positifs ou nuls vérifiant :

(1) $$\sum_{k=0}^{+\infty} p_k = 1 .$$

On désigne par R le rayon de convergence de la série entière $\sum_{k=0}^{+\infty} p_k x^k$ et par f la fonction somme de cette série définie sur $[0, R[$, éventuellement sur $[0, R]$, par :

$$f(x) = \sum_{k=0}^{+\infty} p_k x^k .$$

I.1. Étude de R.

a. Montrer que $R \geqslant 1$ et que, pour tout x appartenant à $[0, 1]$, on a : $0 \leqslant f(x) \leqslant 1$.

b. On pose $s = \sum_{k=1}^{+\infty} \frac{1}{k^2}$. En utilisant la série entière $\sum_{k=1}^{+\infty} \frac{x^k}{k^2}$, donner un exemple de situation où $R = 1$.

Dans toute la suite du problème, on supposera de plus que :

(2) $$p_0 + p_1 < 1 \quad \text{et} \quad p_0 > 0$$

(3) $$R > 1$$

et on étudie la suite (f_n) de fonctions telle que $f_1 = f$ et, pour tout $n \geqslant 2$, $f_n = f \circ f_{n-1}$.

I.2. Étude de $h(x) = f(x) - x$.

Pour x, élément de $[0, R[$, on pose :

$$h(x) = f(x) - x.$$

a. Montrer que h est dérivable dans $[0, R[$ et que h', sa dérivée, est strictement croissante dans cet intervalle.

b. Déduire de la question précédente que l'équation $f(x) = x$ a au plus deux racines dans $[0, R[$.

c. Discuter le nombre de racines de $f(x) = x$ dans $[0, 1]$ suivant le signe du nombre $f'(1) - 1$.

d. Montrer qu'il existe un nombre réel a vérifiant $1 < a \leqslant R$ tel que, pour tout x élément de $[0, a[$, on a $f(x) < R$.

Tournez la page S.V.P.

ENS CACHAN S. Nat

I.3. **Étude de la définition de la suite (f_n).**

a. En utilisant la question précédente I.2.*d*, prouver que f_2 est définie sur un intervalle $[0.\ a_2[$ contenant $[0, 1]$; prouver également que f_n est définie sur un intervalle $[0.\ a_n[$ contenant $[0, 1]$.

Dans les questions *b*, *c*, *d* qui suivent, on étudie l'exemple où $p_k = \left(\frac{1}{2}\right)^{k+1}$, pour tout entier $k \geqslant 0$.

b. Déterminer R et calculer $f(x)$ pour x dans $[0, R[$.

c. Déterminer l'intervalle $[0, a_2[$, ensemble des réels positifs ou nuls tels que $\sum_{k=0}^{\infty} p_k (f(x))^k$ soit convergente et déterminer la somme de cette série.

d. En procédant par récurrence, déterminer l'intervalle ouvert $[0, a_n[$ de longueur maximum sur lequel la fonction f_n est définie. En déduire que toutes les fonctions f_n ne peuvent être définies simultanément sur un intervalle ouvert contenant $[0, 1]$.

I.4. **Étude de la suite $(f_n(0))$.**

a. Montrer que cette suite $(f_n(0))$ est convergente et que sa limite ℓ vérifie

$$0 < \ell \leqslant 1 \quad \text{et} \quad f(\ell) = \ell.$$

b. Préciser ℓ dans l'ensemble des racines de $f(x) = x$ mises en évidence dans I.2.*c*.

PARTIE II

On désire dans cette partie étudier le comportement à l'infini d'un processus démographique. La loi de fécondité est définie par une variable aléatoire X prenant des valeurs entières positives ou nulles, la valeur k étant prise avec la probabilité p_k, où la suite $(p_k)_{k \geqslant 0}$ est supposée vérifier les conditions (1) (2) et (3) de la partie I.

À l'instant initial il existe un seul individu de l'espèce. À la première génération, le nombre d'individus est donné par la loi de X. À leur tour les individus de la première génération se multiplient de façon indépendante et toujours suivant la même loi de fécondité et ainsi de suite.

On note Z_n la variable aléatoire égale au nombre d'individus de la n-ième génération. L'étude de ces variables sera facilitée par l'introduction de la variable U_r (où r est un entier > 0) égale à la somme de r copies indépendantes de la variable X. Si $r = 0$ on convient que $U_0 = 0$.

II.1. **Étude des fonctions génératrices g_r de U_r.**

On pose dans le cas où la série converge :

$$g_r(x) = \sum_{k=0}^{+\infty} P(U_r = k)\, x^k.$$

a. Calculer $P(U_2 = k)$ pour tout entier k et, en comparant avec les coefficients de la série produit de $\sum_{k=0}^{+\infty} p_k x^k$ par elle-même, montrer que :

$$\forall x \in [0, R[,\ g_2(x) = (f(x))^2.$$

b. Montrer, de façon plus générale, que :

$$\forall r > 0, \forall x \in [0, R[,\ g_r(x) = (f(x))^r.$$

ENS CACHAN S. Nat

II.2. **Étude des fonctions génératrices des Z_n.**

Les variables Z_0 et Z_1 vérifient par hypothèse $Z_0 = 1$ et $P(Z_1 = k) = P(U_1 = k)$ pour tout $k \geqslant 0$, et on suppose que, pour tout $n \geqslant 1$, le conditionnement de la variable Z_{n+1} par la variable Z_n vérifie pour tout couple (k, r) d'entiers :

$$P(Z_{n+1} = k / Z_n = r) = P(U_r = k).$$

1. Montrer que, pour tout $n \geqslant 1$ et pour tout $k \geqslant 0$,

$$P(Z_n = k) = \sum_{r=0}^{+\infty} P(Z_{n-1} = r)\, P(U_r = k).$$

2. On considère une famille $(\alpha_{i,j})$ de réels positifs ou nuls, indexée par un couple d'entiers. On suppose, d'une part, que pour tout entier j, la série $\sum_{i=0}^{+\infty} \alpha_{i,j}$ est convergente (de somme β_j) et d'autre part que la série $\sum_{j=0}^{+\infty} \beta_j$ est convergente.

a. Montrer que pour tout triplet de nombres entiers (k, p, n) on a :

$$\sum_{i=k}^{k+p} \left(\sum_{j=0}^{n} \alpha_{i,j} \right) \leqslant \sum_{j=0}^{+\infty} \beta_j .$$

En déduire que, pour tout entier i, $\sum_{j=0}^{\infty} \alpha_{i,j}$ est convergente, puis que la série :

$$\sum_{i=0}^{+\infty} \left(\sum_{j=0}^{+\infty} \alpha_{i,j} \right) \text{ est convergente .}$$

b. Prouver enfin que :

$$\sum_{i=0}^{+\infty} \left(\sum_{j=0}^{+\infty} \alpha_{i,j} \right) = \sum_{j=0}^{+\infty} \left(\sum_{i=0}^{+\infty} \alpha_{i,j} \right).$$

3. En utilisant I.3.*a.*, prouver qu'il existe un intervalle $[0, a_2[$ contenant $[0, 1]$ tel que pour $x \in [0, a_2[$, la série :

$$\sum_{r=0}^{+\infty} p_r \left(\sum_{k=0}^{+\infty} P(U_k = k)\, x^k \right) \text{ soit convergente.}$$

En déduire que si $x \in [0, a_2[$,

$$\sum_{k=0}^{\infty} P(Z_2 = k)\, x^k = \sum_{r=0}^{\infty} p_r (f(x))^r$$

et que la fonction génératrice de Z_2 est égale à la fonction f_2 (voir les fonctions f_n de I) sur l'intervalle $[0, a_2[$.

4. Montrer également que le rayon de convergence de la série $\sum_{k=0}^{+\infty} P(Z_n = k)\, x^k$ est strictement supérieur à 1 et que la fonction génératrice de Z_n n'est autre que la fonction f_n de la suite introduite dans la partie I.

II.3. **Applications aux calculs des moments.**

a. Montrer, sous les hypothèses envisagées, que les moments $E(X)$ et $\sigma^2(X)$ existent. Déterminer ces nombres à l'aide de $f'(1)$ et de $f''(1)$.

b. Déterminer à l'aide de $E(X) = m$ et $\sigma^2(X) = \sigma^2$, les espérances et les variances des variables Z_n (on pourra procéder par récurrence).

Tournez la page S.V.P.

ENS CACHAN S. Nat.

II.4. **Probabilité d'extinction de la lignée.**

a. Déterminer à l'aide de $f_N(0)$ la probabilité P_N d'extinction de la lignée avant la N - 1-ième génération, soit $P(\exists\, k < N + 1,\ Z_k = 0)$.

b. On définit la probabilité d'extinction de la lignée par

$$P(\exists\, N, Z_N = 0) = P\left(\bigcup_{n=1}^{+\infty} (Z_n = 0)\right).$$

Montrez alors que cette probabilité est égale à $\lim_{N \to \infty} P_N$, et est un nombre r solution de $f(r) = r$.

c. Montrez que $r = 1$ si $m \leqslant 1$ et que $r < 1$ si $m > 1$.

PARTIE III. – ÉTUDE D'EXEMPLES

III.1. **Cas d'une variable géométrique.**

On reprend l'étude faite dans I.3. pour les données $P(X = k) = qp^k$ pour tout k entier $\geqslant 0$ (les nombres p et q étant cette fois donnés *distincts* strictement positifs tels que $p + q = 1$).

a. Calculez la fonction génératrice de X ainsi que son rayon de convergence ; en déduire l'expression de E(X). Calculer suivant la valeur de p la probabilité d'extinction de la lignée.

b. Montrer que, pour tout $n \geqslant 1$, la fonction génératrice de Z_n, soit $f_n(x)$, peut s'écrire sous la forme :

$$f_n(x) = \frac{a_n + c_n x}{b_n + d_n x}$$

où les suites (a_n), (b_n), (c_n), (d_n) vérifient $a_1 = q$, $b_1 = 1$, $c_1 = 0$, $d_1 = -p$ et des relations de récurrence de la forme :

$$\begin{cases} a_{n+1} = \alpha_{11} a_n + \alpha_{12} b_n \\ b_{n+1} = \alpha_{21} a_n + \alpha_{22} b_n \\ c_{n+1} = \alpha_{11} c_n + \alpha_{12} d_n \\ d_{n+1} = \alpha_{21} c_n + \alpha_{22} d_n \end{cases} \quad \text{pour tout entier } n \geqslant 1.$$

Calculez la matrice $A = (\alpha_{ij})_{\substack{1 \leqslant i \leqslant 2 \\ 1 \leqslant j \leqslant 2}}$.

c. Calculez à l'aide de p et q les valeurs propres de la matrice A. La matrice A est-elle diagonalisable ? Si oui, exprimez A sous la forme $A = PDP^{-1}$ où D est une matrice diagonale.

d. Calculez alors A^n pour tout $n \geqslant 1$. En déduire la forme explicite de a_n, b_n, c_n, d_n en fonction de p et q, pour tout $n \geqslant 1$. Retrouvez dans ce cas le résultat du III.1.*a.*

III.2. **Cas** $p_0 + p_1 + p_2 = 1$.

a. On considère dans cette question le cas $p_k = 0$ pour tout $k \geqslant 3$. Trouver suivant les valeurs de p_0 et p_2 la probabilité d'extinction de la lignée.

Dans toute la suite on considère le cas particulier $p_0 = p_2 = \frac{1}{4}$, $p_1 = \frac{1}{2}$ et on pose pour tout $n \geqslant 1$, $a_n = \mathrm{P}(Z_n = 0)$.

b. Montrer que la suite $(a_n)_{n \geqslant 0}$ vérifie l'équation de récurrence $a_{n+1} = \left(\frac{1 + a_n}{2}\right)^2$, $a_0 = 0$.

Retrouver dans ce cas particulier le résultat obtenu au III.2.*a.*

c. Soit, pour tout $n \geqslant 0$, $t_n = \dfrac{4}{1 - a_n}$.

Montrer que l'on a pour tout $n \geqslant 0$:

$$t_{n+1} = t_n + 1 + \frac{1}{t_n - 1}.$$

En déduire la double inégalité :

$n + 4 \leqslant t_n \leqslant n + 4 + \ln(n + 2)$ pour tout $n \geqslant 0$, et le fait que la série de terme général $(1 - a_n)$ est divergente.

d. Soit T la variable aléatoire égale au temps de durée de la famille, soit encore $\mathrm{T} = \inf(n \mid Z_n = 0)$. Montrer alors que :

$$\mathrm{E}(\mathrm{T}) = \sum_{n=0}^{+\infty} \mathrm{P}(\mathrm{T} > n) = \sum_{n=0}^{+\infty} (1 - a_n) = +\infty$$

(pour la première inégalité on admettra le fait que II.2.*b.* se généralise à toute famille quelconque $(\alpha_{i,j})$ de réels positifs ou nuls).