84.07B

SESSION 2008

Filière BCPST

MATHÉMATIQUES

Epreuve commune aux ENS de Paris, Lyon et Cachan

Durée : 4 heures

L'usage de calculatrice est interdit.

L'objectif de ce sujet est d'étudier les propriétés de modèles simples pour la croissance d'une sous-population de cellules mutantes au sein d'une population de cellules sauvages.

Le sujet comporte cinq parties relativement indépendantes et de difficultés variées, pouvant être traitées dans le désordre. Toutefois, les résultats, admis ou démontrés, de la deuxième partie intitulée « Fonctions génératrices » sont censés pouvoir servir dans les trois parties qui suivent. Il est donc recommandé de parcourir le sujet dans son intégralité avant de démarrer. Enfin, il est demandé de veiller au soin de la présentation, à la rigueur et à la concision des raisonnements.

Notations

Le logarithme (neperien) est désigné par ln. L'ensemble des entiers naturels est noté $\mathbb{N}$ et l'ensemble des nombres réels $\mathbb{R}$.

Par convention, une *somme* indicée par un ensemble vide est nulle, et un *produit* indicé par un ensemble vide est égal à 1. Pour tout entier naturel n, la notation $n!$ désigne le produit des n premiers entiers non nuls, et en particulier $0! = 1$. On rappelle que le nombre de parties à k éléments d'un ensemble à n éléments est noté C_n^k et vaut $\dfrac{n!}{k!(n-k)!}$.

Une variable aléatoire à valeurs dans $\mathbb{N}$ sera appelée v.a.e., pour *variable aléatoire entière.*

Première partie : un modèle déterministe simple

Dans cette partie, on s'intéresse au comportement moyen (non aléatoire), en fonction du temps (entier), d'une population de cellules composée de deux types, les cellules dites *sauvages* et les cellules dites *mutantes.* On suppose ici que les cellules ne meurent pas et qu'à chaque pas de temps, chaque cellule sauvage produit en moyenne s cellules filles, et chaque cellule mutante m cellules filles, où m et s sont deux réels positifs. Au temps 0, la population est constituée d'une seule cellule, et cette cellule est sauvage.

1. Interpréter le modèle suivant :

$$\left\{\begin{array}{lcl} \mathrm{S}_{n+1} & = & (1+\beta s)\mathrm{S}_n \\ \mathrm{M}_{n+1} & = & (1+m)\mathrm{M}_n + \alpha s \mathrm{S}_n \end{array}\right. \qquad n \geq 0,$$

ainsi que les réels $\alpha \in \left]0\,;1\right]$ et $\beta = 1-\alpha$. Que valent S_0 et M_0 ?

2. Donner une expression de S_n en fonction de β, s et n.

3. ***a)*** Soient deux suites $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ liées pour tout entier n par la relation $v_{n+1} = av_n + u_n$, où a est un nombre réel quelconque. Exprimer v_n en fonction de a, de n, de v_0 et des n premiers éléments de la suite u.

b) Exprimer M_n en fonction de α, β, s, m et n. On distinguera les cas $\beta s = m$ et $\beta s \neq m$.

4. ***a)*** Suivant la position de βs par rapport à m, établir un équivalent asymptotique, lorsque $n \to \infty$, de S_n, M_n, et $\mathrm{M}_n + \mathrm{S}_n$.

b) Montrer que la proportion asymptotique de cellules mutantes vaut 1 si $\beta s \leq m$, et $\alpha s/(s-m)$ si $\beta s > m$.

Deuxième partie : fonctions génératrices

Pour toute v.a.e. X, on définit f sa *fonction génératrice* par :

$$f(x) = \sum_{n=0}^{\infty} p_n \, x^n \qquad x \in [-1;1],$$

où, par définition, $p_n = \mathbb{P}(\mathrm{X} = n)$.

1. Montrer que f est bien définie sur $[-1;1]$, calculer $f(0)$ et $f(1)$.

2. De manière générale, pour toute suite réelle $a = (a_n)_{n \in \mathbb{N}}$, on définit *formellement* la *série entière* de terme général a_n par

$$f(x) = \sum_{n=0}^{\infty} a_n \, x^n,$$

et l'on cherche notamment à déterminer le domaine de définition de f.

Soit $\mathcal{E}(a)$ l'ensemble suivant :

$$\mathcal{E}(a) = \{r \in \mathbb{R} : (a_n r^n)_{n \in \mathbb{N}} \text{ est bornée}\},$$

et $\mathrm{R}(a) \geq 0$ sa borne supérieure.

a) Montrer que si $\mathrm{R}(a) > 0$, alors pour tout $r \in \,]-\mathrm{R}(a);\mathrm{R}(a)[$,

a–i) $r \in \mathcal{E}(a)$,

a–ii) $\lim_{n \to \infty} a_n r^n = 0$

a–iii) la série de terme général $a_n r^n$ converge absolument.

b) Montrer que pour tout $r \notin [-\mathrm{R}(a);\mathrm{R}(a)]$, la série de terme général $a_n r^n$ diverge. En déduire deux inclusions entre les ensembles $]-\mathrm{R}(a);\mathrm{R}(a)[$, $[-\mathrm{R}(a);\mathrm{R}(a)]$ et $\mathcal{E}(a)$.

On appelle $\mathrm{R}(a)$ *rayon de convergence* de la série entière f.

3. Calculer le rayon de convergence des séries entières de terme général :

a) $a_n = c^n$, en fonction de la valeur du nombre réel c ;

b) $a_n = n^\alpha$, en fonction de la valeur du nombre réel α ;

4. Soient f la série entière de terme général a_n et g la série entière de terme général b_n. On suppose que les rayons de convergence $\mathrm{R}(a)$ et $\mathrm{R}(b)$ de f et g respectivement, sont non nuls, et l'on définit $\mathrm{U} = \min(\mathrm{R}(a), \mathrm{R}(b)) > 0$. Soit

$$c_n = \sum_{k=0}^{n} a_k \, b_{n-k} \qquad n \geq 0.$$

a) Montrer que la série de terme général $c_n r^n$ converge absolument pour tout $r \in \,]-\mathrm{U};\mathrm{U}[$.

On admettra par la suite que pour tout $x \in \left]-\mathrm{U}\,;\mathrm{U}\right[$,

$$\sum_{n=0}^{\infty} c_n\, x^n = f(x)g(x).$$

b) Montrer que si X et Y sont deux v.a.e. indépendantes dont les fonctions génératrices sont désignées respectivement par f et g, alors la fonction génératrice de $\mathrm{X}+\mathrm{Y}$ est fg.

5. *a)* Soit $\mathrm{X}_1, \mathrm{X}_2, \ldots$ une suite de v.a.e. indépendantes et de même loi, de fonction génératrice commune g. Pour tout entier $k \geq 1$, donner la fonction génératrice de la v.a.e. $\mathrm{X}_1 + \cdots + \mathrm{X}_k$.

b) Soient N une v.a.e. de fonction génératrice f indépendantes des $\mathrm{X}_1, \mathrm{X}_2, \ldots$ Montrer que la somme $\mathrm{X}_1 + \cdots + \mathrm{X}_\mathrm{N}$ est une v.a.e. de fonction génératrice $f \circ g$.

Dans la suite, **on admettra** que toute série entière f de terme général a_n est dérivable sur $\left]-\mathrm{R}(a)\,;\mathrm{R}(a)\right[$, et que

$$f'(x) = \sum_{n\geq 1} n a_n x^{n-1} \qquad x \in \left]-\mathrm{R}(a)\,;\mathrm{R}(a)\right[.$$

On pourra également se servir des trois égalités suivantes :

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad x \in \mathbb{R},$$

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n} \qquad x \in \left[-1\,;1\right[,$$

$$(1-x)^{-a} = \sum_{n=0}^{\infty} \frac{a(a+1)\cdots(a+n-1)}{n!}\, x^n \qquad a \in \mathbb{R}, x \in \left]-1\,;1\right[.$$

On admettra enfin qu'un développement en série entière est *unique*, c'est-à-dire que s'il existe deux suites réelles a et b, et un nombre réel $\mathrm{U} > 0$, tels que pour tout $x \in \left]-\mathrm{U}\,;\mathrm{U}\right[$,

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n,$$

alors $a_n = b_n$ quel que soit $n \geq 0$.

Troisième partie : la distribution de Luria–Delbrück

On suppose qu'il existe une suite de réels positifs $(p_n)_{n\in\mathbb{N}}$ et une fonction f continue sur $[-1\,;1]$, tels que

$$f(x) = \sum_{n=0}^{\infty} p_n\, x^n \qquad x \in [-1\,;1],$$

et

$$f(x) = (1-x)^{\alpha\frac{1-x}{x}} \qquad x \in \,]-1\,;1\,[\setminus\{0\},$$

où α est un réel strictement positif.

1. Calculer les limites de f en 0 et en 1. En déduire que $(p_n)_{n\in\mathbb{N}}$ est une loi de probabilité sur $\mathbb{N}$, que l'on appellera *distribution* (ou loi) *de Luria–Delbrück.*

2. En dérivant $\ln(f)$, montrer que

$$\frac{f'}{f}(x) = \alpha\sum_{n=0}^{\infty}\frac{x^n}{n+2} \qquad x \in \,]-1\,;1\,[\setminus\{0\}.$$

3. ***a)*** Établir la relation de récurrence suivante :

$$np_n = \alpha\sum_{k=0}^{n-1}\frac{p_k}{n+1-k} \qquad n \geq 1.$$

b) Calculer p_1, p_2, p_3.

4. ***a)*** Écrire $(1-x)\ln(1-x)$ sous la forme d'une série entière sur l'intervalle $]-1\,;1\,[$.

b) Soit $a_0 = 0$ et

$$a_n = \frac{1}{n(n+1)} \qquad n \geq 1.$$

Montrer que $(a_n)_{n\in\mathbb{N}}$ est une loi de probabilité, calculer sa fonction génératrice et son espérance.

5. ***a)*** Soient $X_1, X_2, \ldots$ une suite de v.a.e. indépendantes et de même loi, dont on note g la fonction génératrice commune. Soit N une variable aléatoire de Poisson de paramètre $\theta > 0$, indépendante de la suite $(X_n)_{n\in\mathbb{N}}$. On définit alors la variable aléatoire Y par

$$Y = \sum_{i=1}^{N} X_i.$$

Calculer la fonction génératrice de Y.

b) Trouver g et θ pour que Y suive la loi de Luria–Delbrück.

Quatrième partie : étude succincte du modèle de Luria–Delbrück

On considère une population de cellules sauvages et mutantes. Au temps 0, la population démarre avec une seule cellule, qui est sauvage. À chaque pas de temps $1, 2, 3, \ldots$, une cellule et une seule, prise au hasard (uniformément) dans la population, se divise en deux ; par conséquent, entre les temps $n-1$ et n, le nombre de cellules est n. On se donne un nombre réel $\alpha \in [\,0\,;1\,]$, qui est la *probabilité de mutation*, et l'on note $\beta = 1-\alpha$:

– lorsqu'une cellule sauvage se divise, elle se divise en deux cellules sauvages avec probabilité β ; avec probabilité α, en une cellule sauvage et une cellule mutante ;
– lorsqu'une cellule mutante se divise, elle se divise toujours en deux cellules mutantes.

On note $\mathbb{P}_\alpha$ la probabilité associée à ce modèle.

On désigne par S_n le nombre de cellules sauvages présentes dans la population entre les temps $n-1$ et n, et par $\mathrm{M}_n = n - \mathrm{S}_n$ le nombre de cellules mutantes. Enfin, on empruntera les notations suivantes :

$$p_s(n,k;\alpha) = \mathbb{P}_\alpha(\mathrm{S}_n = k) \quad \text{et} \quad p_m(n,k;\alpha) = \mathbb{P}_\alpha(\mathrm{M}_n = k),$$

pour tous entiers naturels k et n.

1. ***a)*** Donner $p_m(n,k;\alpha)$ pour tout $k \geq n$.

b) Lorsque $\alpha \neq 0$, donner la loi du temps d'apparition de la première cellule mutante, ainsi que son espérance.

2. ***a)*** Montrer que la suite $(\mathrm{M}_n)_{n\in\mathbb{N}}$ a une limite, finie ou infinie, que l'on notera M_∞. On cherche à calculer $\mathbb{P}_\alpha(\mathrm{M}_\infty = +\infty)$.

b) Soient k, n_0, p des entiers non nuls. Établir l'égalité suivante :

$$\mathbb{P}_\alpha(\mathrm{M}_n = k\,,\, \forall n \in \{n_0+1, \ldots, n_0+p+1\} \mid \mathrm{M}_{n_0} = k) = \prod_{i=n_0}^{n_0+p} \frac{(1-\alpha)(i-k)}{i}$$

c) Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres réels appartenant à l'intervalle $]\,0\,;1\,[$. Montrer que la suite $\prod_{k=0}^{n} a_k$ est convergente, et que sa limite est nulle si et seulement si la série de terme général $\ln(a_n)$ diverge.

d) d-i) Montrer que pour tout $k \geq 1$, même si $\alpha = 0$,

$$\mathbb{P}_\alpha(\mathrm{M}_n = k\,,\, \forall n \geq n_0+1 \mid \mathrm{M}_{n_0} = k) = 0.$$

d-ii) En déduire que $\mathbb{P}_\alpha(\mathrm{M}_\infty = k) = 0$ pour tout $k \geq 1$.

e) Donner la valeur de $\mathbb{P}_\alpha(\mathrm{M}_\infty = +\infty \text{ ou } 0)$, et en déduire, selon la valeur de α, celles de $\mathbb{P}_\alpha(\mathrm{M}_\infty = 0)$ et de $\mathbb{P}_\alpha(\mathrm{M}_\infty = +\infty)$.

Cinquième partie : loi du nombre de cellules mutantes

Le but de cette partie est d'établir explicitement la loi de M_n, puis d'obtenir son comportement asymptotique lorsque la probabilité de mutation est faible, de l'ordre de $1/n$. On rappelle avoir défini $p_s(n,k;\alpha) = \mathbb{P}_\alpha(\mathrm{S}_n = k)$ et $p_m(n,k;\alpha) = \mathbb{P}_\alpha(\mathrm{M}_n = k)$.

1. Montrer la relation de récurrence suivante pour *tous* entiers naturels k et n :

$$p_s(n+1,k+1;\alpha) = \frac{k}{n}\beta p_s(n,k;\alpha) + \left(1 - \frac{k+1}{n}\beta\right) p_s(n,k+1;\alpha).$$

2. Montrer que pour tous entiers naturels k et n,

$$p_s(n,k;\alpha) = \frac{1}{(n-1)!}\sum_{i=1}^{k}(-1)^{i-1}\mathrm{C}_{k-1}^{i-1}\prod_{j=1}^{n-1}(j-i\beta)\,.$$

3. Montrer également que pour tous entiers $1 \leq k \leq n$,

$$p_m(n,k;1/2) = \mathrm{C}_{n+k-1}^{k}\left(\frac{1}{2}\right)^{n+k-1}.$$

4. À partir de maintenant, pour tout $k \in \mathbb{N}$ et tout $x \in \,]-1\,;1\,[$, on définit

$$\gamma_k(x;\alpha) = \sum_{n=1}^{\infty} p_s(n,k;\alpha)x^{n-1}.$$

a) Montrer que γ_k est bien définie.

b) Établir l'égalité suivante en utilisant la question 2 et sans justifier les interversions de sommes :

$$\gamma_k(x;\alpha) = c(1-x)^{-\delta}(1-(1-x)^{\eta})^{k-1} \qquad x \in [\,-1\,;1\,]\,,$$

où c, δ, et η sont des nombres réels positifs à déterminer.

5. Montrer que pour tout $k \in \mathbb{N}$,

$$x^{1-k}\gamma_k(x;\alpha) = \sum_{n=0}^{\infty} p_m(n+k,n;\alpha)x^n \qquad x \in [\,-1\,;1\,]\,.$$

6. ***a)*** Pour tout $x \in \,]\,0\,;1\,[$, établir la limite suivante :

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\ln\left(\frac{1-e^{(1-\varepsilon)\ln(1-x)}}{x}\right) = \frac{1-x}{x}\ln(1-x).$$

b) En déduire que pour tout $x \in \,]\,0\,;1\,[$,

$$\lim_{n\to\infty} x^{1-n}\gamma_n(x;\alpha/n) = (1-x)^{\alpha\frac{1-x}{x}}.$$

7. Expliquer pourquoi la distribution $(p_n)_{n\in\mathbb{N}}$ dite de Luria–Delbrück de la troisième partie fournit un candidat pour la limite p_k de $\mathbb{P}_{\alpha/n}(\mathrm{M}_n = k)$, et donner le ou les principaux obstacles à une démonstration mathématique rigoureuse.