64.05B

SESSION 2006

Filière BCPST

MATHÉMATIQUES

Epreuve commune aux ENS de Paris, Lyon et Cachan

Durée : 4 heures

L'usage de calculatrice est interdit.

Le but de ce sujet est de définir et de résoudre partiellement trois modèles concernant l'abondance d'une espèce ou les abondances respectives de plusieurs espèces en coexistence.

Le sujet comporte trois parties indépendantes de longueurs inégales pouvant être traitées dans un ordre quelconque.
Les candidats composeront sur des copies séparées pour chaque partie, en les identifiant clairement. Il est recommandé de veiller au soin de la présentation, à la rigueur et à la concision des raisonnements. Les démonstrations s'appuyant sur des représentations graphiques pourront s'avérer utiles.

Notations

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

Pour tout entier naturel n, la notation $n!$ désigne le produit des n premiers entiers, où par convention $0! = 1$. On rappelle que le nombre de parties à k éléments d'un ensemble à n éléments est noté C_n^k et vaut $\dfrac{n!}{k!(n-k)!}$.

L'espérance mathématique d'une variable aléatoire réelle (v.a.r.) X est notée $\mathbb{E}(\mathrm{X})$ et sa variance $\mathrm{Var}(\mathrm{X})$. La covariance de deux v.a.r. X et Y est notée $\mathrm{Cov}(\mathrm{X}, \mathrm{Y})$.
Pour toute v.a.r. X et tout nombre réel x, $\mathbb{P}(\mathrm{X} \in [\,x\,; x + dx\,[)$ est abrégé en $\mathbb{P}(\mathrm{X} \in dx)$.
L'indicatrice $\mathbf{1}_\mathrm{A}$ d'un événement A est la variable aléatoire qui vaut 1 si A est réalisé et 0 sinon.
En particulier, $\mathbb{E}(\mathbf{1}_\mathrm{A}) = \mathbb{P}(\mathrm{A})$.

Première partie : le partage aléatoire de MacArthur

L'une des plus anciennes façons de modéliser la répartition des ressources (et donc des abondances) entre $n+1$ espèces, consiste à supposer que la quantité totale des ressources disponibles est constante et est divisée aléatoirement entre les $n+1$ espèces présentes. Pour ce faire, on « jette » n points uniformément et indépendamment dans l'intervalle $[\,0\,;1\,]$, qui définissent $n+1$ fragments adjacents aléatoires et de même loi, mais de somme 1. Chacun de ces fragments représente la part des ressources allouée à chaque espèce.

Plus précisément, on considère n variables indépendantes $\mathrm{U}_1, \ldots, \mathrm{U}_n$ uniformes sur $[\,0\,;1\,]$. On désigne par $\mathrm{X}_1 \leqslant \cdots \leqslant \mathrm{X}_n$, le réarrangement croissant des $\mathrm{U}_1, \ldots, \mathrm{U}_n$. Les $n+1$ fragments adjacents sont donc de tailles $\mathrm{X}_1 - \mathrm{X}_0, \mathrm{X}_2 - \mathrm{X}_1, \ldots, \mathrm{X}_n - \mathrm{X}_{n-1}, \mathrm{X}_{n+1} - \mathrm{X}_n$, où par convention $\mathrm{X}_0 = 0$ et $\mathrm{X}_{n+1} = 1$. Il est à nouveau possible de ranger ces fragments dans l'ordre croissant de leurs *tailles*, soit par définition $\mathrm{Y}_1 \leqslant \cdots \leqslant \mathrm{Y}_{n+1}$.

Pour spécifier que le nombre de fragments considérés est $n+1$ (bien que le nombre de points jetés soit n), nous noterons la probabilité $\mathbb{P}_{n+1}$.

1. ***a)*** Montrer que pour pour tous entiers $1 \leqslant k \leqslant n+1$, et tout réel $x \in]\,0\,;1\,[$,

$$\mathbb{P}_{n+1}(\mathrm{X}_k \in dx) = n\,\mathrm{C}_{n-1}^{k-1}\, x^{k-1}(1-x)^{n-k}\, dx.$$

b) En déduire une expression générale, pour tous entiers i et j, de

$$\int_0^1 x^i(1-x)^j\, dx.$$

c) Montrer que $\mathbb{E}_{n+1}(\mathrm{X}_k - \mathrm{X}_{k-1}) = 1/(n+1)$.

d) Expliquer pourquoi en général, bien que les $(\mathrm{Y}_k)_{1\leqslant k\leqslant n+1}$ soient une renumérotation des $(\mathrm{X}_k-\mathrm{X}_{k-1})_{1\leqslant k\leqslant n+1}$, $\mathbb{E}_{n+1}(\mathrm{Y}_k)\neq 1/(n+1)$. Montrer en particulier que

$$\mathbb{E}_{n+1}(\mathrm{Y}_{n+1}) > \frac{1}{n+1}.$$

2. Le but de cette question est de déterminer la loi du plus petit fragment, Y_1. On note I_{n+1} l'intervalle ouvert $]\,0\,;1/(n+1)\,[$.

a) a–i) Montrer que $\mathbb{P}_{n+1}(\mathrm{Y}_1\in \mathrm{I}_{n+1})=1$.

a–ii) Établir que pour $y\in \mathrm{I}_2$, $\mathbb{P}_2(\mathrm{Y}_1>y)=1-2y$.

b) b–i) Soient $a\in\,]\,0\,;1\,[$ et f une fonction réelle de $[\,0\,;1\,]^n$. Montrer que si $\mathrm{U}_1,\dots,\mathrm{U}_n$ sont des v.a. indépendantes uniformes sur $[\,0\,;1\,]$,

$$\mathbb{E}\left(f(\mathrm{U}_1,\dots,\mathrm{U}_n)\mathbf{1}_{\{\mathrm{U}_i<a,\,\forall i\}}\right)=a^n\,\mathbb{E}\left(f(a\mathrm{U}_1,\dots,a\mathrm{U}_n)\right).$$

b–ii) Pour $y\in \mathrm{I}_{n+1}$, justifier l'égalité

$$\mathbb{P}_{n+1}(\mathrm{Y}_1\in dy)=n(n+1)(1-y)^{n-1}\mathbb{P}_n((1-y)\mathrm{Y}_1>y)\,dy.$$

Indication. On pourra se ramener aux cas où $\mathrm{Y}_1=\mathrm{X}_{n+1}-\mathrm{X}_n$.

c) Montrer par récurrence sur $n\geqslant 1$ que pour tout $y\in \mathrm{I}_{n+1}$,

$$\mathbb{P}_{n+1}(\mathrm{Y}_1>y)=(1-(n+1)y)^n.$$

d) En déduire que

$$\mathbb{E}_{n+1}(\mathrm{Y}_1)=\frac{1}{(n+1)^2}.$$

3. *a) a–i)* Soient $a\in\,]\,0\,;1\,[$ et U une v.a. uniforme sur $[\,0\,;1\,]$. Montrer que la loi de U sachant $\{\mathrm{U}>a\}$ est la même que celle de $a+(1-a)\mathrm{U}$.

a–ii) Soit $y\in \mathrm{I}_{n+1}$. Justifier, par exemple à l'aide d'une représentation graphique, que la loi du n-uplet $(\mathrm{Y}_2,\dots,\mathrm{Y}_{n+1})$ sous $\mathbb{P}_{n+1}(\cdot\,|\,\mathrm{Y}_1=y)$, est la même que celle, sous $\mathbb{P}_n$, du n-uplet

$$(y+(1-(n+1)y)\mathrm{Y}_1,\dots,y+(1-(n+1)y)\mathrm{Y}_n).$$

b) Pour $k\in\{1,\dots,n\}$, en déduire une relation entre $\mathbb{E}_{n+1}(\mathrm{Y}_{k+1})$ et $\mathbb{E}_n(\mathrm{Y}_k)$.

c) Pour $k\in\{1,\dots,n\}$, on définit $\alpha_k^{(n)}=n\mathbb{E}_n(\mathrm{Y}_k)$. Donner une relation entre $\alpha_{k+1}^{(n+1)}$ et $\alpha_k^{(n)}$, puis établir le résultat final :

$$\mathbb{E}_n(\mathrm{Y}_{k+1})=\frac{1}{n}\sum_{i=n-k}^{n}\frac{1}{i}\qquad 0\leqslant k\leqslant n-1.$$

Deuxième partie : la métaphore de la cantine

Dans cette partie il s'agit de caractériser la loi du nombre d'espèces représentées dans un échantillon de n individus, et leurs abondances respectives, à l'aide d'un unique paramètre.

Soit θ un réel strictement positif. Des individus numérotés $1, 2, \ldots, n$, arrivent successivement dans une salle de restaurant contenant une infinité de tables infiniment longues. Le premier individu s'assied à une table au hasard. Pour tout entier $k \geqslant 1$, lorsque l'individu $k+1$ arrive, il choisit au hasard un des k convives déjà attablés avec la probabilité $1/(k+\theta)$, et s'assied à la même table, ou occupe une nouvelle table avec la probabilité $\theta/(k+\theta)$.

L'entier K_n désigne le nombre de tables occupées lorsque n convives se sont installés et pour $1 \leq i \leq n$, on note $q_{n,i} = \mathbb{P}(\mathrm{K}_n = i)$. La répartition de ces n convives en K_n tables est une métaphore pour la répartition d'un échantillon de n individus vivants en K_n espèces.

1. *a)* Montrer que

$$q_{n+1,1} = \frac{n!}{(n+\theta)(n-1+\theta)\cdots(1+\theta)} \,.$$

b) Pour tous $2 \leqslant i \leqslant n$, trouver une relation entre $q_{n+1,i}$, $q_{n,i}$ et $q_{n,i-1}$.

2. Soient L_n et P_n les polynômes de degré n suivants

$$\mathrm{P}_n(\mathrm{X}) = \sum_{i=1}^{n} q_{n,i}\, \mathrm{X}^i,$$

$$\mathrm{L}_n(\mathrm{X}) = \prod_{i=0}^{n-1} (\mathrm{X}+i).$$

a) Donner une relation de récurrence vérifiée par (P_n).

b) En déduire que

$$\mathrm{P}_n(\mathrm{X}) = \frac{\mathrm{L}_n(\theta \mathrm{X})}{\mathrm{L}_n(\theta)} \,.$$

On admettra que cette équation caractérise la loi de K_n, mais dans la question suivante, on se concentre sur son espérance et sa variance.

3. *a)* Montrer que $\mathbb{E}(\mathrm{K}_n) = \mathrm{P}'_n(1)$ et en déduire $\mathbb{E}(\mathrm{K}_n)$.

Indication. On pourra prendre le logarithme de P_n.

b) Montrer que $\mathrm{Var}(\mathrm{K}_n) = \mathrm{P}''_n(1) + \mathrm{P}'_n(1) - (\mathrm{P}'_n(1))^2$ et calculer $\mathrm{Var}(\mathrm{K}_n)$.

4. Dans cette question, on cherche à obtenir directement les résultats de la question précédente.

a) Montrer que

$$\mathrm{K}_n = \sum_{i=1}^{n} \varepsilon_i,$$

où les $(\varepsilon_i)_{i=1,\ldots,n}$ sont des variables de Bernoulli indépendantes dont on précisera les probabilités de succès respectives.

b) En déduire $\mathbb{E}(\mathrm{K}_n)$ et $\mathrm{Var}(\mathrm{K}_n)$.

5. ***a)*** Établir la double inégalité

$$1 + \int_1^n \frac{\theta}{\theta + x}\, dx \leqslant \mathbb{E}(\mathrm{K}_n) \leqslant 1 + \int_1^{n-1} \frac{\theta}{\theta + x}\, dx$$

b) Donner un équivalent de $\mathbb{E}(\mathrm{K}_n)$ lorsque $n \to \infty$ (et le justifier).

6. Étudier la différence $\mathrm{Var}(\mathrm{K}_n) - \mathbb{E}(\mathrm{K}_n)$ et en déduire un équivalent de $\mathrm{Var}(\mathrm{K}_n)$ lorsque $n \to \infty$.

Troisième partie : la série logarithmique de Fisher

Le but de cette partie est d'établir un résultat ancien sur la loi du nombre d'individus appartenant à une *même* espèce.

A La fonction Γ

1. Montrer que la fonction $\Gamma : x \longmapsto \displaystyle\int_0^\infty t^{x-1}\,\mathrm{e}^{-t}\, dt$ est bien définie sur $]\,0\,;+\infty\,[$.

2. ***a)*** Pour tout réel positif x, montrer que $\Gamma(x+1) = x\Gamma(x)$.

b) En déduire une expression pour $\Gamma(n)$ lorsque n est un entier naturel non nul.

c) Donner un équivalent de $\Gamma(x)$ lorsque $x \to 0^+$.

Dorénavant, pour tout réel positif θ, nous dirons qu'une v.a.r positive suit la loi $\mathrm{Gamma}(\theta)$ si elle a pour densité de probabilité la fonction

$$t \longmapsto \frac{1}{\Gamma(\theta)} t^{\theta-1}\,\mathrm{e}^{-t} \qquad t > 0.$$

3. Calculer l'espérance et la variance d'une variable aléatoire X de loi $\mathrm{Gamma}(\theta)$.

B Étude de la série logarithmique

Soit f la fonction définie par

$$f : \quad]-\infty\,;+1\,[\longrightarrow \mathbb{R}$$
$$x \longmapsto \ln(1-x)$$

4. Montrer que pour tout $x \in]-1;1[$, la série de terme général $\dfrac{x^n}{n}$ converge.

5. Établir que pour tout entier n, la dérivée n-ième de f est donnée par :

$$f^{(n)}(x) = -\frac{(n-1)!}{(1-x)^n} \qquad x < 1.$$

6. Montrer que pour tout $x < 1$ et tout entier $n \geqslant 1$,

$$f(x) = -\sum_{k=1}^{n} \frac{x^k}{k} - \int_0^x \frac{(x-t)^n}{(1-t)^{n+1}}\, dt$$

7. Pour $x \in]-1;1[$, calculer la somme $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$.

C Abondance d'une espèce

8. Dans cette question *uniquement*, on suppose qu'une espèce $\mathcal{E}$ compte N individus, et qu'il est possible d'observer chacun de ces individus avec la même probabilité p, indépendamment les uns des autres.

a) Donner la loi du nombre X d'individus observés.

b) Soit un réel $d > 0$. Que devient la loi de X lorsque $\mathrm{N} \to \infty$, $p \to 0$ et $p\mathrm{N} \to d$?

On se donne deux réels strictement positifs ρ et θ, et l'on suppose désormais que X suit une loi de Poisson d'espérance D (la vraie densité de population), où pour rendre compte de l'incertitude existant sur elle, D est une v.a.r. indépendante telle que :

$$\mathrm{D} = \rho \mathrm{G},$$

où G suit la loi Gamma(θ).

9. *a)* Montrer que pour k entier naturel et y réel positif,

$$\mathbb{P}(\mathrm{X} = k, \mathrm{G} \in dy) = \frac{1}{\Gamma(\theta)} \frac{\rho^k}{k!} y^{k+\theta-1}\, \mathrm{e}^{-(\rho+1)y}\, dy$$

b) En déduire que

$$\mathbb{P}(\mathrm{X} = k) = \frac{1}{k!} \frac{\Gamma(k+\theta)}{\Gamma(\theta)} \frac{\rho^k}{(\rho+1)^{k+\theta}}.$$

10. *a)* Calculer $\mathbb{P}(\mathrm{G} \in dy \,|\, \mathrm{X} = k)$.

b) Montrer que la v.a. $(\rho+1)\mathrm{D}/\rho$, conditionnée par l'observation de k individus, suit une loi Gamma de paramètre à préciser.

c) Montrer que

$$\mathbb{E}(\mathrm{D} \mid \mathrm{X}=k) = \frac{\rho}{\rho+1}(k+\theta).$$

Commenter la façon dont le nombre d'individus observés prédit la densité de population.

On ne s'intéresse plus dorénavant qu'aux espèces dont au moins un individu a été observé.

11. *a)* Montrer que pour tout entier $k \geqslant 1$ et tout réel $y > 0$

$$\mathbb{P}(\mathrm{X}=k, \mathrm{G} \in dy \mid \mathrm{X} \geqslant 1) = \frac{1}{(1-(\rho+1)^{-\theta})\Gamma(\theta)} \frac{\rho^k}{k!} y^{k+\theta-1} \,\mathrm{e}^{-(\rho+1)y}\, dy$$

b) Prouver la convergence suivante

$$\lim_{\theta \to 0^+} \mathbb{P}(\mathrm{X}=k, \mathrm{G} \in dy \mid \mathrm{X} \geqslant 1) = c\, \frac{\rho^k}{k!}\, y^{k-1} \,\mathrm{e}^{-(\rho+1)y}\, dy,$$

où c est une constante à déterminer.

On appelle $\mathbb{P}^\star$ cette loi limite.

c) En déduire que

$$\mathbb{P}^\star(\mathrm{X}=k) = c\, \frac{x^k}{k} \qquad k \geqslant 1,$$

où $x \in \,]\,0\,;1\,[$ sera exprimé à l'aide de ρ, et que

$$\mathbb{P}^\star(\mathrm{G} \in dy) = c\,\mathrm{e}^{-y} \frac{1-\mathrm{e}^{-\rho y}}{y}\, dy \qquad y > 0.$$

La loi $\mathbb{P}^\star$ est connue sous le nom de *série logarithmique de Fisher.*

d) Calculer $\mathbb{E}^\star(\mathrm{D})$, $\mathbb{E}^\star(\mathrm{X})$, $\mathbb{E}^\star(\mathrm{DX})$, et $\mathrm{Cov}(\mathrm{D},\mathrm{X})$. Quel est le signe de cette covariance ? Commenter.

12. Les espèces connues d'un écosystème sont au nombre de S et leurs abondances sont indépendantes et suivent toutes la loi $\mathbb{P}^\star$.

a) Quel est le nombre attendu M d'espèces singletons (c'est-à-dire dont un seul individu a été observé) ? Quel est le nombre attendu N d'individus observés au total ? Comment peut-on estimer ρ si l'on connaît M et N ?

b) Étudier S/N en tant que fonction de x. Interpréter les cas $x \to 0^+$ et $x \to 1^-$.