Un sujet pour l'épreuve A
(raisonnement et calcul)

20 janvier 2014

1 Présentation

Le texte proposé ci-après est conçu pour l'épreuve A, portant plus particulièrement sur le calcul et le raisonnement.
Dans ce texte, les probabilités apparaissent abondamment mais le contexte modélisant est fixé, et l'enjeu du problème n'est pas de discuter des hypothèses de modélisation mais d'élaborer des conséquences mathématiques du contexte posé *a priori*. C'est pourquoi le contexte modélisant est exposé dès le début de l'énoncé.
Les compétences mobilisées dans ce problème sont essentiellement les suivantes :

- ▷ Raisonner, démontrer, argumenter.
  Cette compétence est notamment présente dans les questions A3c, A3d, A3e, B1e.
- ▷ Calculer, maîtriser le formalisme mathématique.
  Cette compétence est notamment présente dans les questions A1b, A1c, A2b, A3a, A3d, B1c, B2.
- ▷ Mobiliser des connaissances scientifiques pertinentes.
  Cette compétence est notamment présente dans les questions A1b, A1d, A1e, A2a, A3b, B1a, B1b, B1d.
- ▷ Communiquer à l'écrit et à l'oral.
  Cette compétence est présente, par nature, dans l'ensemble des questions.

On rappelle que l'épreuve A dure 2 heures 30 et que l'emploi d'une calculatrice n'y est pas prévu.

2 Énoncé

Le but de ce problème est l'étude de deux modèles de propagation d'un virus. Les deux parties sont indépendantes. Dans tout ce problème, on considère une population $\mathscr{P}$ d'individus, chacun pouvant être porteur d'un virus V. Une unité de temps étant fixée (minute, heure, jour, selon la situation), on note X_0 la variable aléatoire égale au nombre d'individus de $\mathscr{P}$ porteurs de V à l'instant initial de l'étude, et plus généralement X_n la variable aléatoire égale au nombre d'individus porteurs de V au bout de n unités de temps, n étant un entier naturel. La première partie étudie un modèle où X_n est proportionnel à X_{n-1} à chaque instant n. La seconde partie se penche sur un modèle où la dépendance de X_{n+1} en fonction de X_n est fournie par une matrice de transition.

Dans tout ce problème, $E(X)$ désigne, lorsqu'elle existe, l'espérance de la variable aléatoire X.

A. Évolution proportionnelle

Dans cette partie, on propose d'étudier une situation où, à un instant donné n, l'agent contaminant responsable de la transmission de V peut être actif ou inactif (en raison de facteurs extérieurs qu'on ne cherche pas à étudier ici) : à tout instant n, on a la probabilité p qu'il soit actif, et la probabilité $1-p$ qu'il soit inactif, $p \in]0,1[$ étant fixé. On considère de ce fait une suite $(U_n)_{n\in\mathbf{N}}$ de variables aléatoires de même loi de Bernoulli de paramètre p, U_n valant 1 si l'agent est actif à l'instant n, et 0 s'il est inactif. On supposera de plus que les $(U_n)_{n\in\mathbf{N}}$ sont mutuellement indépendantes, et que U_n est indépendante de X_n pour tout $n \in \mathbf{N}$.

On convient que lorsque l'agent contaminant est actif à l'instant $n \in \mathbf{N}$, le nombre d'individus de $\mathscr{P}$ contaminés augmente d'un facteur $\alpha \in]0,1[$ entre les instants n et $n+1$, si bien qu'on a

$$U_n = 1 \Rightarrow X_{n+1} = (1+\alpha)X_n.$$

De façon analogue, on conviendra que

$$U_n = 0 \Rightarrow X_{n+1} = (1-\alpha)X_n.$$

Ce modèle sera déclaré *raisonnable* si la suite (X_n) reste bornée au sens suivant : l'événement

$$\mathscr{B} = \{\text{il existe un réel } B \text{ tel que pour tout } n \in \mathbf{N}, X_n \leqslant B\}$$

(on admet que c'en est bien un) soit quasi-certain (c'est-à-dire, ait une probabilité égale à 1).

A.1. On fixe $n \in \mathbf{N}^*$ dans toute cette question.
a) Établir que $X_n = (1+\alpha)^{U_{n-1}}(1-\alpha)^{1-U_{n-1}}X_{n-1}$.
b) Justifier la formule suivante : $E(X_n) = \big(1+(2p-1)\alpha\big)E(X_{n-1})$.
c) Donner l'espérance de X_n en fonction de X_0.
d) En supposant $E(X_0) > 0$, pour quelles valeurs de p a-t-on $\lim\limits_{n\to+\infty} E(X_n) = +\infty$?
e) Quelle contrainte en découle pour que le modèle étudié soit jugé *raisonnable* ?

A.2. On pose $S_n = \sum\limits_{k=0}^{n-1} U_k$.
a) Quelle est la loi de S_n ?
b) Montrer que $X_n = (1+\alpha)^{S_n}(1-\alpha)^{n-S_n}X_0$.
c) Quelle est, en fonction de X_0, la valeur maximale M_n que peut prendre X_n ?
d) Si $X_0 > 0$, montrer que $\lim\limits_{n\to+\infty} M_n = +\infty$.
Remarque : *le résultat précédent semble empêcher la possibilité que le modèle proposé soit raisonnable. Néanmoins, ce maximum M_n n'est atteint au rang n que si l'agent contaminant s'est montré actif lors des n premiers instants. Si p est faible et n grand, la probabilité d'avoir une telle séquence est infime.*

A.3. Forts de la remarque ci-dessus, nous cherchons dans la suite à quantifier de façon plus rigoureuse le risque que X_n devienne très grand, en évaluant la probabilité $P(X_n > X_0)$ en fonction de $n \in \mathbf{N}^*$ fixé.
a) Montrer que $P(X_n > X_0) = P(S_n > n\theta)$ où $\theta = -\dfrac{\ln(1-\alpha)}{\ln(1+\alpha) - \ln(1-\alpha)}$.
En déduire que pour tout $t > 0$, on a

$$P(X_n > X_0) \leqslant E\left(\mathrm{e}^{tS_n}\right)\mathrm{e}^{-nt\theta}.$$

Indication : l'inégalité de Markov peut servir (mais ce n'est pas la seule méthode possible).
b) Montrer que pour tout $t > 0$

$$P(X_n > X_0) \leqslant \mathrm{e}^{n\phi(t)}$$

où $\phi(t) = \ln(p\mathrm{e}^t + (1-p)) - t\theta$. *Indication :* on peut utiliser l'indépendance des $(U_k)_{0\leqslant k\leqslant n-1}$.
c) Quelle est la limite ℓ de θ lorsque α tend vers 0 ?
d) On admettra dans la suite que α est suffisamment proche de 0 pour qu'on puisse prendre $\theta = \ell$ pour valeur approchée. On suppose de plus que $p = \dfrac{1}{5}$. Montrer que ϕ atteint sur $\mathbf{R}_+^*$ un minimum λ strictement négatif.
e) Que peut-on dire de $P(X_n > X_0)$? Quelle est la nature de la série $\sum_{n\in\mathbf{N}} P(X_n > X_0)$?
Remarque : *on peut démontrer (mais cela n'est pas demandé) que, dans ces conditions, l'événement $\mathscr{B}$ est de probabilité 1 ; le modèle peut ainsi être considéré comme raisonnable pour les valeurs de p et de α choisies précédemment.*

B. Évolution modélisée par une matrice de transition

Soit $N \in \mathbf{N}^*$. Pour une matrice $A = (a_{i,j})_{\substack{0\leqslant i\leqslant N\\ 0\leqslant j\leqslant N}}$ et $n \in \mathbf{N}^*$, on note $(a_{i,j}^{(n)})_{\substack{0\leqslant i\leqslant N\\ 0\leqslant j\leqslant N}}$ les coefficients de la matrice A^n et on dit que la suite de matrices $(A^n)_{n\in\mathbf{N}^*}$ est convergente si et seulement $b_{i,j} = \lim\limits_{n\to+\infty} a_{i,j}^{(n)}$ existe pour tout couple (i,j). Dans ce cas, on note $B = (b_{i,j})_{\substack{0\leqslant i\leqslant N\\ 0\leqslant j\leqslant N}} = \lim\limits_{n\to+\infty} A^n$. On admettra les résultats suivants :

- ▷ si A et A' sont deux matrices carrées et P est une matrice carrée inversible de même taille telles que $A = PA'P^{-1}$, alors $(A^n)_{n\in\mathbf{N}^*}$ converge si et seulement si $(A'^n)_{n\in\mathbf{N}^*}$ converge, et dans ce cas, $\lim\limits_{n\to+\infty} A^n = P(\lim\limits_{n\to+\infty} A'^n)P^{-1}$;
- ▷ si A est une matrice carrée de taille $N+1$ telle que $(A^n)_{n\in\mathbf{N}^*}$ converge vers B, U_0 un vecteur colonne à $N+1$ composantes, et si l'on pose $U_n = A^nU_0$ pour tout $n \in \mathbf{N}^*$, alors les composantes de U_n convergent vers celles de BU_0 lorsque n tend vers $+\infty$.

On suppose dans cette partie que la population $\mathscr{P}$ est de taille N. Le modèle suivant est fondé sur l'hypothèse d'un virus V peu dangereux (la guérison est très rapide) mais très contagieux. On suppose que la propagation de V suit le schéma suivant : si l'on admet qu'à l'instant $n \in \mathbf{N}^*$, on a $X_n = i \in \{0,\ldots,N\}$ individus porteurs de V, et donc $N-i$ individus sains (c'est-à-dire non-porteurs de V), alors

- ▷ chacun des i porteurs devient sain à l'instant $n+1$;
- ▷ chacun des $N-i$ individus sains a une probabilité $p \in]0,1[$ (indépendante de n et de i) de devenir porteur de V, de façon indépendante les uns des autres.

Pour $(i,j) \in \{0,\ldots,N\}^2$, on note $q_{i,j}$ la probabilité conditionnelle $P(X_{n+1} = i | X_n = j)$ que $X_{n+1} = i$ sachant que $X_n = j$.

B.1. On traite ici, afin de se faire une idée du modèle, le cas $N = 2$. On pose pour tout $n \in \mathbf{N}$: $U_n = \begin{pmatrix} P(X_n = 0)\\ P(X_n = 1)\\ P(X_n = 2)\end{pmatrix}$, et on note M la matrice $M = (q_{i,j})_{\substack{0\leqslant i\leqslant 2\\ 0\leqslant j\leqslant 2}}$. On prendra garde au fait que M est une matrice 3×3 dont les lignes et les colonnes sont numérotées de 0 à 2.

a) Pour $j \in \{0,\ldots,N\}$, reconnaître la loi conditionnelle de X_{n+1} sachant $X_n = j$. En déduire que

$$M = \begin{pmatrix} (1-p)^2 & 1-p & 1\\ 2p(1-p) & p & 0\\ p^2 & 0 & 0\end{pmatrix}.$$

b) À l'aide de la formule des probabilités totales, montrer que $U_{n+1} = MU_n$ pour tout $n \in \mathbf{N}$. En déduire l'expression de U_n en fonction de U_0 et d'une puissance de la matrice M.

c) Montrer que 1, $-p$, p^2 sont les valeurs propres de M et déterminer des vecteurs propres associés.

d) Montrer que M est diagonalisable et donner une matrice inversible P et une matrice diagonale D telles que $M = PDP^{-1}$.

e) À l'aide des résultats mentionnés en tête de partie, montrer que pour $k \in \{0,1,2\}$, $\lim\limits_{n\to+\infty} P(X_n = k)$ existe (donner sa valeur). Comment interpréter ce résultat en terme de propagation du virus ?

B.2. La notation M est conservé. Dans le but de généraliser nos résultats à une valeur de N quelconque, on donne une interprétation de M comme matrice d'un endomorphisme d'un espace de polynômes. On note E l'espace vectoriel $\mathbf{R}_2[X]$ des polynômes à coefficients réels de degré inférieur ou égal à 2 et ϕ l'endomorphisme de E dont la matrice dans la base $(1, X, X^2)$ de E est M.

a) Montrer que pour $k \in \{0,1,2\}$, on a $\phi(X^k) = (pX+1-p)^{2-k}$.

b) À l'aide des calculs précédents, donner sous forme factorisée trois polynômes (P_0, P_1, P_2) tels que $\phi(P_k) = (-p)^kP_k$ pour $k \in \{0,1,2\}$.

B.3. On généralise maintenant une partie des résultats de la question B.1 à une valeur de $N \in \mathbf{N}^*$ quelconque. On pose pour tout $n \in \mathbf{N}$: $U_n = \begin{pmatrix} P(X_n = 0)\\ \vdots\\ P(X_n = N)\end{pmatrix}$, M la matrice $M = (q_{i,j})_{\substack{0\leqslant i\leqslant N\\ 0\leqslant j\leqslant N}}$. Dans l'espace vectoriel $E = \mathbf{R}_N[X]$ muni de sa base canonique $(1, X, \ldots, X^N)$, on considère l'endomorphisme ϕ de matrice M.

a) Donner une expression des coefficients de la matrice M (à l'aide de coefficients binomiaux), ainsi qu'une relation entre U_n et U_0 pour tout $n \in \mathbf{N}^*$.
b) Montrer que $\phi(X^k) = (pX + 1 - p)^{N-k}$ pour tout $k \in \{0, \dots, N\}$.
c) En s'inspirant des résultats des deux questions précédentes, quelle hypothèse peut-on faire quant aux valeurs propres de ϕ et aux vecteurs propres correspondants ?
d) Vérifier cette hypothèse par un calcul direct ainsi que du résultat de la question B.3.b.
e) En déduire que M est diagonalisable, et donner une matrice diagonale D telle qu'il existe une matrice P inversible vérifiant $M = PDP^{-1}$ (on ne demande pas d'expliciter P).
f) Donner, de façon directe, un vecteur propre de M associé à la valeur propre 1. On pourra calculer la somme de toutes les colonnes de M.
g) Montrer que $(M^n)_{n\in\mathbf{N}}$ est convergente.

Remarque finale : *un calcul explicite de P et de P^{-1}, par exemple, permettrait de montrer que les composantes U_n convergent, comme dans le cas $N = 2$.*