MATHEMATIQUES
Couples de variables aléatoires discrètes

*L'usage de la calculatrice **n'est pas autorisé** au cours de l'épreuve. Si vous repérez ce qui semble être une erreur d'énoncé, signalez-le sur la copie et poursuivez votre composition en expliquant les raisons des initiatives que vous avez été amené à prendre.*

EXERCICE :

On considère l'espace vectoriel $E = \mathbb{R}^3$ et f l'endomorphisme de E dont la matrice dans la base canonique $\mathcal{B} = (\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$ est la matrice A :

$$A = \begin{pmatrix} 3 & -2 & 3 \\ 1 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

1. Calcul des puissances de A

① Déterminer les valeurs propres λ_1 et λ_2 de l'endomorphisme f, avec $\lambda_1 < \lambda_2$.
Déterminer une base et la dimension de chacun des sous-espaces propres de f.

② La matrice A est-elle inversible ? (On ne demande pas la matrice A^{-1}).

③ Justifier que f n'est pas diagonalisable.

④ Soit u_1 un vecteur de E_{λ_1} de première composante égale à 1 et u_2 un vecteur de E_{λ_2} de seconde composante égale à 1.
Soit $u_3 = (1, 1, 1)$. Montrer que $\mathcal{C} = (u_1, u_2, u_3)$ est une base de E.

⑤ Déterminer la matrice de passage P de la la base $\mathcal{B}$ dans la base $\mathcal{C}$ puis la matrice de passage de la base $\mathcal{C}$ à la base $\mathcal{B}$.

⑥ Montrer que $f(u_3) = u_2 + 2u_3$ et que A est semblable à la matrice :

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

⑦ Prouver que pour tout élément n de $\mathbb{N}^*$ il existe un réel α_n tel que :

$$T^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^n & \alpha_n \\ 0 & 0 & 2^n \end{pmatrix}$$

On donnera le réel α_1 ainsi qu'une relation entre α_{n+1} et α_n

⑧ Montrer que :

$$\forall n \in \mathbb{N}^*,\ \alpha_n = n2^{n-1}$$

En déduire l'écriture matricielle de A^n en fonction de n.

PROBLEME :

Pour tout entier naturel n, I_n désigne l'ensemble des entiers naturels k tels que $0 \leq k \leq n$. On notera par la suite : $I_n = [\![0, n]\!]$ et $I_n^* = [\![1, n]\!]$.

Pour une variable aléatoire X à valeurs dans $\mathbb{N}$, on pose, pour tout réel t pour lequel cela a un sens,

$$g_X(t) = \mathbb{E}(t^X) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)t^k$$

où $\mathbb{E}$ désigne l'espérance. La fonction g_X est appelée *fonction génératrice* de la variable aléatoire X. Dans tout ce problème, on considérera que $0^0 = 1$, ce qui permet par exemple d'affirmer ici que $g_X(0) = \mathbb{P}(X = 0)$.

A. Fonction génératrice d'une variable à valeurs dans I_n

Soit n un entier naturel et X une variable aléatoire réelle à valeurs dans I_n. Pour tout $k \in I_n$, on note $a_k = \mathbb{P}(X = k)$ la probabilité que X prenne la valeur k.

① a) Montrer que g_X est une fonction polynôme à coefficients réels dont on précisera le degré. Quelle est la valeur de $g_X(1)$? Quelle est la valeur de $g'_X(1)$?

b) Montrer que si g_X est donnée, alors la loi de X est entièrement connue.

② Soient (m_1, m_2) deux entiers naturels et (Z_1, Z_2) un couple de variables aléatoires à valeurs respectivement dans I_{m_1} et I_{m_2}. On suppose que Z_1 et Z_2 sont indépendantes.

a) En utilisant pour tout réel t l'expression $g_X(t) = \mathbb{E}(t^X)$, montrer que

$$g_{Z_1+Z_2}(t) = g_{Z_1}(t)g_{Z_2}(t) \ (*)$$

b) On suppose que X est une variable aléatoire qui suit la loi binomiale de paramètres n et p avec $p > 0$. Montrer que sa fonction génératrice g_X est définie par : $g_X(t) = (pt+q)^n \ \forall t \in \mathbb{R}$ où on a posé $q = 1 - p$.
Retrouver par le calcul l'expression de $\mathbb{E}(X)$.

c) Soit Y une variable aléatoire réelle suivant la loi binomiale de paramètres n' et p avec n' un entier naturel. On suppose que X et Y sont indépendantes. Montrer en utilisant A.2.a) et A.1.b) que $X + Y$ suit la loi binomiale de paramètres $n + n'$, p.

③ **Exemple** (cette question est sans incidence sur la suite du problème) On lance deux dés classiques à 6 faces supposés non équilibrés. Le résultat est affichés par le i-ième dé est une variable aléatoire à valeurs dans $I_6^* = [\![1, 6]\!]$. Les variables aléatoires X_1 et X_2 sont supposées indépendantes.
On note pour tout $k \in I_6^*$, $p_k = \mathbb{P}(X_1 = k)$ et $q_k = \mathbb{P}(X_2 = k)$ et on suppose que $p_k > 0$ et $q_k > 0$. On cherche à prouver qu'il est impossible de truquer les deux dés pour que $Y = X_1 + X_2$ soit une variable aléatoire uniforme sur $I' = [\![2, 12]\!]$. Nous raisonnons par l'absurde : on suppose donc dans la suite que Y suit la loi uniforme sur I'.

a) Donner la fonction génératrice de Y qu'on notera R.

b) Vérifier que les racines complexes non réelles de R sont les nombres complexes $e^{i\frac{2k\pi}{11}}$ pour k entier entre 1 et 10.

c) Déduire de la relation (*) du A.2.a) qu'il existe deux polynômes P et Q de degré 5 à coefficients réels tels que pour tout $t \in \mathbb{R}$, $t^2P(t)Q(t) = R(t)$.

d) Montrer une contradiction quant à l'existence de racines réelles de P et de Q. Conclure.

B. Fonction génératrice d'une variable à valeurs dans $\mathbb{N}$

Soit maintenant X une variable aléatoire réelle à valeurs dans $\mathbb{N}$. Pour tout $n \in \mathbb{N}$, on pose

$$a_n = \mathbb{P}(X = n)$$

① Montrer que pour tout $t \in [-1, 1]$, la série $\sum a_n t^n$ est absolument convergente. En déduire que g_X est définie sur $[-1, 1]$ et donner la valeur de $g_X(1)$.

② Soit Y une variable aléatoire réelle à valeurs dans $\mathbb{N}$. Montrer que si X et Y sont indépendantes, alors pour tout $t \in [-1, 1]$, $g_{X+Y}(t) = g_X(t)g_Y(t)$.

③ a) On suppose que X suit une loi géométrique de paramètre $p \in]0, 1[$ à valeurs dans $\mathbb{N}^*$. Calculer $g_X(t)$ pour $t \in [-1, 1]$ (on pourra poser $q = 1 - p$).

b) Même question pour X suivant une loi de Poisson de paramètre $\lambda > 0$.

c) Vérifier dans chacun des deux cas qui précède que g_X est dérivable sur $[-1, 1]$ et que $g'_X(1) = \mathbb{E}(X)$.

C. Généralisation

— Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$, indépendantes et de même loi, dont la fonction génératrice commune est notée f.

— Soit N une variable aléatoire à valeurs dans $\mathbb{N}$, indépendante des variables $(X_n)_{n\geq 1}$ et dont la fonction génératrice est notée h

On pose $S_0 = 0$ et, pour tout $n \in \mathbb{N}^*$, $S_n = \displaystyle\sum_{k=1}^{n} X_k$.

On définit alors $Y = S_N$ (il s'agit donc de la somme d'un nombre aléatoire de variables aléatoires). On admettra que Y est bien un variable aléatoire à valeurs dans $\mathbb{N}$ et on notera g sa fonction génératrice. Enfin on notera ψ_n la fonction génératrice de S_n pour tout $n \geq 0$. On cherche maintenant à déterminer g en fonction de f et h.

On se limitera ici au cas où N prend ses valeurs dans I_s, s étant un entier naturel supérieur à 1, fixé dans toute cette partie. Soit $t \in [-1, 1]$.

① Montrer que, pour tout $n \in I_s$, on a $\psi_n(t) = (f(t))^n$.

② Montrer que, pour tout entier naturel k :

$$\mathbb{P}(Y = k) = \sum_{n=0}^{s} \mathbb{P}((Y = k) \cap (N = n)) = \sum_{n=0}^{s} \mathbb{P}((S_n = k) \cap (N = n))$$

③ En déduire l'égalité :

$$g(t)) = \sum_{k=0}^{\infty} \left(\sum_{n=0}^{s} \mathbb{P}(S_n = k)\mathbb{P}(N = n) \right) t^k$$

④ En déduire que $g(t) = (h \circ f)(t)$.

On admet pour la suite du problème que le résultat obtenu dans la question C.4. est encore valable dans le cas général, c'est-à-dire dans le cas où la variable aléatoire N est à valeurs dans $\mathbb{N}$.

⑤ On suppose ici que les $(X_n)_{n\geq 1}$ suivent la loi géométrique de paramètre $p \in]0,1[$ à valeurs dans $\mathbb{N}^*$ et N la loi géométrique de paramètre $p' \in]0,1[$ à valeurs dans $\mathbb{N}^*$. Montrer que Y suit une loi géométrique dont on précisera le paramètre. <Montrer comment il est possible de la modéliser avec Python.

D. Multiplication d'une bactérie

Une bactérie B est présente dans un milieu M plus ou moins propice à sa reproduction. Elle se reproduit de la façon suivante : chaque individu donne naissance à X nouvelles bactéries B (appelées dans la suite « fils ») puis meurt. On peut donc classer les bactéries par génération : les bactéries d'une génération vont chacune donner naissance à un certain nombre de fils puis disparaître. Les fils de toutes les bactéries de la génération n formeront ainsi la génération $n+1$.
Le but est de déterminer la probabilité que toutes les bactéries B disparaissent du milieu M au bout d'un certain nombre de générations.

Pour $n \in N$, on notera dans la suite Y_n le nombre d'individus formant la génération n de bactéries B présentes dans M (on a donc d'après les hypothèses $Y_0 = 1$ car la génération $n = 0$ ne compte qu'une seule bactérie B). On notera de plus $x_n = \mathbb{P}(Y_n = 0)$ (on a donc $x_0 = 0$).

On admet que :
— les variables aléatoires comptant le nombre de « fils » de chaque bactérie B présente à une génération n donnée sont des variables de même loi. Ces variables sont aussi indépendantes de Y_n.
— X (nombre de « fils » d'une bactérie B fixée, quelle que soit sa génération) suit la loi de Poisson de paramètre $\lambda > 0$, et on notera f la fonction génératrice de X (on admet que pour tout $t \in [-1,1]$, $f(t) = e^{\lambda(t-1)}$) ;

1. Étudier les variations de la suite $(x_n)_{n\in\mathbb{N}}$ et en déduire sa convergence.
   On admettra dans la suite que la probabilité p que la bactérie B disparaisse du milieu M est $p = \lim\limits_{n\to\infty} x_n$.

2. Donner la loi de Y_1 et en déduire que $x_1 = f(x_0)$.

3. Justifier que $Y_2 = \displaystyle\sum_{k=1}^{Y_1} X_k$ où les variables aléatoires $(X_k)_{1\leq k\leq Y_1}$ sont indépendantes et de même loi que X. Déduire du résultat admis en C. la fonction génératrice de Y_2 qu'on exprimera en fonction de f.

4. a) Pour tout $n \geq 1$, montrer que la fonction génératrice de Y_n vérifie : $g_{Y_n} = f^n$ (composée n-ième de f) et en déduire que $x_n = f(x_{n-1})$.
   b) En déduire que $p = f(p)$.

Soit φ la fonction définie sur $[0,1]$ par $\varphi(t) = f(t) - t$.

5. On suppose $\lambda \leq 1$. Etudier les variations de φ sur $[0,1]$ et en déduire que la bactérie B disparaît du milieu M de façon presque certaine.

6. On suppose maintenant que $\lambda > 1$.
   a) Etudier les variations de la fonction $\theta : u \longmapsto 1 - \dfrac{lnu}{u}$ sur $]1, +\infty[$ et en déduire successivement que : $\forall u > 1$, $lnu < u$ et $ue^{-u} - 1 < 0$.
   b) En déduire qu'il existe $\beta \in]0,1[$ tel que $\varphi'(\beta) = 0$. Déterminer les variations de φ sur $[0,1]$.
   c) Montrer qu'il existe un unique $\alpha \in]0,1[$ tel que $f(\alpha) = \alpha$.
   d) Montrer que pour tout $n \in \mathbb{N}$, $x_n \leq \alpha$ et conclure quant à la probabilité de disparition de la bactérie B du milieu M.
7. a) Écrire une fonction Python `poisson(lbd)` de paramètre d'entrée le paramètre $\lambda > 0$ qui simule la réalisation d'une loi de poisson de paramètre λ.
   b) Écrire une fonction `tempsExtinction(lbd,nBMax)` qui retourne la première génération qui voit disparaître la bactérie si c'est le cas et la valeur nulle si Y dépasse une valeur entière `nBMax` de bactéries au-delà de laquelle on peut considérer que la population ne s'éteindra plus (par exemple `nBMax=100` semble raisonnable).

   Par exemple, on notant `LY` la liste des valeurs prises par Y et en prenant `nBMax=100` :
   — Pour $\lambda = 0.8$, si `LY=[1,0]` alors on retourne 1.
   — Pour $\lambda = 0.8$, si `LY=[[1, 2, 2, 4, 3, 1, 1, 0]]` alors on retourne 7.
   — Pour $\lambda = 1.4$, si `LY=[1, 1, 1, 2, 0]` alors on retourne 4.
   — Pour $\lambda = 1.4$, si `LY=[1, 1, 1, 2, 2, 5, 12, 14, 20, 21, 37, 47, 71, 99, 136]` alors on retourne 0.
   c) Ecrire une fonction permettant de valider les réponses fournies en 5. et 6.d).