

**Variables aléatoires discrètes et Algèbre
linéaire**

Le sujet se compose de deux exercices et d'un problème dont les deux parties sont indépendantes. On prendra soin de lire l'ensemble du sujet avant de commencer à composer et, à titre indicatif, on pourra consacrer une heure aux exercices et deux heures au problème.

Il sera tenu compte de la présentation et en particulier de l'encadrement des résultats.

L'usage de la calculatrice n'est pas autorisé au cours de l'épreuve.

Exercice 1 :

Dans cet exercice, n désigne un entier naturel non nul. On dispose d'une pièce dont la probabilité de faire "pile" est $p \in]0; 1[$ et de $(n + 1)$ urnes numérotées de 0 à n .

Pour $k \in \llbracket 0; n \rrbracket$, l'urne n° k contient k boules vertes et $(n - k)$ boules rouges.

On considère l'expérience \mathcal{E} suivante : on lance n fois la pièce, puis on pioche une unique boule dans l'urne dont le numéro correspond au nombre de fois où "pile" a été obtenu.

Par exemple, si on a obtenu quatre "piles" au cours de ces n lancers, on pioche dans l'urne n°4.

On note X la variable aléatoire correspondant au nombre de "piles" obtenues lors des n lancers et Y la variable aléatoire qui vaut 1 si l'on tire une boule verte et 0 sinon.

- ① a) Reconnaître la loi de probabilité de la variable aléatoire X .
On précisera en particulier $X(\Omega)$ et $P(X = k)$ pour tout k de $X(\Omega)$.
Donner l'espérance $E(X)$ et la variance $V(X)$.
- b) En utilisant la formule de Koenig-Huygens, calculer la valeur de $E(X^2)$.
- c) Calculer $P_{(X=0)}(Y = 0)$ et $P_{(X=n)}(Y = 0)$. X et Y sont-elles indépendantes ?
- d) Justifier que, pour tout $k \in \llbracket 0; n \rrbracket$, $P_{(X=k)}(Y = 1) = \frac{k}{n}$.
- e) En déduire, en utilisant le système complet d'événements $(X = k)_{0 \leq k \leq n}$, que :

$$P(Y = 1) = \frac{E(X)}{n}$$

f) Donner la loi de Y et son espérance.

g) Montrer que $E(XY) = \sum_{k=1}^n kP(X = k \cap Y = 1) = \frac{E(X^2)}{n}$.

h) En déduire la covariance du couple (X, Y) .

Exercice 2 :

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

$$g_X(t) = \mathbb{E}(t^X) = \sum_{k \in X(\Omega)} t^k \mathbb{P}(X = k) \text{ (d'après le théorème de Transfert)}$$

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

- ① **La cas des variables aléatoires discrètes finies.** Soient n un entier naturel non nul et X une variable aléatoire réelle à valeurs dans $I_n = \llbracket 0, n \rrbracket$ (appelé support de X).
 Pour tout $k \in I_n$, on pourra noter $a_k = \mathbb{P}(X = k)$ la probabilité que X prenne la valeur k .
- Justifier que g_X est une fonction polynôme dont on précisera le degré. Quelle est la valeur de $g_X(1)$?
 - Justifier la dérivabilité première et seconde de g_X sur \mathbb{R} et montrer que $g'_X(1) = \mathbb{E}(X)$, $g''_X(1) = \mathbb{E}(X(X-1))$. Que vaut $\mathbb{V}(X)$ en fonction de g_X et de ses dérivées ?
 - Montrer que la fonction génératrice d'une variable aléatoire X qui suit la loi binomiale de paramètres n et p avec $0 < p < 1$ est : $g_X(t) = (pt + q)^n$.
 Retrouver grâce à elle les valeurs de l'espérance et de la variance d'une loi binomiale.
- ② **La cas des variables aléatoires discrètes infinies.** si X est une variable aléatoire à valeurs dans \mathbb{N} , nous admettrons que la connaissance de $g_X(t)$ pour tout $t \in [-1, 1]$ caractérise la loi de X (cela permet donc de reconnaître la loi d'une variable aléatoire connaissant sa fonction génératrice).
 Par ailleurs, nous admettrons que si g_X est dérivable en 1, alors comme dans la première question $\mathbb{E}(X)$ existe et vaut $\mathbb{E}(X) = g'_X(1)$.
- Montrer que pour tout $t \in [-1, 1]$, la série $\sum_{n \geq 0} \mathbb{P}(X = n)t^n$ est absolument convergente. En déduire que g_X est définie sur $[-1, 1]$ et donner la valeur de $g_X(1)$.
 - Soit Y une variable aléatoire réelle à valeurs dans \mathbb{N} .
 En utilisant pour tout réel t l'expression $g_X(t) = \mathbb{E}(t^X)$, montrer que si X et Y sont indépendantes, alors pour tout $t \in [-1, 1]$, $g_{X+Y}(t) = g_X(t)g_Y(t)$.
 - On suppose que X suit une loi géométrique de paramètre $p \in]0, 1[$, à valeurs dans \mathbb{N}^* . Montrer que $g_X(t) = \frac{pt}{1-qt}$ pour $t \in [-1, 1]$ (où $q = 1-p$). En déduire l'existence et la valeur de $\mathbb{E}(X)$.
Retrouver ce résultat par la méthode de votre choix.
 - On suppose ici que X suit une loi de Poisson de paramètre $\lambda > 0$.
 - Calculer $g_X(t)$ pour $t \in [-1, 1]$ et retrouver l'existence et la valeur de $\mathbb{E}(X)$.
 - Si X et Y sont deux variables indépendantes qui suivent respectivement une loi de Poisson de paramètre λ et μ . Déterminer $g_{X+Y}(t)$ et en déduire que $X + Y$ suit une loi de Poisson dont on précisera le paramètre.
Retrouver ce résultat par une autre méthode de votre choix.

Problème :

Une éprouvette contient 10 bactéries, 4 sont des bactéries de type A et 6 de type B . On les laisse se reproduire en milliers d'exemplaires, la proportion de bactéries de chaque type restant inchangées. On prélève alors, au hasard, 10 bactéries que l'on met dans une autre éprouvette. On les laisse se reproduire en milliers d'exemplaires dans les mêmes conditions que précédemment, et on recommence l'expérience.
 Que se passe-t-il après un grand nombre d'expériences ?

L'énoncé théorique ci-dessous propose un modèle probabiliste pour répondre à cette question.

Définition. On note N un entier supérieur ou égal à 2 et k_0 un entier de $\llbracket 0, N \rrbracket$. On pose $p = \frac{k_0}{N}$ et $q = 1-p$. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs dans $\llbracket 0, N \rrbracket$ dont les lois de probabilité sont définies de la manière suivante :

- X_0 est la variable certaine égale à k_0
- X_1 suit la loi binomiale de paramètres N et p (Par convention, on dit que la loi binomiale de paramètres N et 0 est la loi de la variable aléatoire certaine égale à 0 et que la loi binomiale de paramètres N et 1 est la loi de la variable aléatoire certaine égale à N)

- Pour tout entier n non nul et tout entier k de $\llbracket 0, N \rrbracket$ tel que $\mathbb{P}(X_n = k) \neq 0$, la loi conditionnelle de X_{n+1} sachant $(X_n = k)$ est la loi binomiale de paramètres N et $\frac{k}{N}$. En d'autres termes :
Pour tout i de $\llbracket 0, N \rrbracket$

$$\mathbb{P}_{(X_n=k)}(X_{n+1} = i) = \binom{N}{i} \left(\frac{k}{N}\right)^i \left(1 - \frac{k}{N}\right)^{N-i} \quad (\text{avec la convention habituelle } 0^0 = 1)$$

- On définit la suite de variable aléatoires $(F_n)_{n \geq 0}$ par $F_n = \frac{X_n}{N}$

Remarque préliminaire

Dans l'exemple introductif, en appelant N le nombre de bactéries prélevées à chaque expérience, k_0 le nombre de bactéries de type A dans la première éprouvette au début de la première expérience et n le numéro de l'expérience, X_n désigne le nombre de bactéries de type A à l'issue du n -ième prélèvement. On a vu qu'il s'agit d'un tirage par poignée de N bactéries mais ce nombre est suffisamment faible au regard du nombre de bactéries dans l'éprouvette (de l'ordre du millier) pour qu'on considère que l'expérience est équivalente à un tirage successifs avec remise des N bactéries prélevées. Il est donc légitime de considérer que X_{n+1} , conditionnée par $(X_n = k)$, suit une loi binomiale de paramètres N et $\frac{k}{N}$ puisqu'on dénombre les succès au cours de N épreuves de Bernoulli « indépendantes » de même paramètre égale à la proportion k/N de bactéries de type A dans l'éprouvette au moment du prélèvement.

Partie 1

Dans cette partie, $N = 3$.

- ① Que dire de la suite $(X_n)_{n \geq 0}$ si $k_0 = 0$? Si $k_0 = 3$?

On suppose désormais, dans la suite de cette partie que $k_0 = 1$.

- ② Etant donné un entier n , déterminer $\mathbb{P}(X_{n+1} = 0)$ en fonction des valeurs de $\mathbb{P}(X_n = k)$, $k \in \llbracket 0, 3 \rrbracket$
Déterminer de même $\mathbb{P}(X_{n+1} = 1)$, $\mathbb{P}(X_{n+1} = 2)$ et $\mathbb{P}(X_{n+1} = 3)$

$$\forall n \in \mathbb{N}, \text{ on pose } U_n = \begin{pmatrix} p(X_n = 0) \\ p(X_n = 1) \\ p(X_n = 2) \\ p(X_n = 3) \end{pmatrix}. \text{ Vérifier que } U_{n+1} = AU_n \text{ où } A = \begin{pmatrix} 1 & \frac{8}{27} & \frac{1}{27} & 0 \\ 0 & \frac{9}{9} & \frac{2}{9} & 0 \\ 0 & \frac{2}{9} & \frac{4}{9} & 0 \\ 0 & \frac{1}{27} & \frac{8}{27} & 1 \end{pmatrix}$$

- ③ a) Soit la matrice ligne $V = (0 \ 1 \ 2 \ 3)$. Calculer VA .
b) Montrer que $\mathbb{E}(X_n) = VU_n$ pour tout entier n . En déduire la valeur de $\mathbb{E}(X_n)$
- ④ a) Soit la matrice ligne $W = (0 \ 2 \ 2 \ 0)$. Calculer WA
b) Soit n un entier naturel. Montrer que $\mathbb{E}(X_n(3 - X_n)) = WU_n$.
En reconnaissant une suite usuelle, en déduire que $\mathbb{E}(X_n(3 - X_n)) = 2 \left(\frac{2}{3}\right)^n$
- c) En déduire $\forall n \in \mathbb{N} \ \mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2) = \left(\frac{2}{3}\right)^n$.
- d) A l'aide de 3. et de 4.c, déterminer $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = i)$ pour tout entier i de $\llbracket 0, 3 \rrbracket$.

- e) En déduire $\lim_{n \rightarrow \infty} \mathbb{P}(F_n = i)$ pour tout $i \in \llbracket 0, 3 \rrbracket$. On dira que la suite $(F_n)_{n \geq 0}$ converge vers une loi limite qu'on précisera. Interpréter.
- f) Déterminer le premier entier n tel que $\mathbb{P}((F_n = 0) \cup (F_n = 1)) > 0,95$. Interpréter.

⑤ a) On pose $Y_2 = \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}$ et $Y_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$

Montrer que Y_2 et Y_3 sont des vecteurs propres de A . Quelles sont les valeurs propres associées ?

- b) Montrer que 1 est valeur propre de A et conclure que A est diagonalisable. Donner ses valeurs propres ainsi qu'une base de vecteurs propres.

En déduire qu'il existe une matrice P inversible telle que $P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{2}{9} & 0 & 0 \\ 0 & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

- c) Calculer A^n pour tout entier n .

- ⑥ a) Préciser U_0 et montrer que la loi de X_n est donnée par :

$$\begin{aligned} p(X_n = 0) &= \frac{2}{3} - \frac{1}{2} \left(\frac{2}{3}\right)^n - \frac{1}{6} \left(\frac{2}{9}\right)^n & p(X_n = 1) &= \frac{1}{2} \left(\left(\frac{2}{3}\right)^n + \left(\frac{2}{9}\right)^n \right) \\ p(X_n = 2) &= \frac{1}{2} \left(\left(\frac{2}{3}\right)^n - \left(\frac{2}{9}\right)^n \right) & p(X_n = 3) &= \frac{1}{3} - \frac{1}{2} \left(\frac{2}{3}\right)^n + \frac{1}{6} \left(\frac{2}{9}\right)^n \end{aligned}$$

- b) Retrouver le résultat du 4.e).

- ⑦ On définit la variable aléatoire T par :

- si pour tout entier n , $(X_n \neq 0)$ et $(X_n \neq 3)$, alors $(T = 0)$ est réalisé.
- sinon, $(T = n)$ est réalisé où n est le plus petit entier k tel que $(X_k = 0)$ ou $(X_k = 3)$

On pose pour tout entier n , $v_n = \mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2)$

- a) Montrer que pour tout entier n non nul, $\mathbb{P}(T = n) = v_{n-1} - v_n$.
- b) Exprimer v_n en utilisant la question 6. et reconnaître la loi de T .
- c) Préciser $\mathbb{P}(T = 0)$.

Donner $\mathbb{E}(T)$ et interpréter quant à la composition des éprouvettes.

- d) Ecrire une fonction Python `simuLT(k0, N)` d'arguments le nombre initial k_0 de bactéries de type A et le nombre N de bactéries prélevées à chaque tirage et qui retourne une simulation de la variable aléatoire T .

Utiliser cette fonction pour écrire une fonction permettant de conforter la valeur de $\mathbb{E}(T)$ obtenue en 7.c).

Partie 2

Dans cette partie, $N \geq 2$ est quelconque et $k_0 \in \llbracket 1, N - 1 \rrbracket$.

On pose pour tout $n \in \mathbb{N}$: $u_n = \mathbb{P}(X_n = 0) + \mathbb{P}(X_n = N)$ et $v_n = 1 - u_n$.

- ① Montrer que pour tout $i \in \llbracket 0, N \rrbracket$, $\mathbb{P}(X_{n+1} = i) = \sum_{k=0}^N \mathbb{P}(X_n = k) \binom{N}{i} \left(\frac{k}{N}\right)^i \left(1 - \frac{k}{N}\right)^{N-i}$.
- ② Montrer que pour tout entier n , $\mathbb{E}(X_{n+1}) = \mathbb{E}(X_n)$.

- ③ a) Montrer que pour tout $n \in \mathbb{N}$, $\mathbb{E}(X_{n+1}(N - X_{n+1})) = \frac{N-1}{N} \mathbb{E}(X_n(N - X_n))$.
- b) En déduire la valeur de $\mathbb{E}(X_n(N - X_n))$ en fonction de n , N et k_0 .
- ④ Montrer que la suite $(u_n)_{n \in \mathbb{N}}$ est croissante et convergente.
- ⑤ a) Soit X une variable aléatoire à valeurs dans \mathbb{R}_+ prenant un nombre fini de valeurs. Rappeler l'inégalité de Markov.
- b) Étudier sur $[1, N-1]$ la fonction f définie par : $\forall x \in [1, N-1]$, $f(x) = x(N-x)$.
- c) En utilisant $\mathbb{E}(X_n(N - X_n))$, montrer que : $\forall n \in \mathbb{N}$, $0 \leq v_n \leq pq \frac{N^2}{N-1} \left(1 - \frac{1}{N}\right)^n$.
- ⑥ a) Quelle est la limite de la suite $(v_n)_{n \in \mathbb{N}}$?
- b) En déduire que pour tout entier $k \in \llbracket 1, N-1 \rrbracket$, $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = 0$.
- c) En utilisant la question 2., montrer que $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = N) = p$. Déterminer $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0)$.
- d) Montrer que la suite $(F_n)_{n \geq 0}$ converge vers une loi de probabilité qu'on déterminera.