

CONCOURS INTERNE D'ATTACHÉ STATISTICIEN DE L'INSEE

ANNÉE 2025

Épreuve de Mathématiques et Statistiques

Durée : 4 heures

Coefficient 3

Le sujet comporte 7 pages (y compris celle-ci)

- Tous documents et appareils électroniques sont interdits.
- L'épreuve est constituée de deux parties indépendantes.
- Il sera tenu compte de la présentation, de la pertinence et de la clarté des justifications.

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

Partie A

Soit n un entier naturel non nul. Une matrice $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{R})$ est dite **stochastique** si elle vérifie :

1. $\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket, a_{i,j} \geq 0$

2. $\forall i \in \llbracket 1, n \rrbracket, \sum_{j=1}^p a_{i,j} = 1$

Le but de cette partie est d'établir quelques propriétés des matrices stochastiques.

1. Pour chacune des matrices suivantes, préciser si elle est ou non stochastique, en justifiant la réponse :

- (a) $A_1 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 2 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$

- (b) $A_2 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{3}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$

- (c) $A_3 = \begin{pmatrix} 0,2 & 0,3 & 0,5 \\ 0 & 0,2 & 0,8 \\ 0,1 & 0,6 & 0,3 \end{pmatrix}$

2. Démontrer que si $\mathbf{p}$ est un vecteur-ligne stochastique de taille n (ou matrice-ligne stochastique de $\mathcal{M}_{1,n}(\mathbb{R})$) et $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice stochastique, alors $\mathbf{p}A$ est un vecteur-ligne stochastique de taille n .
3. Démontrer que si $A, B \in \mathcal{M}_n(\mathbb{R})$ sont deux matrices stochastiques, alors leur produit AB est une matrice stochastique.
4. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice stochastique. On note $Sp(A)$ l'ensemble de ses valeurs propres (spectre de A).

- (a) Démontrer que $1 \in Sp(A)$ et déterminer un vecteur propre u de A associé à cette valeur propre.

- (b) Démontrer que pour toute valeur propre réelle ou complexe λ de A , on a $|\lambda| \leq 1$.

- (c) Soit $A = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix}$, avec $0 \leq p, q \leq 1$. Démontrer que $p - q$ est une valeur propre de A . La matrice A est-elle diagonalisable ?

- (d) La matrice $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$ est-elle diagonalisable ?

- (e) Quelle conclusion pouvez-vous en tirer ?

Partie B

Une puce se déplace aléatoirement entre trois positions différentes notées 1, 2 et 3. On note $X_n \in \{1, 2, 3\}$ sa position à la date $n \in \mathbb{N}$, et on suppose que pour tout couple d'entiers $(i, j) \in \llbracket 1, 3 \rrbracket^2$, la probabilité conditionnelle $p_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i)$ est égale au coefficient situé à la ligne numéro i et la colonne numéro j de la matrice

$$\mathbf{P} = \begin{pmatrix} 0 & p & 1-p \\ p & 0 & 1-p \\ p & 1-p & 0 \end{pmatrix}$$

où $p \in [0; 1]$ est fixé.

Pour tout entier naturel n , on note de plus $\pi_n = (\pi_n(1), \pi_n(2), \pi_n(3))$, avec $\pi_n(i) = \mathbb{P}(X_n = i)$ pour $i \in \llbracket 1, 3 \rrbracket$. Le vecteur π_n représente donc la loi de probabilité de la variable X_n . Pour $n = 0$, on note plus simplement $\pi_0 = (p_1, p_2, p_3)$.

5. On suppose, seulement dans cette question, que $p = 0$.

- (a) Calculer $\mathbb{P}(X_n = 1)$.
- (b) Calculer $\mathbb{P}(X_n = 2)$ (distinguer les cas n pair et n impair).
- (c) Calculer $\mathbb{P}(X_n = 3)$ (distinguer les cas n pair et n impair).

6. On rappelle que pour tout couple d'entiers $(i, j) \in \llbracket 1, 3 \rrbracket^2$, on pose

$$p_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i)$$

- (a) Démontrer que pour tout entier naturel n et pour tout j dans $\llbracket 1, 3 \rrbracket$:

$$\pi_{n+1}(j) = \sum_{i=1}^3 \pi_n(i) p_{i,j}$$

- (b) En déduire que pour tout entier naturel n , $\pi_{n+1} = \pi_n \mathbf{P}$, puis que $\pi_n = \pi_0 \mathbf{P}^n$.

7. On s'intéresse dans cette question au comportement asymptotique de la suite $(\pi_n)_{n \in \mathbb{N}}$.
On dit qu'une suite $(A_n)_{n \in \mathbb{N}}$ de matrices de $\mathcal{M}_3(\mathbb{R})$ converge vers une matrice A de $\mathcal{M}_3(\mathbb{R})$ si chaque coefficient de A_n a pour limite le coefficient correspondant de A .

- (a) Démontrer que si la suite $(\mathbf{P}^n)_{n \in \mathbb{N}}$ converge, alors la suite $(\pi_n)_{n \in \mathbb{N}}$ converge également.
- (b) En supposant que cette condition soit vérifiée, et en notant π et $\tilde{\mathbf{P}}$ les limites respectives des suites $(\pi_n)_{n \in \mathbb{N}}$ et $(\mathbf{P}^n)_{n \in \mathbb{N}}$, démontrer les égalités

$$\pi = \pi_0 \tilde{\mathbf{P}}$$

$$\pi = \pi \mathbf{P}$$

- (c) En déduire un vecteur propre de ${}^t \mathbf{P}$ et préciser sa valeur propre associée.
- (d) Comment interpréter le vecteur π de la question précédente ?

8. On suppose dans cette question que $p = \frac{1}{2}$.

- (a) Démontrer que $Sp(\mathbf{P}) = \left\{ 1, -\frac{1}{2} \right\}$.
- (b) La matrice $\mathbf{P}$ est-elle diagonalisable ? Justifier.
- (c) Déterminer les sous-espaces propres associés aux valeurs propres obtenues à la question (a).

- (d) On note $Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$. En utilisant la question précédente, justifier sans faire de calculs que Q est inversible et que $Q^{-1}\mathbf{P}Q$ est une matrice diagonale que l'on explicitera.
- (e) Calculer Q^{-1} .
- (f) Pour n entier naturel, calculer $\mathbf{P}^n$ et en déduire $\pi_n(i)$ pour tout $1 \leq i \leq 3$.
- (g) La suite $(\pi_n)_{n \in \mathbb{N}}$ est-elle convergente? Si oui, quelle est sa limite et comment l'interpréter?

9. On suppose maintenant que $p \notin \left\{0, \frac{1}{2}, 1\right\}$.

- (a) Démontrer que $Sp(\mathbf{P}) = \{1, p-1, -p\}$.
- (b) Déterminer les sous-espaces propres associés aux valeurs propres obtenues à la question précédente.
- (c) On note $Q = \begin{pmatrix} 1 & p-1 & 1 \\ 1 & p-1 & -p \\ 1 & 1 & -p \end{pmatrix}$. En utilisant la question précédente, justifier sans faire de calculs que Q est inversible et que $Q^{-1}\mathbf{P}Q$ est une matrice diagonale que l'on explicitera.
- (d) On pose

$$R = \begin{pmatrix} 2p-p^2 & p^2-p+1 & 1-p^2 \\ 0 & -1-p & 1+p \\ 2-p & -2+p & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Démontrer l'égalité $QR = (1+p)(2-p)I_3$, où I_3 est la matrice identité de $\mathcal{M}_3(\mathbb{R})$.

- (e) En déduire Q^{-1} .
- (f) Démontrer que la suite $(\pi_n)_{n \in \mathbb{N}}$ converge et calculer sa limite. Cette limite dépend-elle de la loi initiale π_0 ?
10. Soit n un entier naturel. Démontrer, pour $i_0, i_1, \dots, i_n$ dans $\llbracket 1, 3 \rrbracket$, l'égalité suivante :

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p_{i_0} p_{i_0, i_1} \dots p_{i_{n-1}, i_n}$$

11. On s'intéresse dans cette question au temps d'attente avant que la puce n'atteigne la position 2 pour la première fois, autrement dit à la variable

$$T = \inf \{n \in \mathbb{N}, X_n = 2\}$$

- (a) Déterminer la loi de T lorsque $p = 0$. Démontrer que dans ce cas, $\mathbb{E}(T) = 2p_1 + p_3$.
Dans ce qui suit, on suppose désormais que $0 < p < 1$ et que $\pi_0 = (1, 0, 0)$.
- (b) Déterminer $\mathbb{P}(T = 0)$.
- (c) Pour $n \in \mathbb{N}^*$ calculer $\mathbb{P}(T = 2n)$.
- (d) Pour $n \in \mathbb{N}$, calculer $\mathbb{P}(T = 2n + 1)$.
- (e) Quelle est la probabilité que T soit pair? Quelle est la probabilité que T soit impair?
- (f) Calculer le temps d'attente moyen $\mathbb{E}(T)$.

Partie A - La mantisse entière

On définit la fonction "logarithme décimal" par la relation : $\forall x > 0, \log x = \frac{\ln x}{\ln 10}$, où $\ln$ désigne la fonction "logarithme népérien".

Pour x réel, $\lfloor x \rfloor$ désigne la partie entière de x , c'est à dire le plus grand entier relatif plus petit ou égal à x ; $\{x\}$ est sa partie fractionnaire définie par $\{x\} = x - \lfloor x \rfloor$.

Soient x et y deux réels strictement positifs. Montrer que :

1. $10^{\log x} = x$.
2. $\log xy = \log x + \log y$.
3. $\log 10^x = x$.
4. $1 \leq 10^{\{\log x\}} < 10$.
5. $x = 10^{\{\log x\}} \cdot 10^{\lfloor \log x \rfloor}$.
6. Supposons que $x = s \cdot 10^p$ avec $s \in [1, 10[$ et $p \in \mathbb{Z}$.
 - (a) Montrer que $\log s = \{\log x\}$.
 - (b) Montrer que $p = \lfloor \log x \rfloor$.
7. Supposons que $x = 0,0031415927$. Calculer $\lfloor \log x \rfloor$ et $10^{\{\log x\}}$.

On appelle $me(x) = \lfloor 10^{\{\log x\}} \rfloor$ la **mantisse entière** de x , ou premier chiffre significatif de x .

Soient x un réel strictement positif et $n \in \mathbb{N}$, on pose :

- $p_n = \lfloor 10^n x \rfloor$,
 - $a_n = p_n \cdot 10^{-n}$,
 - $b_n = (p_n + 1) \cdot 10^{-n}$,
 - $c_n = p_n - 10 \cdot p_{n-1}$, si $n \geq 1$.
8. Si $x = 0,0031415927$, calculer p_5, a_5, b_5, c_5 .
 9. Montrer que $c_n \in \{0 \dots 9\}$.
 10. Montrer que $a_n \leq x \leq b_n$.
 11. Montrer que les suites $(a_n)_n$ et $(b_n)_n$ sont adjacentes.
 12. Montrer que pour $n \geq 1$: $a_n = a_0 + \sum_{k=1}^n c_k 10^{-k}$.
 13. Exprimer x en fonction de a_0 et de la suite $(c_k)_k$.
 14. Soit $(d_k)_k$ une suite non nulle telle que $\forall k \in \mathbb{N}, d_k \in \{0 \dots 9\}$.
 - (a) Montrer que la série $\sum_{k \geq 0} \frac{d_k}{10^k}$ converge vers un réel z .
 - (b) Exprimer $me(z)$ en fonction de la suite $(d_k)_k$.

Partie B - Quelques résultats intermédiaires

Considérons une variable aléatoire réelle X (dont on dira qu'elle suit une loi de Benford) qui prend ses valeurs dans l'intervalle d'entiers $\llbracket 1; 9 \rrbracket$ et dont la loi est déterminée par :

$$\forall k \in \llbracket 1; 9 \rrbracket, \quad \mathbb{P}(X = k) = p_k = \log(1 + 1/k).$$

15. Montrer que $\sum_{k=1}^9 p_k = 1$.

16. Montrer que $\mathbb{E}(X) = 9 - \log 9!$.

Considérons une variable aléatoire réelle X (dont on dira qu'elle suit la loi de Haddock) qui prend ses valeurs dans l'intervalle d'entiers $\llbracket 1; 9 \rrbracket$ et dont la loi est déterminée par :

$$\forall k \in \llbracket 1; 9 \rrbracket, \quad \mathbb{P}(X = k) = q_k = c \cdot \left(\frac{1}{k} - \frac{1}{k+1} \right),$$

où c est un réel.

17. Déterminer la valeur de c pour que $\sum_{k=1}^9 q_k = 1$.

18. Calculer q_1 et q_2 .

19. Exprimer $\mathbb{E}(X)$ en fonction d'un nombre harmonique (somme partielle de la série harmonique).

Considérons une variable aléatoire réelle X , de fonction de répartition F_X continue.

20. Montrer que pour tout $k \in \mathbb{Z} : \mathbb{P}(\lfloor X \rfloor = k) = F_X(k+1) - F_X(k)$.

21. Montrer que pour tout $t \in [0, 1[: \mathbb{P}(\{X\} \leq t) = \sum_{n \in \mathbb{Z}} \mathbb{P}(n \leq X \leq n+t)$.

22. Montrer que pour tout $k \in \{1 \dots 9\} : \mathbb{P}(me(X) = k) = \mathbb{P}(\log k \leq \{\log X\} < \log(k+1))$.

Partie C - Lois de mantisses entières uniformes

Considérons une variable aléatoire réelle Z qui suit une loi uniforme sur le segment $]0, 1]$.

23. Donner l'expression de la fonction de répartition de $\log Z$.

24. Représenter graphiquement la fonction de répartition de $\log Z$.

25. Soit $t \in [0, 1[$.

(a) Montrer que $\mathbb{P}(\{\log Z\} \leq t) = \sum_{n=1}^{\infty} \mathbb{P}(-n \leq \log Z \leq -n+t)$.

(b) En déduire que $\mathbb{P}(\{\log Z\} \leq t) = \frac{1}{9}(10^t - 1)$.

26. Représenter graphiquement la fonction de répartition de $\{\log Z\}$.

27. Montrer que $me(Z)$ suit une loi uniforme.

Considérons une variable aléatoire réelle X qui suit une loi uniforme sur le segment $[0, 1[$, et Y une variable aléatoire à valeurs dans $\mathbb{Z}$. On pose $Z = 10^{X+Y}$.

28. Soit $k \in \llbracket 1; 9 \rrbracket$, montrer que $me(Z) = k \iff \log k \leq X < \log(k+1)$.
29. Quelle est la loi de $me(Z)$? Donner son espérance.

Soit X une variable aléatoire qui suit une loi exponentielle de paramètre $\lambda > 0$. On pose $Z = 10^X$.

30. Donner la loi de $me(Z)$.
31. Reconnaître la loi de $me(Z)$ pour $\lambda = \ln 10$.
32. Effectuer un développement limité en $\lambda = 0$ de $\mathbb{P}(me(Z) = k)$, pour $k \in \llbracket 1; 9 \rrbracket$.
33. Montrer que la loi de $me(Z)$ converge vers la loi de Benford quand $\lambda \rightarrow 0$. Interpréter.

Soit Z une variable aléatoire qui suit une loi géométrique de paramètre $0 < p < 1$. On pose $q = 1 - p$.

34. Montrer que pour tout couple d'entiers naturels $m \leq n$: $\sum_{k=m}^n q^{k-1} p = q^{m-1} - q^n$.
35. Montrer que $\mathbb{P}(me(Z) = 1) = \sum_{k=0}^{+\infty} (q^{10^k-1} - q^{2 \cdot 10^k-1})$.
36. Montrer que $\mathbb{P}(me(Z) = 1) > p$.
37. Si $p = \frac{1}{2}$, est-ce que $me(Z)$ peut suivre une loi uniforme?