INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ANALYSE DE L'INFORMATION

Concours interne d'attaché statisticien de l'Insee

SESSION 2023

ÉPREUVE DE MATHÉMATIQUES ET STATISTIQUES

Durée 4 heures

Coefficient 3

Sans documents – L'usage de la calculatrice est interdit

Le sujet comprend 7 pages

Exercice 1 .... 4 pts

Soit E un espace vectoriel sur $\mathbb{R}$ de dimension finie $n \geq 2$.
On note Id_E l'application identité de E.
Pour tout endomorphisme f de E, on note $f^0 = Id_E$ et pour tout entier naturel k, $f^{k+1} = f \circ f^k$.
Pour tout entier n non nul, $\mathcal{M}_n(\mathbb{R})$ désigne l'espace des matrices carrées avec n lignes, n colonnes et à coefficients réels.

Partie A

Dans cette partie, l'espace E est de dimension 2 et est rapporté à sa base canonique (e_1, e_2) :

$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

On considère l'endomorphisme f dont la matrice A dans la base (e_1, e_2) est :

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

1. Montrer que f est diagonalisable.
2. On note λ_1, λ_2 les valeurs propres de f telles que $\lambda_1 \leq \lambda_2$.
   Déterminer une base (u_1, u_2) de E formée de vecteurs propres de f respectivement associés aux valeurs propres λ_1 et λ_2.
   Donner la matrice D de f dans cette nouvelle base.
3. Soit un entier $m \geq 1$.
   Déterminer la matrice de f^m dans la base canonique. (On cherchera à exprimer A^m en fonction de D, P et P^{-1}, P étant la matrice de passage de la base canonique à la base (u_1, u_2)).
4. Pour toute matrice H de $\mathcal{M}_2(\mathbb{R})$, on note $\mathcal{C}(H) = \{M \in \mathcal{M}_2(\mathbb{R}), HM = MH\}$.
   (a) Donner deux éléments distincts de $\mathcal{C}(H)$.
   (b) Montrer que pour tout H, $\mathcal{C}(H)$ est un sous espace vectoriel de $\mathcal{M}_2(\mathbb{R})$.
   (c) En écrivant $M = \begin{pmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{pmatrix}$, trouver toutes les matrices M de $\mathcal{C}(D)$, où D est la matrice trouvée à la question 2.
   (d) Soit R une matrice de $\mathcal{M}_2(\mathbb{R})$ telle que $R^2 = D$. Montrer que R et D commutent (i.e. $RD = DR$).
   (e) Déterminer alors toutes les matrices R de $\mathcal{M}_2(\mathbb{R})$ vérifiant $R^2 = D$.
   On notera $\mathcal{R}(D)$ cet ensemble.
   (f) En déduire que toutes les matrices M de $\mathcal{M}_2(\mathbb{R})$ telles que $M^2 = A$ vérifient $P^{-1}MP \in \mathcal{R}(D)$.

Partie B

Dans cette partie, l'espace E est de dimension finie $n \geq 2$.
On considère une matrice A de $\mathcal{M}_n(\mathbb{R})$ qui possède n valeurs propres distinctes positives ou nulles.
Soit P une matrice inversible de $\mathcal{M}_n(\mathbb{R})$ telle que :

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = P^{-1}AP.$$

Pour toute matrice M de $\mathcal{M}_n(\mathbb{R})$, on note $\Delta_M = P^{-1}MP$. On note aussi $\mathcal{C}(A) = \{M \in \mathcal{M}_n(\mathbb{R}), AM = MA\}$.

1. Montrer que pour toute matrice M de $\mathcal{M}_n(\mathbb{R})$, $AM = MA$ si et seulement si $D\Delta_M = \Delta_M D$.
2. Pour toute matrice B de $\mathcal{M}_n(\mathbb{R})$, montrer que $DB = BD$ si et seulement si B est une matrice diagonale de $\mathcal{M}_n(\mathbb{R})$.
3. En déduire que $\mathcal{C}(A) = \left\{P\tilde{D}P^{-1}, \tilde{D} \,\text{matrice diagonale}\right\}$.
4. On cherche à déterminer l'ensemble de toutes les matrices de $\mathcal{M}_n(\mathbb{R})$ dont le carré vaut D.
   (a) Montrer que si la matrice R de $\mathcal{M}_n(\mathbb{R})$ satisfait $R^2 = D$, alors elle est diagonale.
   (b) En déduire l'ensemble de toutes les matrices R de $\mathcal{M}_n(\mathbb{R})$ satisfaisant $R^2 = D$.
5. En déduire l'ensemble des matrices M de $\mathcal{M}_n(\mathbb{R})$ telles que $M^2 = A$.

Exercice 2 .... 6 pts

Partie A

On note pour tout entier naturel n, $W_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n(x)\,\mathrm{d}x$.

1. Montrer que $(W_n)_{n\in\mathbb{N}}$ est monotone. En déduire que $(W_n)_{n\in\mathbb{N}}$ est convergente.
2. À l'aide d'une intégration par parties, montrer que :

$$\forall n \in \mathbb{N}, n \geq 2, W_n = \frac{n-1}{n}\, W_{n-2}.$$

3. Justifier que pour tout entier n, $W_n > 0$; puis montrer que $\displaystyle\lim_{n\to+\infty} \frac{W_{n+1}}{W_n} = 1$.
4. Montrer que pour tout entier n, $W_{2n} = \dfrac{(2n)!}{2^{2n}(n!)^2}\dfrac{\pi}{2}$ et que $W_{2n+1} = \dfrac{2^{2n}(n!)^2}{(2n+1)!}$.
5. En déduire que $\displaystyle\lim_{n\to+\infty} \frac{2^{4n}(n!)^4}{n((2n)!)^2} = \pi$.

Partie B

1. Soit f la fonction définie sur $]1, +\infty[$ par :

$$f(x) = \frac{1}{x - \frac{1}{2}} + \ln(x-1) - \ln(x).$$

   (a) Calculer la dérivée f' de f.

   (b) Calculer la limite de f en 1 et en $+\infty$.

   (c) Dresser le tableau de variations de f et donner le signe de f sur $]1, +\infty[$.

2. Soit $(a_n)_{n\in\mathbb{N}^*}$ et $(b_n)_{n\in\mathbb{N}^*}$ les suites définies par $a_n = \dfrac{n!\mathrm{e}^n}{n^n\sqrt{n}}$ et $b_n = \ln(a_n)$.

   (a) Pour tout entier $n \geq 2$, montrer que $b_n - b_{n-1} = (n - \frac{1}{2})f(n)$, où f est la fonction définie en B-1.

   (b) En déduire que la suite $(a_n)_{n\in\mathbb{N}^*}$ converge vers un réel $\ell \geq 0$.

3. Soit g la fonction définie sur $]2, +\infty[$ par :

$$g(x) = f(x) + \frac{1}{5x^2(x - \frac{1}{2})}$$

   où f est la fonction définie en B-1.

   (a) Vérifier que l'expression de la dérivée g' de g est donnée par :

$$g'(x) = \frac{-7x^2 + 16x - 4}{20x^3(x-1)(x - \frac{1}{2})^2}.$$

   (b) Dresser le tableau de variations de g et donner le signe de g.

Partie C

On définit pour tout entier $n \geq 2$, $c_n = \displaystyle\sum_{k=2}^{n} \frac{1}{k^2}$.

1. Pour tout entier $k \geq 2$, montrer que : $\dfrac{1}{k^2} \leq \displaystyle\int_{k-1}^{k} \frac{1}{x^2} dx$.
2. En déduire que pour tout entier $n \geq 2$, $c_n \leq \displaystyle\int_{1}^{n} \frac{1}{x^2} dx$.
3. Pour tout entier $n \geq 2$, montrer que : $c_n \leq 1$.
4. Montrer que $(c_n)_{n\geq 2}$ converge vers un réel $c \leq 1$.
5. À l'aide de l'égalité démontrée en B-2-a, montrer que pour tout $n \geq 2$:
$$b_n - b_{n-1} \geq -\frac{1}{5n^2}.$$
6. En déduire que pour tout entier $n \geq 2$, $b_n \geq -\dfrac{1}{5} c_n + 1$.
7. En déduire que la limite ℓ de la suite $(a_n)_{n\in\mathbb{N}^*}$ est strictement positive.
8. En calculant $\dfrac{a_n^4}{a_{2n}^2}$, déterminer la limite de $\dfrac{2^{4n}(n!)^4}{n((2n)!)^2}$.
9. En conclure à l'aide des questions précédentes et de la partie A, la valeur de ℓ.

Problème .... 10 pts

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. La fonction $\log_2$ est définie pour tout réel $t > 0$ par $\log_2(t) = \ln(t)/\ln(2)$.

Partie A – Incertitude d'un événement

Pour tout événement A de probabilité non nulle, on appelle incertitude de A la quantité $i(A) = \log_2(1/\mathbb{P}(A))$.

1. On lance une pièce telle que la probabilité d'obtenir Pile est égale à $p \in]0,1[$. On note A l'événement « Le résultat du lancer est Pile » et $\overline{A}$ son contraire.
   Représenter graphiquement la fonction $p \to p \cdot i(A) + (1-p) \cdot i(\overline{A})$. Vérifier que la fonction possède un maximum en $p = \dfrac{1}{2}$.
2. (a) Donner la valeur de $i(\Omega)$.
   (b) Montrer que si un événement A a la même probabilité que son contraire $\overline{A}$, alors son incertitude est égale à 1.
   (c) Soient A et B deux événements tels que $A \subset B$ et $\mathbb{P}(A) > 0$. Comparer $i(A)$ et $i(B)$.
   (d) Soient A et B deux événements de probabilités non nulles. Montrer que :
   — $i(A \cup B) \leq \min(i(A), i(B))$;
   — si $\mathbb{P}(A \cap B) > 0$, alors $i(A \cap B) \geq \max(i(A), i(B))$.
   (e) Soient A et B deux événements indépendants de probabilités non nulles. Montrer que $\mathbb{P}(A \cap B) > 0$ et exprimer $i(A \cap B)$ à l'aide de $i(A)$ et $i(B)$.

Partie B – Entropie des variables aléatoires discrètes

Pour X une variable aléatoire discrète définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles, on pose sous réserve d'existence :

$$H(X) = \sum_{x \in X(\Omega)} P(X = x) \cdot i(X = x),$$

où $X(\Omega)$ désigne le support de la loi de X. On dit que $H(X)$ désigne l'entropie de X (ou incertitude moyenne de X).

1. Soit n un entier naturel et U une variable aléatoire qui suit une loi uniforme sur l'ensemble des entiers compris entre 0 et n : $\{0, \ldots, n\}$.
   (a) Calculer $H(U)$.
   (b) Soit B une variable de Bernoulli de paramètre $p \in]0,1[$, indépendante de U.
   Calculer $H(UB)$.
   Comparer ce résultat à la valeur de $H(U)$.
   Commenter ce résultat, à la lumière des valeurs limites de p.
2. Montrer que $H(X) \geq 0$ et que $H(X) = 0$ si et seulement si X est presque sûrement constante.
3. (a) Montrer que pour réel $t > 0$, $\ln t \leq t - 1$.
   (b) Soient n un entier naturel, $(p_0, \ldots, p_n)$ et $(q_0, \ldots, q_n)$ deux $(n+1)$-uplets de réels strictement positifs tels que $\sum_{i=0}^{n} p_i = 1$ et $\sum_{i=0}^{n} q_i = 1$.
   Montrer que :
   $$\sum_{i=0}^{n} q_i \ln(p_i/q_i) \leq 0.$$

(c) Soit X une variable aléatoire à valeurs dans l'ensemble $\{0, \dots, n\}$.
Montrer que $H(X) \leq \log_2(n+1)$.
Commenter ce résultat.

4. Soit G une variable aléatoire qui suit une loi géométrique de paramètre $p \in]0, 1[$:

$$\forall k \geq 1, \qquad p_k = \mathbb{P}(G = k) = (1-p)^{k-1} p.$$

(a) Calculer la valeur de $m = \mathbb{E}(G)$.

(b) En admettant son existence, calculer la valeur de $H(G)$.

(c) Soit X une variable aléatoire strictement positive et à valeurs entières ($q_k := \mathbb{P}(X = k) > 0$ pour tout $k \in \mathbb{N}^*$). Supposons que X admet une espérance notée m et que $H(X)$ existe.

i. En appliquant l'inégalité $\ln t \leq t - 1$ à p_k/q_k, montrer que pour tout entier $k \geq 1$:
$$q_k \ln(p) + (k-1) q_k \ln(1-p) - q_k \ln(q_k) \leq p_k - q_k \,;$$

ii. Etablir que $H(X) \leq H(G)$, puis que $H(X) = H(G)$ si et seulement si X et G suivent la même loi ;

iii. Commenter ce résultat.

5. Pour X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles, on pose sous réserve d'existence :

$$H(X, Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \mathbb{P}(X = x; Y = y) \cdot i(X = x; Y = y),$$

où $X(\Omega)$ et $Y(\Omega)$ désignent les supports des lois de X et Y. On dit que $H(X, Y)$ désigne l'entropie du couple (X, Y).

(a) Soient X et Y deux variables de Bernoulli indépendantes et de paramètres $1/2$. On pose $S = X + Y$ et $P = X \cdot Y$.

i. Est-ce que les variables S et P sont indépendantes ?

ii. Calculer $H(S)$, $H(P)$ et $H(S, P)$. Comparer $H(S, P)$ et $H(S) + H(P)$. Commenter ce résultat.

(b) Sous réserve d'existence des différentes entités, montrer que $H(X, Y) \geq \max(H(X), H(Y))$.

(c) Sous réserve d'existence des différentes entités, montrer que $H(X, Y) \leq H(X) + H(Y)$.
On montrera dans un premier temps que :

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \mathbb{P}(X = x; Y = y) \log_2\left(\frac{\mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)}{\mathbb{P}(X = x; Y = y)}\right) \leq 0.$$

(d) Supposons que X et Y sont indépendantes. Sous réserve d'existence des différentes entités, montrer que $H(X, Y) = H(X) + H(Y)$.

C – Entropie des variables aléatoires continues

Pour X une variable aléatoire à densité définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles et de densité f_X, on pose sous réserve d'existence :

$$H(X) = \int_{\mathbb{R}} f_X(t) \log_2(1/f_X(t))dt.$$

1. Si N_0 suit une loi normale centrée et réduite de densité fournie par la relation $f_{N_0}(t) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}$, déterminer la valeur de $H(N_0)$ (on admettra son existence).
2. Si N suit une loi normale de moyenne m et de variance σ^2, déterminer la valeur de $H(N)$ (on admettra son existence).
3. Soit X_0 une variable aléatoire suivante une loi exponentielle de paramètre λ, dont la densité est donnée par la relation $f_{X_0}(t) = \lambda \mathrm{e}^{-\lambda t}\mathbf{1}_{[0;+\infty[}(t)$.

   (a) En admettant l'existence de $H(X_0)$, déterminer sa valeur.

   (b) Soit X une variable aléatoire à valeurs dans $\mathbb{R}_+^*$, admettant une densité f. On suppose que $H(X)$ existe et que X admet une espérance égale à $1/\lambda$. Vérifier que :

   $$H(X_0) = -\frac{1}{\ln 2}\int_0^{+\infty} f(t)\ln(f_{X_0}(t))dt.$$

   Montrer que $H(X) \leq H(X_0)$.
   Commenter ce résultat.