CONCOURS ADMINISTRATEUR EXTERNE DE L'INSEE

Session 2022

ÉPREUVE DE MATHÉMATIQUES

DURÉE : 4 heures

L'énoncé comporte 8 pages, numérotées de 1 à 8

Tous documents et appareils électroniques interdits

Partie 1 : algèbre-analyse

Cette partie est constituée de deux exercices indépendants.

Exercice 1

Définitions et notations :

- **On note $\binom{n}{k}$ le coefficient du binôme défini, pour $(k,n) \in \mathbb{N}^2$ vérifiant $0 \leqslant k \leqslant n$, par $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$.**
- **On note E l'espace vectoriel des fonctions de classe C^1 de $[0,1]$ dans $\mathbb{R}$.**
- **À toute fonction f appartenant à E, on associe la fonction $B_n(f)$ définie sur $[0,1]$ par :**

$$\forall n \in \mathbb{N}, \quad \forall x \in [0,1], \quad B_n(f)(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.$$

Remarque
On pourra, pour simplifier certains calculs, utiliser une variable aléatoire Z qui suit la loi binomiale $\mathcal{B}(n,x)$, où $x \in [0,1]$

1. **On suppose dans cette question que f est définie sur $[0,1]$ par $f(x) = x^2$.**
   (a) **Calculer, pour tout entier naturel n, $B_n(f)$.**
   (b) **Montrer que $B_n(f)$ converge uniformément vers la fonction f quand n tend vers $+\infty$.**

   Dans la suite, on revient au cas général où f est une fonction quelconque de classe C^1 de $[0,1]$ dans $\mathbb{R}$.

2. **Calculer $\sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^k (1-x)^{n-k}$ en fonction de x et de n.**
3. **Soit $x \in [0,1]$ et a un réel strictement positif.**
   On pose : $I = \{k \in [\![0,n]\!] \ ; \ |k-nx| \geqslant na\}$ et $J = \{k \in [\![0,n]\!] \ ; \ |k-nx| < na\}$.
   (a) **Justifier l'existence d'un réel K positif tel que :**
   $$\forall (x,y) \in [0,1]^2, \quad |f(x) - f(y)| \leqslant K|x-y|.$$
   (b) **Établir, pour tout entier naturel n non nul, l'inégalité suivante :**
   $$\sum_{k \in I} \binom{n}{k} x^k (1-x)^{n-k} \leqslant \frac{x(1-x)}{na^2}.$$
   (c) **En déduire le résultat suivant :**
   $$\forall x \in [0,1], \quad \forall n \in \mathbb{N}^*, \quad |f(x) - B_n(f)(x)| \leqslant K\left(a + \frac{1}{4a^2 n}\right).$$
   (d) **Montrer que $B_n(f)$ converge uniformément vers f quand n tend vers $+\infty$.**
4. **Soit f une fonction de classe C^1 sur $[0,1]$ telle que, pour tout entier naturel n, $\int_0^1 t^n f(t)dt = 0$.**
   Montrer que $f = 0$.
5. **Soit g une fonction de classe C^1 sur $\mathbb{R}_+$ telle que :**
   - $\lim_{x \to +\infty} g(x) = 0$.
   - **Il existe $\ell \in \mathbb{R}$ tel que $\lim_{t \to +\infty} e^t g'(t) = \ell$.**
   - **Pour tout réel x strictement positif, $\int_0^{+\infty} e^{-tx} g(t)dt = 0$.**

   Montrer que $g = 0$.

Exercice 2

Les trois premières parties sont indépendantes entre elles. La quatrième fait la synthèse de ces trois parties.

L'espace de référence est $E = M_n(\mathbb{C})$. On rappelle que tout élément de E est ***trigonalisable***, c'est-à-dire semblable à une matrice triangulaire supérieure dont les éléments diagonaux sont les valeurs propres.

1ère partie

Soit f une ***forme linéaire*** définie sur $M_n(\mathbb{C})$. On considère les deux propriétés :

- f est commutative vis-à-vis du produit matriciel, c'est-à-dire : $f(AB) = f(BA)$.
- f est invariante par similitude, c'est-à dire que si P est semblable à Q : $f(P) = f(Q)$.

1. Montrer que : $\boxed{i \Rightarrow ii.}$

2. On veut montrer que : $\boxed{ii \Rightarrow i.}$

   a) Montrer que, si A ou B sont inversibles, AB est semblable à BA et conclure.
   b) Montrer que, dans le cas opposé, il existe $\lambda \in \mathbb{C}$ tel que $A - \lambda I_n$ est inversible.
   c) Conclure.

3. On suppose maintenant que f vérifie les propriétés i et ii. On note $E_{i,j}$ les matrices élémentaires n'ayant qu'un seul 1 à la ligne *i* et la colonne *j* et 0 ailleurs.

   a) Calculer $E_{i,j}E_{j,i}$ et $E_{i,i}E_{i,j}$.
   b) En déduire les valeurs de $f(E_{i,i})$ et $f(E_{i,j})$ $pour\ i \neq j$.
   c) En déduire que : $\exists\, \alpha \in \mathbb{C},\ \forall\ A \in E : f(A) = \alpha\, Tr(A)$.
   d) Que peut-on dire de la réciproque du résultat ainsi démontré ?

2ème partie

On suppose maintenant que f est une forme linéaire non nulle mais ne vérifiant pas nécessairement les propriétés i et ii précédentes.
On pose $H = Ker f$. Soit $A_0 \in E$ telle que : $f(A_0) \neq 0$.

4. Montrer que : $E = H \oplus Vect\{A_0\}$.

5. On considère une application s de E dans $\mathbb{R}^+$ vérifiant les propriétés suivantes :

$$\begin{cases} \forall\ A, B \in E : s(A+B) \leq s(A) + s(B) \\ \forall\ \lambda \in \mathbb{C} : s(\lambda A) = |\lambda| s(A) \end{cases}.$$

   s s'appelle ***semi-norme*** sur E. On notera qu'**on n'a pas** l'implication :

$$\forall\ A \in E : s(A) = 0 \Rightarrow A = 0.$$

   On suppose que : $\boxed{\forall\ Q \in H : s(Q) = 0.}$

   a) Montrer que : $\forall\ A \in E, \forall\ Q \in H : s(A+Q) = s(A)$.
   b) En déduire qu'il existe $\beta \geq 0$ **tel que :** $\forall\ A \in E : s(A) = \beta\, |f(A)|$.

3ème partie

Soit N une matrice ***nilpotente*** c'est-à-dire telle qu'il existe $p \in \mathbb{N},\ p \geq 2$ vérifiant $N^p = 0$.

6. Montrer que N est semblable à une matrice triangulaire supérieure dont la diagonale est nulle.

7. Réciproquement, montrer que toute matrice triangulaire supérieure de diagonale nulle est nilpotente.

8. Montrer que toute matrice de diagonale nulle ***non nécessairement triangulaire*** est la somme de deux matrices nilpotentes.

9. On admet que ***toute matrice de trace nulle est semblable à une matrice de diagonale nulle***. Montrer que le sous-espace vectoriel de E engendré par les matrices nilpotentes est l'ensemble des matrices de trace nulle.

4ème partie

Soit s une semi-norme définie sur E (cf. question 5) non identiquement nulle qu'on suppose ***invariante par similitude*** (définition analogue à celle de la première partie).

10.
   a) Soit N une matrice nilpotente de E, semblable d'après la question 6 à une matrice triangulaire supérieure T de diagonale nulle.

   En interprétant T comme la matrice représentative d'un endomorphisme g de $\mathbb{C}^n$ relativement à une base $(e_1, \ldots, e_n)$ et en définissant une nouvelle base sous la forme : $e'_1 = e_1,\ e'_k = \mu_k e_k\ (k = 2, \ldots, n)$, montrer que, pour tout $q \in \mathbb{N}^*$, on peut trouver une matrice triangulaire supérieure T_q semblable à T dont les éléments $t_{i,j}$ vérifient :

   $$|t_{i,j}| < \frac{1}{q}.$$

   On cherchera les coefficients μ_k appropriés.

   b) Montrer que : $s(T_q) \to 0$ quand $q \to +\infty$.

   c) En déduire que : $s(N) = 0$.

11.
   a) En déduire que s s'annule sur l'ensemble des matrices de trace nulle.

   b) Conclure de ce problème que toute semi-norme s invariante par similitude est de la forme :

   $$\forall\ A \in E : s(A) = \gamma |Tr(A)|,$$

   où γ est un réel positif ou nul fixé.

Partie 2 : probabilités-statistiques

Cette partie est constituée de deux exercices indépendants.

Exercice 1

Toutes les variables aléatoires de cet exercice sont supposées être définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Soit a et b deux réels, $b > 2$ et X une variable aléatoire de densité :

$$f(t) = \begin{cases} \dfrac{b}{(x-a)^{b+1}} \text{ si } x \geqslant a+1 \\ \\ 0 \text{ sinon} \end{cases}$$

On dit que X suit la loi de Pareto $\mathcal{P}(a, b)$.
On suppose que a et b sont deux paramètres inconnus que l'on désire estimer.
On considère donc un $n-$échantillon $X_1, \ldots, X_n$ de n variables aléatoires indépendantes suivant toutes la même loi que X.

Préliminaire

1. Déterminer la fonction de répartition de X.
2. Étudier l'existence de l'espérance de X et en cas d'existence, la calculer.

Partie 1

On suppose dans cette partie que le paramètre b est connu et on désire construire un estimateur de a.
On pose $M_n = \min(X_1, \ldots, X_n)$ et $A_n = M_n - 1$.

3. Montrer que M_n suit une loi de Pareto dont on précisera les paramètres.
4. Calculer $\mathbb{E}(|A_n - a|)$.
5. En déduire que A_n est un estimateur convergent de a.
6. Soit $\alpha \in]0, 1[$.
   Déterminer en fonction de n, de b et de A_n un intervalle de confiance pour a au risque α.

Partie 2

On suppose dans cette partie que le paramètre a est connu et on désire construire un estimateur de b.
On pose :

$$B_n = \frac{n}{\sum\limits_{i=1}^{n} \ln(X_i - a)}$$

7. Donner la loi de $\ln(X - a)$.
8. En déduire que B_n converge en probabilité vers b.

Partie 3

On suppose dans cette partie que a et b sont inconnus et on désire estimer b.

$$\hat{B}_n = \frac{n}{\sum\limits_{k=1}^{n} \ln(X_i - A_n)} \text{ et } R_n = \frac{1}{n}\sum_{i=1}^{n} \ln(X_i - a) - \frac{1}{n}\sum_{i=1}^{n} \ln(X_i - A_n)$$

On admettra le résultat suivant :
si (Z_n) est une suite de variables aléatoires qui converge en loi vers la loi de Z et (T_n) une suite de variable aléatoire qui converge en probabilité vers une variable certaine T, alors les suites de variables (Z_nT_n) et (Z_n+T_n) convergent en loi respectivement vers les lois des variables alétoires ZT et $Z+T$.

9. (a) Montrer que, pour tout $x > 0$, $\ln(x) \leqslant x - 1$
   (b) En déduire l'inégalité suivante :
$$R_n \leqslant \frac{1}{n}\sum_{i=1}^{n}\frac{A_n - a}{X_i - A_n}$$
   (c) En déduire que $\sqrt{n}R_n$ converge en probabilité vers 0.
10. (a) Montrer que $\sqrt{n}\left(\frac{1}{b} - \frac{1}{n}\sum_{i=1}^{n}\ln(X_i - a)\right)$ converge en loi vers $\mathcal{N}(0, \frac{1}{b^2})$.
   (b) En déduire que $\sqrt{n}\left(\frac{1}{b} - \frac{1}{n}\sum_{i=1}^{n}\ln(X_i - a)\right) + \sqrt{n}R_n$ converge en loi et donner la loi limite.
   (c) Simplifier $b\hat{B}_n\left(\sqrt{n}\left(\frac{1}{b} - \frac{1}{n}\sum_{i=1}^{n}\ln(X_i - a)\right) + \sqrt{n}R_n\right)$.
   (d) Montrer que $\hat{B}_n$ converge en probabilité vers b.
   (e) En déduire que $\sqrt{n}(\hat{B}_n - b)$ converge en loi et donner la loi limite.
   Donner un intervalle de confiance pour b au risque $0,05$.

Exercice 2

Chacune des parties de ce problème utilise les résultats du prologue et des parties précédentes.

Prologue

Soit $Z \sim> N(0,\ 1)$. Calculer $E(e^{s\,Z})$ pour $s \in \mathbb{R}$.

■ ■ ■

On considère le modèle linéaire $Y_i = bx_i + U_i$, pour $i \in \mathbb{N}$, où les x_i sont des observations ***non aléatoires***, les U_i sont des variables aléatoires indépendantes et de même loi $N(0,\ \sigma^2)$ $(\sigma \neq 0)$ et b un paramètre inconnu.

On estime ce modèle sur les données relatives à la période $i = 1,\ \ldots,\ n$.

1^{ère} partie

1. Calculer l'estimateur des moindres carrés ordinaires de b, soit $\hat{b}_n$, et donner sa loi.
2. On suppose **dans cette question seulement** que : $\dfrac{1}{n}\sum_{k=1}^{n} x_i^2 \to \mu_{2,x}$ quand $n \to +\infty$.
   Étudier la convergence en loi de $\sqrt{n}(\hat{b}_n - b)$ quand $n \to +\infty$.
3. **Pour cette question seulement**, étude du cas particulier $x_i = i$. Déterminer α pour que $n^{\alpha}(\hat{b}_n - b)$ converge en loi vers une loi non dégénérée (différente d'une constante) quand $n \to +\infty$.

On revient au cas général.

4. Pour $i \in \{1,\ 2,\ \ldots,\ n\}$, on pose : $\hat{U}_i = Y_i - \hat{b}_n x_i$.
   a) Expliciter la loi de $\hat{U}_i$.
   b) Calculer $E\left(\sum_{i=1}^{n} \hat{U}_i^2\right)$.
   c) En déduire un estimateur sans biais de σ^2, noté $\hat{\sigma}_n^2$.
5.
   a) Calculer $Cov(\hat{b}_n,\ \hat{U}_k)$ pour $k \in \{1,\ \ldots,\ n\}$.
   b) Montrer que $\hat{b}_n$ et $\hat{U}_k$ sont indépendants.

2^{ème} partie

6. On considère un instant $\boxed{j > n.}$
   a) Quel est le prédicteur naturel de Y_j fondé sur ce modèle ? On le notera $\hat{Y}_j$.
   b) Calculer $E(\hat{Y}_j - Y_j)$ et $E(\hat{Y}_j - Y_j)^2$.

En réalité, la variable d'intérêt (expliquée par le modèle) est une variable $R_i > 0$ (par exemple le revenu du ménage *i),* définie par : $Y_i = \mathrm{Ln}\, R_i$. Le prédicteur intuitif de R_j (pour $j > n$) est alors :

$$\hat{R}_j = e^{\hat{Y}_j}.$$

7. Calculer $E(\hat{R}_j - R_j)$. Qu'en conclut-on ?

8. On pose : $\hat{\hat{R}}_j = e^{\lambda_j}\, \hat{R}_j$.
Déterminer, en fonction de $\sigma^2,\, x_i\,(i = 1, \dots, N)$ et x_j, la valeur de λ_j qui assure que $E(\hat{\hat{R}}_j - R_j) = 0$.

3ème partie

9. On réalise un tirage à probabilités égales d'un indice parmi $\{1, 2, \dots, n\}$. On appelle I la variable aléatoire représentant l'indice tiré. On suppose que I est **indépendant** des U_i. On admet que $\hat{U}_I$ (avec l'indice aléatoire I) est une variable aléatoire.
   a) Exprimer sous forme sommatoire la densité g de la loi de $\hat{U}_I$.
   b) En déduire la valeur de $E\,\hat{U}_I$.
   c) En déduire la valeur de $V\,\hat{U}_I$.

10. Montrer que $\hat{b}_n$ et $\hat{U}_I$ sont indépendants.
On pourra utiliser les fonctions de répartition.

11. On revient au cadre de la 2ème partie. On pose : $Y_j^* = \hat{b}_n x_j + \hat{U}_I$ et : $R_j^* = e^{Y_j^*}$.
   a) Calculer $E\, e^{\hat{U}_I}$.
   b) En déduire $E\,R_j^*$.
   c) Donner une expression approchée de $E\,R_j^*$ quand les x_i^2 et x_j^2 sont petits par rapport à $\sum_{k=1}^{n} x_k^2$ et conclure.

■ ■ ■ ■ ■ ■