CONCOURS INTERNE 2020
POUR LE RECRUTEMENT D'ADMINISTRATEURS DE L'INSEE

COMPOSITION DE MATHÉMATIQUES

(durée: 4 heures)

L'énoncé comporte 6 pages, numérotées de 1 à 6.

Tout document ou appareil électronique est interdit.

L'épreuve est constituée de deux parties indépendantes,
chacune comptant pour moitié dans la note finale

Tournez la page S.V.P.

Partie 1 : Algèbre et Analyse

Cette partie est constituée de deux exercices indépendants.

Exercice 1

Soient E un espace euclidien muni de sa norme usuelle $\|.\|$ et u un endomorphisme *symétrique* de E, dont **toutes les valeurs propres sont** >0.

On définit l'application : $q : x \in E-\{0\} \to q(x) = \dfrac{[<x, u(x)>]^2}{\|u(x)\|^2\|x\|^2}$.

L'objet de ce problème est la recherche et la détermination des *extrema* de q.

1. Montrer que q est bien définie sur $E-\{0\}$.

2.
   a) Pour $\lambda \in \mathbb{R}^*$ et $x \in E-\{0\}$, calculer $q(\lambda x)$.
   b) En déduire que les *extrema* de q sur $E-\{0\}$ (s'ils existent) sont ceux de q sur la *sphère unité* $S(0,1) = \{x \in E \,;\, \|x\| = 1\}$.
   On ne demande pas à ce stade de démontrer l'existence de ces extrema.

3.
   a) Montrer que : $\forall\, x \in E-\{0\} : q(x) \leq 1$.
   b) En déduire la valeur du maximum de q et déterminer les éléments en lesquels il est atteint.

4.
   a) Montrer que, en exprimant *x* par ses composantes x_i sur une base adéquate de E, on peut écrire : $q(x) = \dfrac{(\sum_{i=1}^{n} \lambda_i x_i^2)^2}{(\sum_{i=1}^{n} \lambda_i^2 x_i^2)(\sum_{i=1}^{n} x_i^2)}$, où les λ_i sont les valeurs propres de u.
   b) Que devient cette expression quand on exprime $q(y)$ où $y \in S(0,1)$?

5.
   a) Montrer que 0 ne peut être le minimum de *q*.
   b) On note respectivement λ_m et λ_M la plus petite (resp. la plus grande) valeur propre de u. Montrer que : $\forall\, x \neq 0 : q(x) \geq \dfrac{\lambda_m^2}{\lambda_M^2}$.
   c) Que peut-on en conclure si toutes les valeurs propres de u sont égales ? Qu'est alors u ?

On suppose dans la suite que les valeurs propres de u ne sont pas toutes identiques.

6. On se place ici dans le cas $\boxed{Dim\, E = 2.}$
   a) Montrer que la recherche des *extrema* de q sur la sphère unité équivaut à la recherche des *extrema* d'une fonction d'une seule variable réelle, notée $\tilde{q}$.
   b) En étudiant cette fonction $\tilde{q}$, déterminer *analytiquement* (*via* les conditions du 1er ordre) les points de la sphère unité en lesquels q atteint son minimum (au moyen de leurs coordonnées sur la base considérée) et calculer la valeur de ce minimum.

7. On revient au cas où la dimension de E est quelconque, égale à $n \geq 3$. On numérote les valeurs propres de u en ordre croissant : $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$.

a) Montrer que la recherche des *extrema* de q sur la sphère unité équivaut à la recherche des *extrema* de la fonction de *n-1* variables : $\widetilde{q}(y_1, \dots, y_{n-1}) = \dfrac{\left(\sum\limits_{i=1}^{n-1}(\lambda_i - \lambda_n)y_i^2 + \lambda_n\right)^2}{\sum\limits_{i=1}^{n-1}(\lambda_i^2 - \lambda_n^2)y_i^2 + \lambda_n^2}$.

b) Montrer que les conditions du 1[er] ordre pour la recherche des *extrema* de $\widetilde{q}$ s'écrivent sous la forme :

$$\forall j \in \{1, \dots, n-1\} : y_j(\lambda_j - \lambda_n)\left(2 - \frac{P(y)}{Q(y)}(\lambda_j + \lambda_n)\right) = 0, \quad \textbf{(S)}$$

avec : $\begin{cases} P(y) = \sum\limits_{i=1}^{n-1}(\lambda_i - \lambda_n)y_i^2 + \lambda_n \\ Q(y) = \sum\limits_{i=1}^{n-1}(\lambda_i^2 - \lambda_n^2)y_i^2 + \lambda_n^2 \end{cases}$.

c) On se place dans le cas particulier **où toutes les valeurs propres de u sont *distinctes*.**

Montrer que la résolution du système **(S)** se ramène à un problème à deux variables comme dans le cas où $Dim\, E = 2$.

En admettant que les conditions du 1[er] ordre conduisent effectivement à un minimum, déterminer alors le minimum de $\widetilde{q}$, puis celui de q sur la sphère unité.

Exercice 2:

Dans tout l'exercice, n désigne un entier supérieur ou égal à 2.
On considère une suite réelle $(b_n)_{n\geqslant 2}$ à termes strictement positifs et la suite $(u_{n,k})$ définie par :

$$\forall k \in \mathbb{N}^*, \quad \forall n \in \mathbb{N},\ n \geqslant 2, \quad u_{n,k} = \frac{1}{\left(1+\dfrac{kb_n}{n}\right)^n}$$

1. Montrer que pour tout n fixé, la série de terme général $u_{k,n}$ converge.

   On pose alors $S_n = \displaystyle\sum_{k=1}^{+\infty} \frac{1}{\left(1+\dfrac{kb_n}{n}\right)^n}$

2. Établir l'inégalité suivante :

$$S_n \geqslant \frac{1}{e^{b_n}-1}$$

3. On considère la fonction f_n définie sur $\mathbb{R}_+$ par :

$$\forall x \geqslant 0, \quad f_n(x) = \frac{1}{\left(1+\dfrac{xb_n}{n}\right)^n}$$

   (a) Montrer que $S_n \leqslant \displaystyle\int_0^{+\infty} f_n(x)\mathrm{d}x$.

   (b) En déduire que $b_n S_n \leqslant 2$.

4. On pose :

$$S_{n,p} = \sum_{k=1}^{p} \frac{1}{\left(1+\dfrac{kb_n}{n}\right)^n}$$

   On suppose que la suite (b_n) converge vers un réel ℓ strictement positif et on admet le résultat suivant :

   Si la suite $(S_{n,p})_{p\geqslant 1}$ converge uniformément vers S_n quand p tend vers $+\infty$, alors :

$$\lim_{n\to+\infty}\left(\lim_{p\to+\infty} S_{n,p}\right) = \lim_{p\to+\infty}\left(\lim_{n\to+\infty} S_{n,p}\right)$$

   (a) Justifier l'existence d'un réel b strictement positif tel que :

$$0 \leqslant u_{k,n} \leqslant \frac{1}{\left(1+\dfrac{kb}{n}\right)^n}$$

   (b) Étudier les variations de la fonction h définie sur $\mathbb{R}_+^*$ par :

$$h(x) = \frac{1}{\left(1+\dfrac{kb}{x}\right)^x}$$

   (c) En déduire l'encadrement suivant :

$$0 \leqslant u_{n,k} \leqslant \frac{1}{\left(1+\dfrac{kb}{2}\right)^2}$$

5. (a) Montrer que la suite $(S_{n,p})_{p\geqslant 1}$ converge uniformément vers S_n quand p tend vers $+\infty$.

   (b) En déduire, en fonction de ℓ, la valeur de $\displaystyle\lim_{n\to+\infty} S_n$.

Partie 2 : Probabilités et 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 $(\Omega, \mathcal{A}, \mathbb{P})$.
On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables aléatoires indépendantes suivant toutes la même loi et une variable N, à valeur dans $\mathbb{N}^*$ et indépendante des variables X_k.
On considère les deux applications Z_n et Z définies sur Ω par :

$$\forall n \in \mathbb{N}^*,\ \forall \omega \in \Omega, \quad Z_n(\omega) = \max\big(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)\big) \text{ et } \forall \omega \in \Omega,\ Z(\omega) = \max\big(X_1(\omega), X_2(\omega), \ldots, X_{N(\omega)}\big)$$

On admet que Z_n et Z sont des variables aléatoires.

1. **On suppose dans cette question que les variables aléatoires X_k suivent toutes la loi géométrique de paramètre p, $(0 < p < 1)$ et que N suit la loi de Mengoli, c'est-à-dire que :**

$$\forall m \in \mathbb{N}^*, \quad \begin{cases} \mathbb{P}([X_k = m]) = p(1-p)^{m-1} \\ \\ \mathbb{P}([N = m]) = \dfrac{1}{m(m+1)} \end{cases}$$

   (a) **On considère un réel x appartenant à $]0, 1[$.**

   i. **Établir, pour tout entier naturel m non nul, la formule suivante :**

$$\sum_{k=1}^{m} \frac{x^k}{k} = -\ln\left(1-x\right) - \int_0^x \frac{t^m}{1-t}\mathrm{d}t$$

   ii. **En déduire la formule suivante :**

$$\sum_{k=1}^{+\infty} \frac{x^k}{k} = -\ln\left(1-x\right)$$

   iii. **En déduire également l'égalité suivante :**

$$\sum_{k=1}^{+\infty} \frac{x^{k+1}}{k(k+1)} = x + (1-x)\ln\left(1-x\right)$$

   (b) i. **Calculer, pour tout entier naturel m non nul, $\mathbb{P}([Z_n \leqslant m])$.**

   ii. **Établir, pour tout entier naturel m non nul, l'égalité suivante :**

$$\mathbb{P}([Z \leqslant m]) = \sum_{k=1}^{+\infty} \frac{(1-(1-p)^m)^k}{k(k+1)}$$

   iii. **En déduire la loi de Z.**

2. **On suppose dans cette question que les variables aléatoires X_k suivent toutes la loi uniforme sur $[0, 1]$ et que la loi de N est donnée par :**

$$\forall m \in \mathbb{N}^*, \quad \mathbb{P}([N = m]) = \frac{1}{(e-1)m!}$$

   (a) **Vérifier que la relation précédente définit bien une loi de probabilité.**

   (b) i. **Déterminer la fonction de répartition de Z.**

   ii. **Vérifier que Z admet une densité et en donner une densité.**

   iii. **Calculer $\mathbb{E}(Z)$.**

   (c) **On considère dans cette question une variable aléatoire Y, indépendante des X_k, qui suit la loi géométrique de paramètre $1 - \dfrac{1}{e}$ et on pose $T = Y - Z$.**
   Calculer, pour tout réel x, $\mathbb{P}([T > x]$ et reconnaître la loi de T.

Exercice 2 :

L'objet de ce problème est le calcul de la médiane d'un « mélange » (défini ci-après) de lois lognormales en fonction des médianes de ces dernières, soit de manière exacte (mais au moyen d'une équation non soluble analytiquement), soit de manière approchée.

Les parties 2 et 3 de ce problème sont indépendantes mais utilisent les résultats de la première partie.

1ère partie : généralités sur la loi lognormale

On dit qu'une variable aléatoire Z à valeurs >0 suit une *loi lognormale* si $X = Ln\ Z$ suit une loi normale $\mathcal{N}(m,\sigma^2)$ ($\sigma^2 > 0$ *est la* ***variance*** *de la loi normale*).

1. Soit Z_1 une variable aléatoire à valeurs >0 telle que : $X_1 = Ln\ Z_1$ ~> $\mathcal{N}$(0,1).
   a. Déterminer la fonction de répartition de la loi de Z_1. On l'exprimera au moyen de la fonction de répartition *H* de la loi $\mathcal{N}$ (0,1).
   b. Écrire la densité de la loi de Z $_1$.

2. Soit Z_2 une variable aléatoire à valeurs >0 telle que : $X_2 = Ln\ Z_2$ ~> $\mathcal{N}(m,\sigma^2)$
   a. Calculer le quantile d'ordre $\alpha \in]0,1[$ de la loi de Z_2. On le notera q_α et on l'exprimera au moyen de la fonction de répartition *H* de la loi $\mathcal{N}$ (0,1) (*on rappelle que celle-ci est continue et strictement croissante*).
   b. En déduire la valeur de la médiane μ de la loi de Z_2.
   c. Montrer que la loi lognormale est entièrement déterminée par la connaissance de sa médiane μ et d'un autre quantile *quelconque* d'ordre $\alpha \in]0,1[/ \{\frac{1}{2}\}$.

2ème partie

On considère ici deux familles de variables aléatoires : $A_1 = \{X_1, ..., X_{n_1}\}$ et $A_2 = \{Y_1, ..., Y_{n_2}\}$. Toutes les variables considérées sont indépendantes entre elles. Les variables X_i (resp. Y_j) suivent la même loi lognormale de densité f_1 et de fonction de répartition F_1 (resp. f_2 et F_2), issues de lois normales $N(m_k, \sigma_k^2), k = 1, 2$. On note μ_k les médianes de ces lois lognormales.

Conditionnellement au tirage des variables X_i et Y_j, on tire de manière équiprobable dans $A_1 \cup A_2$ une variable aléatoire, qui sera notée Z (égale donc à l'une des variables X_i ou Y_j).

3. Déterminer la loi de Z par une densité et/ou sa fonction de répartition.

4.
   a. Écrire, en fonction de *H* (définie en 1.a) et des paramètres μ_k et σ_k l'équation permettant de déterminer la médiane μ de la loi de Z.
   b. Donner une expression explicite de cette équation sous forme intégrale en utilisant la formulation de *H sous forme intégrale* et au moyen des mêmes paramètres.

3ème partie

5. On considère deux lois de probabilités P et P_0 sur $\mathbb{R}$, admettant des densités respectives f et f_0, supposées continues et > 0 sauf éventuellement en un point. On définit un *indicateur d'écart* entre P et P_0 : $K_{P/P_0} = \int_{-\infty}^{+\infty} -\ln\left(\frac{f(x)}{f_0(x)}\right) f_0(x)\,dx$ lorsque cette intégrale a un sens.

   a. Montrer que : $K_{P/P_0} \geq 0$.

   b. Montrer que $K_{P/P_0} = 0$ si et seulement si $P = P_0$.

   c. En déduire la solution du problème $Max \int_{-\infty}^{+\infty} [\ln f(x)]\, f_0(x)\,dx$ quand f décrit l'ensemble des densités de probabilités sur $\mathbb{R}$.

   Nota : cette définition se transpose sans difficulté au cas de probabilités définies seulement sur $\mathbb{R}^{+*}$.

6. On considère ici le cas où la loi P_0 est définie par sa densité : $f_0 = \sum_{i=1}^{n} \alpha_i f_i$ où les f_i sont les densités de lois lognormales P_i issues de lois normales $N(m_i, \sigma_i^2)$ avec $\sigma_i > 0$ et les $\alpha_i \in]0, 1[$ tels que : $\sum_{i=1}^{n} \alpha_i = 1$.

   On cherche la loi lognormale P de densité f (issue d'une loi $N(m, \sigma^2)$) *approximant* la loi P_0, c'est-à-dire *minimisant* K_{P/P_0}.

   a. Montrer que ce problème équivaut à la recherche des paramètres m et σ solutions du problème :

   $$Min\left[Ln\ \sigma + \frac{1}{2\sigma^2}\sum_{i=1}^{n} \alpha_i \int_0^{+\infty} (Ln\ x - m)^2 f_i(x)\,dx\right]$$

   b. Calculer $\int_0^{+\infty} (Ln\ x - m)^2 f_i(x)\,dx$.

   c. En déduire que le problème à résoudre équivaut à :

   $$\boxed{Min\left[Ln\ \sigma + \frac{1}{2\sigma^2}\sum_{i=1}^{n} \alpha_i [\sigma_i^2 + (m_i - m)^2]\right]}$$

   d. Écrire les conditions du 1er ordre pour la résolution de ce problème de minimisation.

   e. En déduire les valeurs des paramètres solutions des équations obtenues (*on admettra qu'elles définissent bien un minimum*).

   f. Calculer la médiane de la loi P^* réalisant l'optimum en fonction de celles des lois P_i.