ECRICOME 2025

Dans les questions faisant intervenir des instructions en langage **Python** on prendra soin d'importer les bibliothèques nécessaires lors de leur première utilisation.
Pour traiter les questions d'informatique, les candidats sont invités à se référer aux annexes fournies en fin de sujet. Ils ne sont pas limités à l'utilisation des seules fonctions mentionnées dans ces annexes.

EXERCICE 1

On considère la matrice $A = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{pmatrix}$.

On note 0_3 la matrice nulle de $\mathcal{M}_3(\mathbb{R})$ et I_3 la matrice identité de $\mathcal{M}_3(\mathbb{R})$.
Pour toute matrice C de $\mathcal{M}_3(\mathbb{R})$, on note E_C l'ensemble des matrices M de $\mathcal{M}_3(\mathbb{R})$ telles que $CM + MC = 0_3$.

1. Déterminer les ensembles E_{0_3} et E_{I_3}.
2. Montrer que, pour toute matrice C de $\mathcal{M}_3(\mathbb{R})$, l'ensemble E_C est un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$.
3. Soit M une matrice de E_A. Montrer que ${}^tM$ appartient à E_A.
4. 4.a. Justifier que A est diagonalisable.
   4.b. Soit λ un réel.
   Montrer que si λ est valeur propre de A, alors :
   $$\lambda^3 - 9\lambda = 0$$
   4.c. Déterminer une matrice diagonale D de $\mathcal{M}_3(\mathbb{R})$, dont les coefficients diagonaux sont classés dans l'ordre croissant, et une matrice inversible P de $\mathcal{M}_3(\mathbb{R})$, dont les coefficients diagonaux sont tous égaux à 1, telles que $D = P^{-1}AP$.
5. Calculer P^2. En déduire une expression de P^{-1}.
6. 6.a. Soit $N = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ une matrice de $\mathcal{M}_3(\mathbb{R})$.
   Montrer que N appartient à E_D si, et seulement si, $N = \begin{pmatrix} 0 & 0 & c \\ 0 & e & 0 \\ g & 0 & 0 \end{pmatrix}$.
   6.b. En déduire une base $\mathscr{B}$ de l'espace vectoriel E_D et préciser la dimension de E_D.
7. 7.a. Soit $M \in \mathcal{M}_3(\mathbb{R})$. On pose $N = P^{-1}MP$.
   Montrer que M appartient à E_A si, et seulement si, N appartient à E_D.
   7.b. En déduire une base de l'espace vectoriel E_A exprimée à l'aide de P et des matrices de $\mathscr{B}$.
8. Déterminer l'ensemble des matrices M de $\mathcal{M}_3(\mathbb{R})$ telles que $(A+M)^2 = A^2 + M^2$.
   On ne cherchera pas à expliciter les coefficients de M.
9. Soit φ l'endomorphisme de $\mathcal{M}_3(\mathbb{R})$ défini par : $\forall M \in \mathcal{M}_3(\mathbb{R}),\ \varphi(M) = AM + MA$.
   En utilisant certains résultats des questions précédentes, déterminer le rang de φ.

EXERCICE 2

1. 1.a. Montrer que, pour tout entier naturel n, l'intégrale $\displaystyle\int_0^{+\infty} t^n e^{-t}\,dt$ est convergente.
   On note, pour tout entier naturel n, $I_n = \displaystyle\int_0^{+\infty} t^n e^{-t}\,dt$.
   1.b. Calculer I_0 et I_1.
2. Montrer que, pour tout réel x positif, l'intégrale $\displaystyle\int_0^{+\infty} \frac{e^{-t}}{1+xt}\,dt$ est convergente.

On considère la fonction F définie sur $[0;+\infty[$ par :
$$\forall x \in [0;+\infty[,\ F(x) = \int_0^{+\infty} \frac{e^{-t}}{1+xt}\,dt$$

3. Expliciter la valeur de $F(0)$.
4. Soient x et y deux réels positifs tels que $x \leqslant y$. Montrer que $F(y) \leqslant F(x)$.
   Que peut-on en déduire sur la fonction F ?
5. 5.a. Pour tout réel x positif, calculer l'intégrale $\displaystyle\int_0^1 \frac{1}{1+xt}\,dt$. *On distinguera les cas $x = 0$ et $x > 0$.*

5.b. Montrer que, pour tout réel x positif :
$$0 \leqslant \int_0^1 \frac{e^{-t}}{1+xt}dt \leqslant \int_0^1 \frac{1}{1+xt}dt$$

5.c. Montrer que pour tout réel x strictement positif :
$$0 \leqslant \int_1^{+\infty} \frac{e^{-t}}{1+xt}dt \leqslant \frac{1}{x}\int_1^{+\infty} e^{-t}dt$$

5.d. A l'aide des questions précédentes, déterminer la limite de $F(x)$ lors x tend vers $+\infty$.

6. Soit x un réel positif. On admet que l'intégrale $\displaystyle\int_0^{+\infty} \frac{t^2 e^{-t}}{1+xt}dt$ est convergente.

6.a. Montrer que :
$$F(x) - \int_0^{+\infty} e^{-t}(1-xt)dt = x^2\int_0^{+\infty} \frac{t^2 e^{-t}}{1+xt}dt$$

6.b. En déduire que :
$$0 \leqslant F(x) - I_0 + xI_1 \leqslant x^2 I_2$$

7. 7.a. En déduire que la fonction F admet le développement limité à l'ordre 1 suivant au voisinage de 0 :
$$F(x) \underset{x\to 0}{=} 1 - x + o(x)$$

7.b. Montrer que F est dérivable en 0 et déterminer $F'(0)$.

8. On admet que la fonction F est continue sur $[0;+\infty[$.
En tenant compte des propriétés démontrées dans cet exercice, tracer l'allure de la courbe représentative de F. On fera figurer sa tangente au point d'abscisse 0.

Exercice 3

Toutes les variables aléatoires de cet exercice sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Soit n un entier naturel non nul.
La population active d'un territoire est divisée en n catégories socioprofessionnelles, numérotées de 1 à n.
Pour tout entier i compris entre 1 et n, on note X_i la variable aléatoire égale au revenu mensuel, en milliers d'euros, d'un individu choisi au hasard avec équiprobabilité au sein de la catégorie socioprofessionnelle numéro i. On suppose que la variable aléatoire X_i admet pour densité la fonction f_i définie sur $\mathbb{R}$ par :
$$\forall x \in \mathbb{R},\ f_i(x) = \begin{cases} \dfrac{i}{x^{i+1}} & \text{si } x \geqslant 1 \\ 0 & \text{si } x < 1 \end{cases}$$

On note F_i la fonction de répartition de X_i.

Partie I.

1. Vérifier que, pour tout entier i compris entre 1 et n, la fonction f_i est une densité de probabilité.

2. 2.a. Déterminer les entiers naturels, compris entre 1 et n, tels que X_i admet une espérance, et déterminer alors l'espérance de X_i.

2.b. En justifiant la réponse, classer les numéros de catégorie socioprofessionnelle dans l'ordre de leur revenu mensuel moyen, du moins élevé au plus élevé.
On ne considère que les entiers i pour lesquels l'espérance de X_i est bien définie.

3. Montrer que pour tout entier i compris entre 1 et n et pour tout réel x :
$$F_i(x) = \begin{cases} 1 - \dfrac{1}{x^i} & \text{si } x \geqslant 1 \\ 0 & \text{si } x < 1 \end{cases} \tag{1}$$

4. Soit U une variable aléatoire à densité de loi uniforme sur $]0;1[$.
Soit i un entier compris entre 1 et n. On pose $V_i = \dfrac{1}{U^{1/i}}$.

4.a. Montrer que V_i suit la même loi que X_i.

4.b. Écrire une fonction en langage **Python** nommée **simulX**, prenant en argument d'entrée l'entier i, et renvoyant une simulation de la variable aléatoire X_i.

Partie II.

Soit p un réel de $]0;1[$. On choisit un individu au hasard dans la population et on note Y la variable aléatoire égale au numéro de la catégorie socioprofessionnelle à laquelle cet individu appartient. On suppose que la variable aléatoire $Y-1$ suit la loi binomiale $\mathscr{B}(n-1;p)$.

5. En utilisant uniquement la fonction **random** du module **numpy.random**, écrire une fonction en langage **Python** nommée **simulY**, prenant en arguments d'entrée les paramètres n et p, et renvoyant une simulation de la variable aléatoire Y.
6. Recopier et compléter la fonction, en langage **Python**, nommée **loiY**, prenant en arguments d'entrée les paramètres n et p, et renvoyant une liste $(p_1, \ldots, p_n)$ où pour i compris entre 1 et n, p_i est une valeur approchée de $\mathbb{P}\big([Y=i]\big)$.

```python
def loiY(n,p):
    N=10000
    loi=[0]*n
    for k in ......:
        y=simulY(n,p)
        loi[...] ...
    return loi
```

7. Écrire une fonction **Python**, prenant en arguments d'entrée les paramètres n et p, permettant d'afficher un diagramme en bâtons représentant approximativement la loi de Y. On représentera les valeurs prises de Y en abscisses et les probabilités correspondantes en ordonnées. *On pourra utiliser les fonctions définies dans les questions précédentes, et l'annexe fournie en fin de sujet.*
8. Dans cette question, on suppose que chaque profession est identifiée de manière unique par un numéro appelé code PCS.
Par ailleurs, les différentes professions sont regroupées dans six grandes catégories socioprofessionnelles, que l'on identifie par un entier de à 6.
On dispose d'une base de données comportant trois tables nommées **individu**, **departement** et **profession**, décrites ci-dessous.
   - La table **individu** contient des informations sur tous les individus de la population active française. Chaque entrée correspond donc à un individu. La table comporte les attributs suivants.
     - ⋆ **i_nom** (de type **TEXT**) : le nom de l'individu.
     - ⋆ **i_prenom** (de type **TEXT**) : le prénom de l'individu.
     - ⋆ **i_departement** (de type **INTEGER**) : le numéro de département où réside l'individu.
     - ⋆ **i_insee** (de type **INTEGER**) : le numéro INSEE (ou numéro de sécurité sociale) de l'individu.
     - ⋆ **i_code_profession** (de type **INTEGER**) : le code PCS identifiant la profession actuelle de l'individu.
   - La table **departement** contient des informations sur les départements français. Chaque entrée correspond donc à un département. La table comporte les attributs suivants.
     - ⋆ **d_numero** (de type **INTEGER**) : le numéro du département.
     - ⋆ **d_nom** (de type **TEXT**) : le nom du département.
     - ⋆ **d_population** (de type **INTEGER**) : le nombre d'habitants vivant dans le département.
   - La table **profession** contient des informations sur toutes les professions recensées dans la base de données. Chaque entrée correspond donc à une profession différente. La table comporte les attributs suivants.
     - ⋆ **p_pcs** (de type **INTEGER**) : le code PCS permettant d'identifier la profession.
     - ⋆ **p_categorie** (de type **INTEGER**) : le numéro de la catégorie socioprofessionnelle (de 1 à 6) à laquelle la profession se trouve attachée.
     - ⋆ **p_intitule** (de type **TEXT**) : l'intitulé de la profession (par exemple, **chirurgien dentiste**).

   8.a. Que doit vérifier la clé primaire d'une table dans une base de données ?

   8.b. Pour chacune des trois tables de la base de données de cet exemple, indiquer sans justifier un attribut pouvant jouer le rôle de clé primaire.

   8.c. Dresser le schéma relationnel de la base de données décrite ci-dessus, en mettant en évidence les relations qui existent entre les tables et les attributs permettant d'établir ces relations.
   On s'assurera que chaque table est reliée à au moins l'une des deux autres tables.

   8.d. Écrire une requête SQL renvoyant tous les codes PCS des professions exercées dans le département de l'Eure-et-Loir (numéro 28). Chaque code PCS ne pourra apparaître qu'une seule fois.
   *On pourra utiliser la commande **SELECT DISTINCT** décrite dans l'annexe fournie en fin de sujet.*

   8.e. Écrire une requête SQL permettant d'obtenir le numéro de catégorie socioprofessionnelle (entre 1 et 6) de chaque individu.
   La requête devra renvoyer deux attributs pour chaque individu : son numéro INSEE et la catégorie socioprofessionnelle à laquelle est rattachée sa profession.

Partie III.

Soit p un réel de $]0;1[$.
Un institut réalise un sondage selon le protocole suivant :

- On choisit une catégorie socioprofessionnelle de manière aléatoire (mais sans équiprobabilité), et on note Y la variable aléatoire égale au numéro de la catégorie choisie.
Comme dans la partie II, on suppose que $Y-1$ suit la loi binomiale $\mathscr{B}(n-1;p)$.
- On sélectionne alors un individu au hasard (avec équiprobabilité) dans la catégorie socioprofessionnelle choisie à l'étape précédente, et on note Z_n la variable aléatoire égale à son revenu mensuel, en milliers d'euros.

Enfin, on rappelle que, pour tout entier i compris entre 1 et n, la fonction de répartition F_i de la variable aléatoire X_i est donnée par l'égalité (1) montrée à la question 3.
On note G_n la fonction de répartition de la variable aléatoire Z_n.

9. Expliciter $G_n(x)$ pour tout réel x strictement inférieur à 1.

10. Soit x un réel supérieur ou égal à 1.

    10.a. Justifier que pour tout entier i compris entre 1 et n :

$$\mathbb{P}_{[Y=i]}\big([Z_n \leqslant x]\big) = F_i(x)$$

    10.b. Montrer que :

$$G_n(x) = \sum_{k=0}^{n-1} F_{k+1}(x) \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

    10.c. En déduire que :

$$G_n(x) = 1 - \frac{\big(p+(1-p)x\big)^{n-1}}{x^n}$$

11. Justifier que Z_n est une variable aléatoire à densité.

12. En utilisant les fonctions **`simulX`** et **`simulY`** définies aux questions 4.b et 5, écrire une fonction en langage **`Python`** nommée **`sondage`**, prenant en arguments d'entrée les paramètres n et p, et renvoyant une simulation de la variable aléatoire Z_n.

13. Dans cette question uniquement, on suppose $p = \dfrac{1}{n}$.

    13.a. Montrer que, pour tout x réel :

$$G_n(x) = \begin{cases} 1 - \dfrac{1}{x}\left(1 - \dfrac{x-1}{nx}\right)^{n-1} & \text{si } x \geqslant 1 \\ 0 & \text{si } x < 1 \end{cases}$$

    13.b. Montrer que la suite de variables aléatoires $(Z_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire dont on déterminera la fonction de répartition.

Annexe A - Fonctions Python utiles

Manipulation de listes.

- L'opérateur de concaténation +, appliquée entre deux listes, renvoie la liste obtenue en plaçant les éléments de la seconde liste à la suite de ceux de la première liste.
  Par exemple, `[1,2,5]+[4,3]` renvoie la liste `[1,2,5,4,3]`.
- L'opérateur *, appliqué entre une liste `L` et un entier `n`, renvoie la liste obtenue en concaténant `n` fois la liste `L` avec elle-même.
  Par exemple, `[1,4,2]*3` renvoie la liste `[1,4,2,1,4,2,1,4,2]`.
- La fonction `len` prend en argument d'entrée une liste et renvoie le nombre d'éléments dans cette liste.
- La commande `L.append(x)` permet d'inclure l'élément `x` à la fin de la liste `L`.

La bibliothèque numpy.

- Exemple d'importation : `import numpy as np`.
- Les opérations +, `-, *, /, **`, lorsqu'elles sont possibles, peuvent être réalisées entre deux tableaux numpy de dimensions compatibles et agissent alors **coefficient par coefficient**.
- Les fonctions `np.sqrt` (racine carrée), `np.abs` (valeur absolue), `np.log` (logarithme népérien) et `np.exp` (exponentielle) s'appliquent à une quantité numérique ou à un tableau numpy de nombres. Dans ce dernier cas, les fonctions sont appliquées à chaque élément du tableau donné en argument d'entrée.
- La fonction `np.linspace`, prenant en arguments d'entrée deux flottants `a` et `b` et un entier `n`, renvoie un tableau `numpy` contenant `n` éléments régulièrement espacés entre `a` et `b` inclus.

La bibliothèque matplotlib.pyplot.

- Exemple d'importation : `import matplotlib.pyplot as plt`.
- La fonction `plt.plot` prend en arguments d'entrée deux listes `x` et `y` de même longueur ou deux tableaux `numpy` `x` et `y` à une ligne et de même longueur, et renvoie une figure constituée de la ligne brisée joignant les points du plan de coordonnées (x_i, y_i), où x_i et y_i sont respectivement les coefficients des tableaux `x` et `y`.
- La fonction `plt.bar` prend en arguments d'entrée deux listes `x` et `y` de même longueur ou deux tableaux `numpy` `x` et `y` à une ligne et de même longueur, et produit un diagramme en bâtons, les coefficients de `x` indiquant les abscisses auxquelles sont centrés les bâtons, et les coefficients de `y` déterminant la hauteur de chaque bâton en ordonnée.
- La fonction `plt.show`, employée sans argument d'entrée, permet l'affichage d'une figure préalablement tracée, par exemple avec les fonctions `plt.plot` ou `plt.bar`.

Le module numpy.random.

- Exemple d'importation : `import numpy. random as rd`.
- La fonction `rd.random`, appelée sans argument d'entrée, renvoie une simulation d'une variable aléatoire de la loi uniforme sur l'intervalle $[0; 1[$. Il est également possible de spécifier les dimensions d'un tableau `numpy` en argument d'entrée pour obtenir un tableau dont les coefficients sont des simulations de variables aléatoires indépendantes de la loi uniforme sur $[0; 1[$.
- La fonction `rd.randint` prend deux entiers n et p (avec $p > n$) en arguments d'entrée et renvoie une simulation d'une variable aléatoire de la loi uniforme discrète sur $[\![n; p-1]\!]$.

Annexe B - Commande SQL utile

La commande `SELECT DISTINCT`. La commande SQL `SELECT DISTINCT`, utilisée à la place de la commande `SELECT`, permet de ne retenir qu'une occurrence de chaque valeur dans une colonne données, même si cette colonne comporte des valeurs qui se répètent.
Exemple. Une table `etablissement` contient les données suivantes concernant plusieurs établissements scolaires et la ville où ils se situent.

id	nom	ville
1	Edouard Herriot	Livry-Gargan
2	Diderot	Lyon
3	Edouard Herriot	Lyon
4	Louise Michel	Champigny-sur-Marne
5	Diderot	Paris

La requête

```
SELECT DISTINCT nom FROM etablissement;
```

permet d'obtenir les différents noms d'établissements, sans répétition :

```
Edouard Herriot
Diderot
Louise Michel
```

La commande `SELECT DISTINCT` peut éventuellement être combinée avec la commande `WHERE` pour filtrer une partie des données.

★★★★★★★ Fin ★★★★★★★