

Conception : ESCP BS – HEC Paris

MATHÉMATIQUES 2 APPROFONDIES

FILIÈRE ÉCONOMIQUE ET COMMERCIALE

VOIE GÉNÉRALE

Jeudi 25 avril 2024, de 14 h. à 18 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à encadrer dans la mesure du possible les résultats de leurs calculs.
Aucun document n'est autorisé. L'utilisation de toute calculatrice et de tout matériel électronique est interdite. Seule l'utilisation d'une règle graduée est autorisée.
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.

Toutes les variables aléatoires de ce problème sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On rappelle que pour tout x réel, $\lfloor x \rfloor$ désigne la partie entière de x , c'est à dire le nombre entier relatif $\lfloor x \rfloor$ vérifiant $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

Si A est un ensemble fini non vide, on note $\text{Card}(A)$ le nombre d'éléments de A . Si $A = \emptyset$, on convient que $\text{Card}(A) = 0$.

Pour les programmes en Python, on suppose importées :

- la bibliothèque `numpy` sous l'alias `np`,
- la bibliothèque `numpy.random` sous l'alias `rd`,
- la bibliothèque `matplotlib.pyplot` sous l'alias `plt`.

Le mot **FIN** marque la fin de l'énoncé.

Première partie

Soit f une densité de probabilité sur \mathbb{R} , X une variable aléatoire réelle ayant f pour densité et F sa fonction de répartition.

Soit $(X_i)_{i \in \mathbb{N}^*}$ un échantillon de X , c'est-à-dire une suite de variables aléatoires réelles mutuellement indépendantes, de même loi que X .

Pour tout entier $n \geq 1$, on ordonne, pour chaque $\omega \in \Omega$, le n -uplet $(X_1(\omega), X_2(\omega), \dots, X_n(\omega))$ en :

$$X_{(n,1)}(\omega) \leq X_{(n,2)}(\omega) \leq \dots \leq X_{(n,n)}(\omega),$$

et on admet que les applications $X_{(n,k)}$ ainsi définies sont des variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour tous k, n entiers tels que $1 \leq k \leq n$, la variable aléatoire $X_{(n,k)}$ donne la k -ième valeur obtenue en classant dans l'ordre croissant les valeurs prises par les n premières X_i .

On remarque en particulier que $X_{(n,1)} = \min_{1 \leq i \leq n} (X_i)$ et $X_{(n,n)} = \max_{1 \leq i \leq n} (X_i)$.

1. Soit $n \in \mathbb{N}^*$.

(a) Exprimer, à l'aide de F , la fonction de répartition $F_{(n,n)}$ de la variable aléatoire $X_{(n,n)}$. En déduire que $X_{(n,n)}$ admet comme densité la fonction $f_{(n,n)}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{(n,n)}(x) = nf(x)F(x)^{n-1}.$$

(b) Par une méthode similaire, montrer que $X_{(n,1)}$ admet comme densité $f_{(n,1)}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{(n,1)}(x) = nf(x)(1 - F(x))^{n-1}.$$

2. Soit $x \in \mathbb{R}$ et soient n et k deux entiers tels que $1 \leq k \leq n$.

Pour tout $i \in \mathbb{N}^*$, on note $Y_{i,x}$ la variable aléatoire indicatrice de l'événement $[X_i \leq x]$. C'est à dire

$$Y_{i,x} = \begin{cases} 1 & \text{si } X_i \leq x \\ 0 & \text{sinon} \end{cases}$$

(a) Montrer que $Y_{i,x}$ suit une loi de Bernoulli dont on précisera le paramètre.

(b) Montrer l'égalité entre les événements :

$$[X_{(n,k)} \leq x] = \left[\sum_{i=1}^n Y_{i,x} \geq k \right].$$

(c) En déduire que la fonction de répartition $F_{(n,k)}$ de la variable aléatoire $X_{(n,k)}$ est donnée par :

$$\forall x \in \mathbb{R}, \quad F_{(n,k)}(x) = \sum_{\ell=k}^n \binom{n}{\ell} F(x)^\ell (1 - F(x))^{n-\ell}.$$

3. Pour tout triplet d'entiers (n, k, ℓ) tel que $0 \leq k \leq n$ et $0 \leq \ell \leq n$, on introduit les fonctions polynomiales $b_{n,\ell}$ et $a_{n,k}$ définies respectivement par, pour tout $z \in \mathbb{R}$:

$$b_{n,\ell}(z) = \binom{n}{\ell} z^\ell (1 - z)^{n-\ell} \quad \text{et} \quad a_{n,k}(z) = \sum_{\ell=k}^n b_{n,\ell}(z).$$

On convient également que pour tout $n \in \mathbb{N}^*$, la fonction $b_{n-1,n}$ est identiquement nulle.

(a) Soient $n \in \mathbb{N}^*$ et $\ell \in \llbracket 1, n \rrbracket$. Montrer que pour tout $z \in \mathbb{R}$:

$$b'_{n,\ell}(z) = n(b_{n-1,\ell-1}(z) - b_{n-1,\ell}(z)).$$

(b) Soient $n \in \mathbb{N}^*$ et $k \in \llbracket 1, n \rrbracket$. Déduire de la question précédente une expression de $a'_{n,k}$ en fonction de l'une des fonctions $b_{m,\ell}$ avec m et ℓ à déterminer en fonction de n et k .

(c) Soient $n \in \mathbb{N}^*$ et $k \in \llbracket 1, n \rrbracket$. Montrer que la variable aléatoire $X_{(n,k)}$ admet comme densité $f_{(n,k)}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{(n,k)}(x) = n \binom{n-1}{k-1} f(x) F(x)^{k-1} (1 - F(x))^{n-k}.$$

4. Soit x un réel tel que $0 < F(x) < 1$.

(a) Soient $\alpha \in]F(x), 1[$ et $\varepsilon > 0$ tel que $\alpha - F(x) > \varepsilon$.

Montrer que pour tout entier n suffisamment grand, on a l'inclusion d'événements :

$$\left[\sum_{i=1}^n Y_{i,x} \geq \lfloor \alpha n \rfloor \right] \subset \left[\left| \frac{1}{n} \sum_{i=1}^n Y_{i,x} - F(x) \right| \geq \varepsilon \right],$$

et en déduire que :

$$\lim_{n \rightarrow \infty} F_{(n, \lfloor \alpha n \rfloor)}(x) = 0.$$

(b) Soient $\alpha \in]0, F(x)[$ et $\varepsilon > 0$ tel que $F(x) - \alpha > \varepsilon$.

Montrer que pour tout $n \in \mathbb{N}^*$:

$$\left[\sum_{i=1}^n Y_{i,x} \geq \lfloor \alpha n \rfloor \right] \supset \left[\left| \frac{1}{n} \sum_{i=1}^n Y_{i,x} - F(x) \right| < \varepsilon \right],$$

et en déduire que :

$$\lim_{n \rightarrow \infty} F_{(n, \lfloor \alpha n \rfloor)}(x) = 1.$$

5. On suppose pour cette question qu'il existe un intervalle ouvert $I =]a, b[$, avec éventuellement $a = -\infty$ ou $b = +\infty$, tel que f est continue sur I , $f(x) > 0$ pour tout $x \in I$ et $f(x) = 0$ pour tout $x \in \mathbb{R} \setminus I$.

(a) Montrer que la fonction de répartition associée F établit une bijection entre les intervalles I et $]0, 1[$.

(b) On note F^{-1} la bijection réciproque de la fonction F restreint à I et on fixe $\alpha \in]0, 1[$.

Montrer que la suite de variables aléatoires $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en loi vers une variable aléatoire certaine égale à $F^{-1}(\alpha)$.

(c) En déduire que la suite de variables aléatoires $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en probabilité vers une variable aléatoire certaine égale à $F^{-1}(\alpha)$.

6. Dans cette question, on souhaite écrire une fonction Python d'en-tête `def tri(T)` : permettant de trier dans l'ordre croissant les valeurs d'un tableau unidimensionnel T (de type `numpy.array`) sans utiliser de fonction de tri prédéfinie dans Python.

Pour cela, on utilise un algorithme de tri appelé *tri par insertion* et défini de la façon suivante :

- On agit sur les coefficients du tableau T , que l'on suppose numérotés de 0 à $s - 1$.
- Pour chaque k compris entre 1 et $s - 1$, avant l'étape k les coefficients numérotés 0 à $k - 1$ de T ont déjà été classés dans l'ordre croissant, les coefficients suivants étant inchangés. L'étape k consiste alors à insérer la valeur x du coefficient numéroté k de T en bonne position parmi les $k + 1$ premiers coefficients de T afin qu'à l'issue de l'étape k les coefficients numérotés 0 à k de T soient classés dans l'ordre croissant, les coefficients suivants étant inchangés.

Par exemple en partant du tableau T dont le contenu est `[38 28 35 24 31 15]`, les étapes de l'algorithme de tri par insertion donnent successivement :

```
[38 28 35 24 31 15]
[28 38 35 24 31 15]
[28 35 38 24 31 15]
[24 28 35 38 31 15]
[24 28 31 35 38 15]
[15 24 28 31 35 38]
```

Le tri doit être «en place », c'est à dire que le tableau se retrouve trié en fin de fonction au lieu de retourner un autre tableau, version triée du premier.

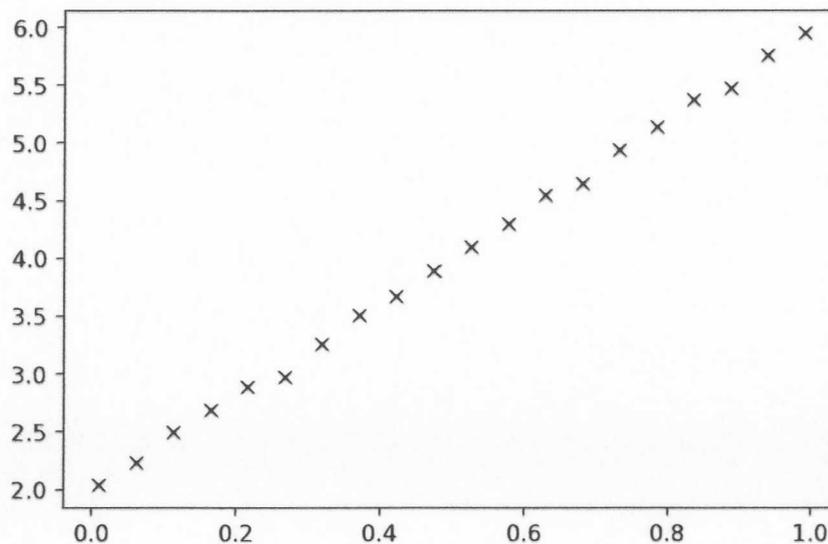
Écrire la fonction `tri` demandée en mettant en œuvre l'algorithme de tri par insertion décrit ci-dessus.

7. On écrit le programme Python suivant à la suite de la fonction `tri` :

```
1 def ech(n):
2     *****
3
4 def X(n,k):
5     res = ech(n)
6     *****
7
8 n = 2000 ; na = 20
9 A = np.linspace(0.01, 0.99, na)
10 B = np.zeros(na)
11 for k in range(na):
12     a = A[k]
13     B[k] = X(n, int(np.floor(a*n)))
14 plt.plot(A, B, "xk")
15 plt.show()
```

La fonction `ech`, que l'on suppose déjà écrite pour les trois questions suivantes (*a*, *b* et *c*), renvoie, sous la forme d'un tableau unidimensionnel de longueur n , une réalisation du n -uplet de variables aléatoires (X_1, X_2, \dots, X_n) .

- (a) Compléter la fonction `X` afin qu'elle renvoie une réalisation de la variable aléatoire $X_{(n,k)}$ lorsqu'on lui fournit en entrée deux entiers de type Python `int` n et k tels que $1 \leq k \leq n$.
On pourra utiliser la fonction `tri`.
- (b) En exécutant le programme, on obtient l'affichage suivant :



À quoi correspondent les points affichés sur cette représentation graphique ?

- (c) Utiliser la représentation graphique pour déterminer la loi commune des X_i simulées par la fonction `ech`.
- (d) Compléter le corps de la fonction `ech` par une ou plusieurs lignes de code pour que le script produise une représentation graphique de valeurs simulées proche de la représentation graphique ci-dessus.

Deuxième partie

On s'intéresse dans cette partie au cas où $(X_i)_{i \in \mathbb{N}^*}$ est une famille de variables aléatoires mutuellement indépendantes suivant toutes la même loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda \in]0, +\infty[$.

8. Soit $n \in \mathbb{N}^*$ fixé.

(a) On note (T_1, \dots, T_n) une famille de variables aléatoires mutuellement indépendantes telles que :

pour tout $i \in \llbracket 1, n \rrbracket$, T_i suit la loi exponentielle $\mathcal{E}((n-i+1)\lambda)$.

Pour tout $k \in \llbracket 1, n \rrbracket$, on note de plus $S_k = T_1 + \dots + T_k$.

Montrer par récurrence sur $k \in \llbracket 1, n \rrbracket$ que la variable aléatoire S_k suit la loi de densité $g_{n,k}$ définie par :

$$\forall x \in \mathbb{R}, \quad g_{n,k}(x) = \begin{cases} n \binom{n-1}{k-1} \lambda e^{-\lambda(n-k+1)x} (1 - e^{-\lambda x})^{k-1} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

(b) Soit $k \in \llbracket 1, n \rrbracket$.

Vérifier que S_k suit la même loi que $X_{(n,k)}$ lorsque les X_i suivent toutes la loi $\mathcal{E}(\lambda)$.

9. Soit (n, k) un couple d'entiers tel que $1 \leq k \leq n$.

Montrer que $X_{(n,k)}$ admet une espérance et une variance données par :

$$\mathbb{E}(X_{(n,k)}) = \frac{1}{\lambda} \sum_{i=1}^k \frac{1}{n-i+1} \quad \text{et} \quad \mathbb{V}(X_{(n,k)}) = \frac{1}{\lambda^2} \sum_{i=1}^k \frac{1}{(n-i+1)^2}.$$

10. (a) Montrer que pour tout entier $j \geq 1$ et tout $x \in [j, j+1]$, on a :

$$0 \leq \frac{1}{j} - \frac{1}{x} \leq \frac{2}{x^2}.$$

(b) En déduire l'encadrement, pour $(n, k) \in \mathbb{N}^2$ tel que $1 \leq k \leq n$:

$$\frac{1}{\lambda} \ln\left(\frac{n+1}{n-k+1}\right) \leq \mathbb{E}(X_{(n,k)}) \leq \frac{1}{\lambda} \ln\left(\frac{n+1}{n-k+1}\right) + \frac{2}{\lambda} \left(\frac{1}{n-k+1} - \frac{1}{n+1}\right).$$

(c) Montrer que pour tout $\alpha \in]0, 1[$:

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_{(n, \lfloor \alpha n \rfloor)}) = -\frac{1}{\lambda} \ln(1 - \alpha).$$

11. Montrer que pour tout $\alpha \in]0, 1[$:

$$\lim_{n \rightarrow \infty} \mathbb{V}(X_{(n, \lfloor \alpha n \rfloor)}) = 0.$$

12. (a) Soit $\varepsilon > 0$. Montrer que :

$$\mathbb{P}\left(\left|X_{(n, \lfloor \alpha n \rfloor)} + \frac{1}{\lambda} \ln(1 - \alpha)\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \mathbb{E}\left(\left(X_{(n, \lfloor \alpha n \rfloor)} + \frac{1}{\lambda} \ln(1 - \alpha)\right)^2\right).$$

(b) En déduire que la suite de variables aléatoires $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en probabilité vers une variable aléatoire certaine égale à $-\frac{1}{\lambda} \ln(1 - \alpha)$.

13. Quelle question de la partie 1 permet de retrouver le résultat précédent ?

14. On souhaite estimer la valeur du paramètre λ à l'aide de l'observation de l'échantillon $(X_i)_{i \in \mathbb{N}^*}$. Pour cela on introduit pour tout $m \in \mathbb{N}^*$ et $t \in]0, +\infty[$, la variable aléatoire :

$$V_{m,t} = \frac{\text{Card}(\{i \in \llbracket 1, m \rrbracket, X_i > t\}) + 1}{m}$$

(a) Montrer que pour tout $t \in]0, +\infty[$, la suite de variables aléatoires $(V_{m,t})_{m \in \mathbb{N}^*}$ converge en probabilité vers une constante que l'on précisera.

(b) Soient $\varepsilon \in \mathbb{R}_+^*$, $m \in \mathbb{N}^*$ et $t \in]0, +\infty[$. Montrer que :

$$\mathbb{P}(|-\ln(V_{m,t}) - \lambda t| \leq \varepsilon) \geq \mathbb{P}(|V_{m,t} - e^{-\lambda t}| \leq e^{-\lambda t}(1 - e^{-\varepsilon}))$$

En déduire que la suite de variables aléatoires $(-\ln(V_{m,t}))_{m \in \mathbb{N}^*}$ converge en probabilité vers une constante que l'on précisera.

(c) On fixe deux réels t_1 et t_2 tels que $0 < t_1 < t_2$.

Pour tout $m \in \mathbb{N}^*$, on pose :

$$W_m = \frac{1}{t_2 - t_1} \ln \left(\frac{\text{Card}(\{i \in \llbracket 1, m \rrbracket, X_i > t_1\}) + 1}{\text{Card}(\{i \in \llbracket 1, m \rrbracket, X_i > t_2\}) + 1} \right)$$

Montrer que W_m est un estimateur convergent de λ .

Troisième partie

15. On admet **pour cette question** le théorème suivant :

Théorème : Soit $n \in \mathbb{N}^*$ fixé.

Il existe une famille $(T_{n,1}, \dots, T_{n,n})$ de variables aléatoires mutuellement indépendantes telles que :

- Pour tout $i \in \llbracket 1, n \rrbracket$, la variable aléatoire $T_{n,i}$ suit la loi exponentielle $\mathcal{E}((n - i + 1)\lambda)$;
- Pour tout $k \in \llbracket 1, n \rrbracket$, on a $X_{(n,k)} = T_{n,1} + \dots + T_{n,k}$.

En quoi ce théorème donne-t-il un résultat plus fort que ce qui a été prouvé en question 8b ?

Le but de cette partie est de démontrer ce théorème dans le cas où $n = 2$.

Soit $q \in]0, 1[$ et soit $p = 1 - q$.

On dit qu'une variable aléatoire Z à valeurs dans \mathbb{N} suit la loi $\mathcal{L}(q)$ si :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(Z = k) = pq^k.$$

16. On suppose que Z suit la loi $\mathcal{L}(q)$. Donner la valeur de l'espérance et de la variance de Z . Déterminer $\mathbb{P}(Z \geq \ell)$ pour tout $\ell \in \mathbb{N}$.

17. Soit $n \in \mathbb{N}^*$ et soit (Z_1, \dots, Z_n) une famille de variables aléatoires mutuellement indépendantes de loi $\mathcal{L}(q)$.

On pose $U_n = \min(Z_1, \dots, Z_n)$.

(a) Calculer $\mathbb{P}(U_n \geq \ell)$ pour tout $\ell \in \mathbb{N}$.

(b) En déduire que U_n suit une loi $\mathcal{L}(r)$ pour un certain paramètre r à déterminer.

18. Soient Z_1, Z_2 deux variables aléatoires indépendantes et de même loi $\mathcal{L}(q)$. Soient k_1, k_2 deux entiers naturels. Montrer que :

$$\mathbb{P}([Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2]) = \frac{1}{1+q} q^{2k_1} q^{k_2}.$$

19. Soient X_1, X_2 deux variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$.

Soient x, y deux réels positifs ou nuls.

(a) Soit m un entier naturel non nul. On pose $Z_1 = \lfloor mX_1 \rfloor$ et $Z_2 = \lfloor mX_2 \rfloor$.

Montrer que les variables aléatoires Z_1 et Z_2 suivent toutes les deux la loi $\mathcal{L}(q)$, où q est un paramètre que l'on précisera.

(b) On conserve les notations de la question précédente et on pose de plus

$k_1 = \lfloor mx \rfloor$ et $k_2 = \lfloor my \rfloor$.

Montrer les inclusions entre événements :

$$[Z_1 \geq k_1 + 1] \subset [X_1 > x] \subset [Z_1 \geq k_1],$$

$$[Z_2 - Z_1 \geq k_2 + 2] \subset [X_2 - X_1 > y] \subset [Z_2 - Z_1 \geq k_2].$$

(c) En déduire l'encadrement :

$$\frac{1}{1 + e^{-\lambda/m}} e^{-2\lambda \frac{\lfloor mx \rfloor + 1}{m}} e^{-\lambda \frac{\lfloor my \rfloor + 2}{m}} \leq \mathbb{P}([X_1 > x] \cap [X_2 - X_1 > y]) \leq \frac{1}{1 + e^{-\lambda/m}} e^{-2\lambda \frac{\lfloor mx \rfloor}{m}} e^{-\lambda \frac{\lfloor my \rfloor}{m}}.$$

(d) Montrer que :

$$\mathbb{P}([X_1 > x] \cap [X_2 - X_1 > y]) = \frac{1}{2} e^{-2\lambda x} e^{-\lambda y}.$$

20. Soient X_1, X_2 deux variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$.

On note $X_{(1)}$ et $X_{(2)}$ les variables aléatoires notées précédemment $X_{(2,1)}$ et $X_{(2,2)}$.

Autrement dit pour tout $\omega \in \Omega$, si $X_1(\omega) \leq X_2(\omega)$, alors $X_{(1)}(\omega) = X_1(\omega)$ et $X_{(2)}(\omega) = X_2(\omega)$, et dans le cas contraire $X_{(1)}(\omega) = X_2(\omega)$ et $X_{(2)}(\omega) = X_1(\omega)$.

(a) Montrer que $\mathbb{P}(X_1 = X_2) = 0$. En déduire que les variables aléatoires $X_{(1)}$ et $X_{(2)} - X_{(1)}$ sont presque-sûrement à valeurs strictement positives.

(b) Montrer que pour tous réels positifs x et y :

$$\mathbb{P}([X_{(1)} > x] \cap [X_{(2)} - X_{(1)} > y]) = 2\mathbb{P}([X_1 > x] \cap [X_2 - X_1 > y]).$$

(c) Montrer que les variables aléatoires $T_1 = X_{(1)}$ et $T_2 = X_{(2)} - X_{(1)}$ sont indépendantes, de loi respective $\mathcal{E}(2\lambda)$ et $\mathcal{E}(\lambda)$.

Conclure quant à l'objectif de cette partie.

FIN

Conception : ESSEC – HEC Paris

MATHÉMATIQUES APPROFONDIES

FILIÈRE ÉCONOMIQUE ET COMMERCIALE

VOIE GÉNÉRALE

Mercredi 24 avril 2024, de 14 h. à 18 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à encadrer dans la mesure du possible les résultats de leurs calculs.
Aucun document n'est autorisé. L'utilisation de toute calculatrice et de tout matériel électronique est interdite. Seule l'utilisation d'une règle graduée est autorisée.
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.

Notations

Dans tout le texte, on adopte les notations suivantes :

- Si A est un ensemble fini non vide, on note $\#A$ le nombre d'éléments de A . Si $A = \emptyset$, on convient que $\#A = 0$.
- Pour tout $n \in \mathbb{N}^*$, on note I_n la matrice identité de $\mathcal{M}_n(\mathbb{R})$.
- Pour tout $(n, m) \in \mathbb{N}^* \times \mathbb{N}^*$ et tout $(i, j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket$ le coefficient sur la i -ème ligne et la j -ème colonne d'une matrice $A \in \mathcal{M}_{n,m}(\mathbb{R})$ est noté $A_{i,j}$.
- La transposée d'une matrice A est notée tA . Lorsque $A = [a] \in \mathcal{M}_1(\mathbb{R})$, où $a \in \mathbb{R}$, on identifie A au réel a . Si $U \in \mathcal{M}_{n,1}(\mathbb{R})$, $n \in \mathbb{N}^*$, on note $\|U\|$ sa norme euclidienne associée au produit scalaire canonique, c'est-à-dire
$$\|U\|^2 = {}^tUU.$$
- Si X est une variable aléatoire réelle, on note $\mathbb{E}(X)$ son espérance et $\mathbb{V}(X)$ sa variance, si elles existent.
- Pour tout $n \in \mathbb{N}^*$, on note \mathcal{O}_n l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$ (on rappelle qu'une matrice $Q \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si ${}^tQQ = I_n$).
- Pour tout $n \in \mathbb{N}$, $\mathbb{R}_n[x]$ désigne l'espace des polynômes à coefficients dans \mathbb{R} de degré inférieur ou égal à n .

- Pour tous $i \in \mathbb{N}$ et $k \in \mathbb{N}$, $\delta_{i,k}$ désigne le symbole de Kronecker défini par : $\delta_{i,k} = \begin{cases} 1 & \text{si } i = k, \\ 0 & \text{si } i \neq k. \end{cases}$
- On note dans tout le problème $\|x\|$ la norme d'un vecteur x , la nature du vecteur entre les doubles barres suffisant à préciser de quelle norme il s'agit.

Pour tout $n \in \mathbb{N}^*$, on adopte les deux définitions suivantes tout au long de l'énoncé :

- On dit qu'une matrice $P \in \mathcal{M}_n(\mathbb{R})$ est une matrice de permutation s'il existe une bijection σ de $\llbracket 1; n \rrbracket$ dans $\llbracket 1; n \rrbracket$ telle que $\forall (i, j) \in \llbracket 1; n \rrbracket^2$, $P_{i,j} = \delta_{i,\sigma(j)}$.
- On dit qu'une matrice est de Hadamard si elle est carrée, si tous ses coefficients appartiennent à $\{-1, 1\}$ et si ses vecteurs colonnes sont deux à deux orthogonaux.

L'énoncé comporte trois parties essentiellement indépendantes.

Le mot **FIN** marque la fin de l'énoncé.

Partie I : existence des matrices de Hadamard.

Soit $n \in \mathbb{N}$ tel que $n \geq 2$.

On suppose dans toute cette partie I qu'il existe une matrice de Hadamard $H \in \mathcal{M}_n(\mathbb{R})$. On note $c_1, \dots, c_n \in \mathcal{M}_{n,1}(\mathbb{R})$ ses vecteurs colonnes et $\ell_1, \dots, \ell_n \in \mathcal{M}_{1,n}(\mathbb{R})$ ses vecteurs lignes.

1. Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice de permutation. Montrer que P est orthogonale et que tP est aussi une matrice de permutation.
2. Soit $Q \in \mathcal{M}_n(\mathbb{R})$. Montrer l'équivalence suivante

$$Q \text{ est une matrice de Hadamard} \Leftrightarrow {}^tQQ = nI_n \text{ et } \forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad Q_{i,j}^2 = 1.$$

3. Montrer que tH est une matrice de Hadamard.
4. Soit $D \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonale ne comportant que les valeurs -1 et 1 sur sa diagonale et $P \in \mathcal{M}_n(\mathbb{R})$ une matrice de permutation. Montrer que PH, DH, HP et HD sont toutes des matrices de Hadamard.
5. En déduire qu'il existe une matrice de Hadamard de taille $n \times n$ dont la première ligne est $(1, \dots, 1)$.
6. En déduire que n est pair.
7. Soit $S \in \mathcal{M}_n(\mathbb{R})$ une matrice de Hadamard telle que $\forall j \in \llbracket 1; n \rrbracket, S_{1,j} = 1$. Montrer que

$$\forall (i, m) \in \llbracket 2; n \rrbracket^2 \quad \text{tels que } i \neq m \quad \sum_{k=1}^n (S_{i,k} + 1)(S_{m,k} + 1) = n.$$

8. En déduire que soit $n = 2$ soit n est un multiple de 4.
9. Dans le cas où $n > 2$, montrer qu'il existe une matrice de Hadamard de taille $n \times n$ dont les trois premières lignes sont de la forme

$$\left(\begin{array}{ccc|ccc|ccc|ccc} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ 1 & \dots & 1 & 1 & \dots & 1 & -1 & \dots & -1 & -1 & \dots & -1 \\ 1 & \dots & 1 & -1 & \dots & -1 & 1 & \dots & 1 & -1 & \dots & -1 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \end{array} \right),$$

(on ne demande pas dans cette question de montrer que les quatre blocs verticaux comportent le même nombre de colonnes).

Indication : on peut observer que la permutation de certaines colonnes (ou la multiplication par -1 de certaines colonnes) dans une matrice de Hadamard donne encore une matrice de Hadamard, d'après la question 4 ci-dessus.

10. Montrer que les quatre blocs de la matrice de Hadamard de la question précédente comportent le même nombre de colonnes et en déduire à nouveau que n est nécessairement divisible par 4.
11. Pour toutes matrices $A \in \mathcal{M}_n(\mathbb{R})$ et $B \in \mathcal{M}_k(\mathbb{R})$ où $n \geq 1$ et $k \geq 1$, on note $A \otimes B$ la matrice appartenant à $\mathcal{M}_{kn}(\mathbb{R})$ définie bloc par bloc comme suit

$$A \otimes B = \begin{pmatrix} A_{1,1}B & \cdots & A_{1,n}B \\ A_{2,1}B & \cdots & A_{2,n}B \\ \vdots & \vdots & \vdots \\ A_{n,1}B & \cdots & A_{n,n}B \end{pmatrix},$$

(ainsi, chacun des blocs $A_{i,j}B$ est de taille $k \times k$).

Montrer que si A et B sont des matrices de Hadamard, alors $A \otimes B$ l'est aussi.

12. Montrer que pour tout $m \in \mathbb{N}^*$, il existe une matrice de Hadamard dans $\mathcal{M}_{2^m}(\mathbb{R})$.
13. La fonction `test_hadamard()` ci-dessous est écrite en PYTHON. Elle est incomplète et a comme paramètre d'entrée une matrice M de coefficients entiers représentée par un tableau bidimensionnel (de type array)

Compléter les parties soulignées en pointillé afin que la fonction `test_hadamard()`

- renvoie la valeur -2 si la matrice M n'est pas carrée,
- renvoie la valeur -1 si la matrice M est carrée mais au moins l'un de ses coefficients n'appartient pas à $\{-1, 1\}$,
- renvoie la valeur 0 si la matrice M est carrée et ses coefficients appartiennent tous à $\{-1, 1\}$ mais M n'est pas une matrice de Hadamard,
- renvoie la valeur 1 si M est une matrice de Hadamard.

On reproduira sur la copie le programme après l'avoir complété.

```
import numpy as np
def test_hadamard(M):
    n, p = np.shape(M)
    if _____:
        return -2
    else:
        for i in range(0, _____):
            for j in range(0, _____):
                if _____ != 1 :
                    return -1
            for j in range(0, _____):
                for k in range(_____, n):
                    if np.dot(_____) != 0:
                        return 0
    return 1
```

14. On voudrait maintenant écrire en PYTHON une fonction `rand_hadam()` qui cherche une matrice de Hadamard dont les coefficients de la première ligne sont tous égaux à 1 (ainsi, chacune des autres lignes comporte autant de 1 que de -1) et de taille $4m \times 4m$ où m est un entier donné en paramètre. Cette fonction a aussi un autre paramètre d'entrée `Nmax` désignant le nombre maximal de matrices à tester. On procède comme suit :
- (a) On construit aléatoirement une matrice ne contenant que des -1 et 1 comme coefficients de la manière suivante : pour chaque ligne (à partir de la deuxième ligne), on choisit successivement et aléatoirement les coefficients dans l'ensemble $\{-1, 1\}$ jusqu'à ce que le nombre de la valeur 1 ou le nombre de la valeur -1 dans la ligne atteint $n/2 = 2m$. Si le nombre de coefficients égaux à 1 (respectivement -1) dans la ligne atteint $n/2$, on attribue à tous les coefficients restants de la ligne la valeur -1 (respectivement $+1$).

- (b) En utilisant la fonction `test_hadamard()` on teste si la matrice ainsi construite est de Hadamard. Si oui, la fonction renvoie cette matrice. Sinon, on refait la construction d'une matrice de la même manière que dans (a) et cela autant de fois que nécessaire jusqu'à l'obtention d'une matrice de Hadamard, sans toutefois dépasser un nombre maximal de matrices testées désigné ici par `Nmax`. Si ce nombre maximal est atteint sans trouver une matrice de Hadamard, la fonction `rand_hadam()` renvoie la matrice nulle de taille $4m \times 4m$.

Compléter les parties soulignées en pointillé de la fonction `rand_hadam()` ci-dessous. *On reproduira sur la copie le programme après l'avoir complété (sans les commentaires).*

```
import numpy.random as rd
def rand_hadam(m, Nmax):
    n = 4*m
    for tst in range(0, Nmax):
        matpm = np.ones((n, n), dtype = int)
        for i in range(_____, n):
            nb_un = 0
            j = 0
            while 2*nb_un < n and _____ < n :
                val = rd.randint(0,2)
                nb_un += _____
                matpm[i, j] = 2*val - 1
                j = _____
            if (2*nb_un == n):
                for k in range(j, n):
                    matpm[i, k] = _____
        if (test_hadamard(matpm) == 1):
            return _____
    return np.zeros((n, n), dtype = int)
```

Partie II : variables aléatoires deux à deux indépendantes sur un espace probabilisé fini.

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé où Ω désigne un univers fini, $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω et \mathbb{P} une probabilité. On pose $\Omega = \{\omega_1, \dots, \omega_n\}$ où l'entier n , $n \geq 2$, désigne ici et dans la suite de cette partie le nombre d'éléments de Ω . On suppose que $\mathbb{P}(A) \neq 0$ pour tout $A \subset \Omega$ non vide et on pose

$$\forall i \in \llbracket 1; n \rrbracket \quad p_i = \mathbb{P}(\{\omega_i\}).$$

Toutes les variables aléatoires considérées dans la suite de cette partie II sont définies sur cet espace probabilisé. On suppose qu'il existe un entier naturel $\ell \geq 2$ et des variables aléatoires réelles Z_1, \dots, Z_ℓ vérifiant les deux propriétés suivantes

- $\forall (i, j) \in \llbracket 1; \ell \rrbracket^2$ avec $i \neq j$, Z_i et Z_j sont indépendantes (autrement dit, Z_1, \dots, Z_ℓ sont deux à deux indépendantes),
- $\forall i \in \llbracket 1; \ell \rrbracket$, $\#Z_i(\Omega) \geq 2$.

Pour tout $i \in \llbracket 1; \ell \rrbracket$, on pose $z_i = (Z_i(\omega_1), \dots, Z_i(\omega_n)) \in \mathbb{R}^n$

Dans toute la suite, pour tous vecteurs $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ et $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ on note

$$\langle u; v \rangle = \sum_{k=1}^n p_k u_k v_k.$$

On pose dans la suite $u_0 = (1, \dots, 1) \in \mathbb{R}^n$.

15. Montrer que l'application $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \langle x; y \rangle$ est un produit scalaire sur \mathbb{R}^n .

16. Soit $i \in \llbracket 1; \ell \rrbracket$. Montrer qu'il existe un et un seul couple $(a_i, b_i) \in \mathbb{R}_+^* \times \mathbb{R}$ vérifiant

$$\mathbb{E}(a_i Z_i + b_i) = 0 \text{ et } \mathbb{V}(a_i Z_i + b_i) = 1.$$

Dans la suite de cette partie, on pose pour tout $i \in \llbracket 1; \ell \rrbracket$: $X_i = a_i Z_i + b_i$ et $x_i = (X_i(\omega_1), \dots, X_i(\omega_n)) \in \mathbb{R}^n$, où le couple (a_i, b_i) est celui de la question 16.

17. Soient V et W deux variables aléatoires réelles sur Ω .

On pose $v = (V(\omega_1), \dots, V(\omega_n)) \in \mathbb{R}^n$, $w = (W(\omega_1), \dots, W(\omega_n)) \in \mathbb{R}^n$. Montrer que

$$\mathbb{E}(V) = \langle v; u_0 \rangle \quad \text{et} \quad \mathbb{E}(VW) = \langle v; w \rangle.$$

18. En déduire les relations pour tout $(i, j) \in \llbracket 1; \ell \rrbracket^2$ tels que $i \neq j$

$$\langle x_i; u_0 \rangle = 0 \quad , \quad \langle x_i; x_i \rangle = 1 \quad \text{et} \quad \langle x_i; x_j \rangle = 0$$

19. En déduire que

$$3 \leq \ell + 1 \leq n.$$

20. Soit Z une variable aléatoire réelle d'espérance nulle. On pose $z = (Z(\omega_1), \dots, Z(\omega_n)) \in \mathbb{R}^n$.

On suppose que $Z(\Omega) = \{\alpha_1, \dots, \alpha_m\}$ où $m = \#Z(\Omega) > 2$.

(a) Montrer qu'il existe des m réels β_1, \dots, β_m non tous nuls tels que

$$\sum_{k=1}^m \mathbb{P}(Z = \alpha_k) \beta_k = 0 \quad \text{et} \quad \sum_{k=1}^m \mathbb{P}(Z = \alpha_k) \alpha_k \beta_k = 0$$

(b) Montrer que l'application $T : Q \in \mathbb{R}_{m-1}[x] \mapsto (Q(\alpha_1), \dots, Q(\alpha_m)) \in \mathbb{R}^m$ est un isomorphisme de $\mathbb{R}_{m-1}[x]$ dans \mathbb{R}^m .

(c) En déduire qu'il existe un polynôme Q de degré inférieur ou égal à $m - 1$ tel que

$$\mathbb{E}(Q(Z)) = 0, \quad \mathbb{E}(Q(Z)Z) = 0, \quad \text{et} \quad Q(Z)(\Omega) \neq \{0\}.$$

21. On pose $r = \#\{i \in \llbracket 1, \ell \rrbracket \mid \#X_i(\Omega) > 2\}$.

(a) Montrer que

$$\ell + r \leq n - 1.$$

(b) En déduire que

$$r \leq \frac{n-1}{2}.$$

22. On suppose de plus dans cette question que $\ell = n - 1$ (on rappelle que nécessairement $n \geq 3$). On considère la matrice carrée réelle M de taille $n \times n$ dont les coefficients sont définis par

$$\forall (i, j) \in \llbracket 1; n-1 \rrbracket \times \llbracket 1; n \rrbracket \quad M_{i,j} = X_i(\omega_j) \text{ et } M_{n,j} = 1,$$

On considère aussi la matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$ définie par $\forall i \in \llbracket 1; n \rrbracket \quad D_{i,i} = \sqrt{p_i}$.

(a) Montrer que : $\forall i \in \llbracket 1; n-1 \rrbracket \quad \#X_i(\Omega) = 2$.

(b) Soit $i \in \llbracket 1; n-1 \rrbracket$.

Soient α_i et β_i deux réels tels que $X_i(\Omega) = \{\alpha_i, \beta_i\}$ et $\alpha_i > \beta_i$ et soit $\theta_i \in]0, 1[$ tel que

$$\mathbb{P}(X_i = \alpha_i) = \theta_i \text{ et } \mathbb{P}(X_i = \beta_i) = 1 - \theta_i.$$

Montrer les deux relations :

$$\alpha_i = \sqrt{\frac{1-\theta_i}{\theta_i}}, \quad \beta_i = -\frac{1}{\alpha_i}.$$

(c) Montrer que la matrice MD est orthogonale.

(d) Soit Y la variable aléatoire définie par $\forall j \in \llbracket 1; n \rrbracket \quad Y(\omega_j) = \frac{1}{p_j}$. Montrer que

$$Y = 1 + \sum_{k=1}^{n-1} X_k^2$$

(e) En déduire que pour tout $i \in \llbracket 1; n-1 \rrbracket$

$$\mathbb{E}(X_i^3) = \sum_{k=1}^n X_i(\omega_k).$$

(f) On reprend les notations de la question 22b et on pose

$$\forall i \in \llbracket 1; n-1 \rrbracket \quad m_i = \#\{k \in \llbracket 1, n \rrbracket \mid X_i(\omega_k) = \alpha_i\}.$$

Montrer que les trois affirmations suivantes sont nécessairement vraies :

(A1) pour tout $i \in \llbracket 1; n-1 \rrbracket$, $\alpha_i = \sqrt{\frac{n-m_i-1}{m_i-1}}$ et $\theta_i = \frac{m_i-1}{n-2}$.

(A2) $n \geq 4$,

(A3) $2 \leq m_i \leq n-2$,

(g) On suppose dans cette question que pour tout $i \in \llbracket 1; n \rrbracket \quad p_i = \frac{1}{n}$.

Montrer les deux assertions suivantes :

(a) n est nécessairement pair et $\theta_i = 1/2$.

(b) M est une matrice de Hadamard.

Partie III : deux propriétés des matrices de Hadamard

Soit $n \geq 3$ un entier naturel. On considère la fonction

$$F : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \rightarrow \mathbb{R} \\ A & \mapsto \sum_{i=1}^n \sum_{j=1}^n |A_{i,j}|. \end{cases}$$

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, on note $\mathbb{S}(A)$ la matrice appartenant à $\mathcal{M}_n(\mathbb{R})$ définie par

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad \mathbb{S}(A)_{i,j} = \text{sgn}(A_{i,j})$$

où sgn est la fonction signe définie par : $\text{sgn}(x) = 1$ si $x \geq 0$ et $\text{sgn}(x) = -1$ si $x < 0$.

On considère le produit scalaire $\langle \cdot; \cdot \rangle$ défini sur $\mathcal{M}_n(\mathbb{R})$ par :

$$\forall A, B \in \mathcal{M}_n(\mathbb{R}), \langle A; B \rangle = \sum_{i=1}^n \sum_{j=1}^n A_{i,j} B_{i,j},$$

ainsi que la norme associée $\|\cdot\|$ définie par

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \|A\| = \sqrt{\langle A; A \rangle} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n A_{i,j}^2}.$$

On note φ l'application définie

$$\varphi : \begin{cases} \mathbb{R}^{n^2} & \rightarrow \mathcal{M}_n(\mathbb{R}) \\ (x_1, \dots, x_{n^2}) & \mapsto [x_{(i-1)n+j}]_{(i,j) \in \llbracket 1; n \rrbracket^2} \end{cases}$$

\mathbb{R}^{n^2} est un espace euclidien muni du produit scalaire canonique.

23. Montrer que φ est une application linéaire de \mathbb{R}^{n^2} dans $\mathcal{M}_n(\mathbb{R})$ et que

$$\forall x \in \mathbb{R}^{n^2}, \|\varphi(x)\| = \|x\|.$$

24. Montrer que pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on a $\|A\|^2 = \text{Tr}({}^tAA)$.

25. Montrer que pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on a

$$\sum_{i=1}^n \left(\sum_{j=1}^n A_{i,j}^2 \right)^{1/2} \leq F(A) \leq n \left(\sum_{i=1}^n \sum_{j=1}^n A_{i,j}^2 \right)^{1/2}. \quad (1)$$

26. On pose $\mathcal{K}_n = \{x \in \mathbb{R}^{n^2} \mid \varphi(x) \in \mathcal{O}_n\}$. Montrer que \mathcal{K}_n est un fermé borné de \mathbb{R}^{n^2} .

27. Montrer que

$$\forall x \in \mathbb{R}^{n^2}, \forall y \in \mathbb{R}^{n^2}, |F(\varphi(x)) - F(\varphi(y))| \leq n\|x - y\|.$$

28. En déduire que F admet un minimum global et un maximum global (atteints) sur \mathcal{O}_n .

Dans toute la suite, on note M_n^- (respectivement M_n^+) la valeur minimale (respectivement maximale) de F sur \mathcal{O}_n :

$$M_n^- = \min_{Q \in \mathcal{O}_n} F(Q), \quad M_n^+ = \max_{Q \in \mathcal{O}_n} F(Q).$$

29. Montrer que $M_n^- = n$.

30. Trouver toutes les matrices $Q \in \mathcal{O}_n$ qui vérifient $F(Q) = n$.

31. Montrer que pour toute matrice $Q \in \mathcal{O}_n$ on a :

$$F(Q) = n\sqrt{n} - \frac{\sqrt{n}}{2} \|Q - \frac{1}{\sqrt{n}}\mathbb{S}(Q)\|^2.$$

32. En déduire que $M_n^+ \leq n\sqrt{n}$.

Montrer que l'égalité est réalisée si et seulement s'il existe au moins une matrice de Hadamard dans $\mathcal{M}_n(\mathbb{R})$.

33. En déduire que si n n'est pas un multiple de 4 alors $M_n^+ < n\sqrt{n}$.

34. On suppose dans cette question que $n = 3$. Pour tous réels α et β on pose

$$U(\alpha, \beta) = \begin{bmatrix} \alpha & \beta & \beta \\ \beta & \alpha & -\beta \\ -\beta & \beta & -\alpha \end{bmatrix}.$$

Trouver tous les couples (α, β) pour lesquels $U(\alpha, \beta) \in \mathcal{O}_3$ et en déduire un encadrement de M_3^+ (on note que $3\sqrt{3} \approx 5,196152$).

35. On suppose dans cette question que n est pair.

(a) Soit $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ deux vecteurs de \mathbb{R}^n tels que $\forall i \in \llbracket 1; n \rrbracket \quad |x_i| = |y_i| = 1$ et $\langle x; y \rangle \neq 0$. Montrer que $|\langle x; y \rangle| \geq 2$.

(b) Soit x, y, u et v quatre vecteurs de \mathbb{R}^n . Montrer que

$$|\langle x; y \rangle - \langle u; v \rangle| \leq \sqrt{2(\|y\|^2\|x - u\|^2 + \|u\|^2\|y - v\|^2)}.$$

(c) Soit $Q \in \mathcal{O}_n$ telle que $\mathbb{S}(Q)$ n'est pas de Hadamard. Montrer que

$$\|Q - \frac{1}{\sqrt{n}}\mathbb{S}(Q)\|^2 \geq \frac{2}{n^2}.$$

(d) En déduire que s'il n'existe aucune matrice de Hadamard dans $\mathcal{M}_n(\mathbb{R})$, alors

$$M_n^+ \leq n\sqrt{n} - \frac{1}{n\sqrt{n}}.$$

FIN

