PROBLÈME N°1

Construction de triangles

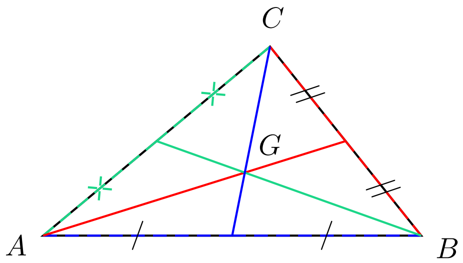
Dans un plan affine euclidien orienté, on considère deux points distincts B et C et un point M n'appartenant pas à la droite (BC).

Question 1.

Pour chacune des assertions suivantes, déterminer s'il existe un point A qui la vérifie.
On précisera pour chaque cas le nombre de solutions et on prendra soin de fournir toutes les explications et justifications utiles.

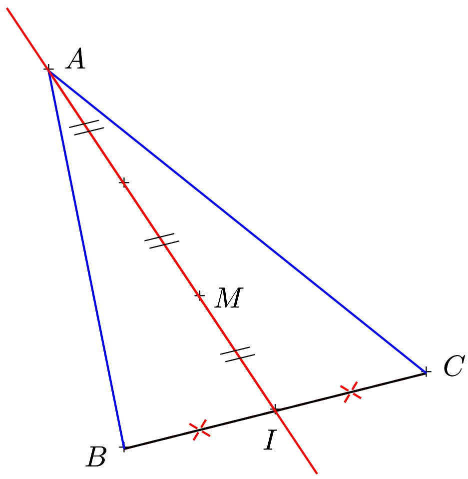
1.1.
M est le centre de gravité du triangle ABC.

Propriété 1
Les trois médianes d'un triangle sont concourantes en un point G qui est le **centre de gravité** du triangle.
G se situe aux deux tiers de chaque médiane à partir du sommet.



Analyse : supposons qu'un tel triangle existe, on note I le milieu du segment $[BC]$. Le point M, centre de gravité du triangle, est situé aux deux tiers de la médiane de sommet A, on a donc $\overrightarrow{AM} = \frac{2}{3}\overrightarrow{AI}$, ou encore $\overrightarrow{IA} = 3\overrightarrow{IM}$. Le point A ainsi déterminé est unique.

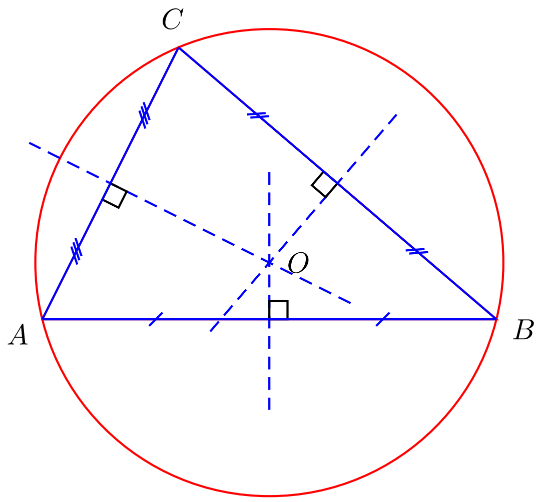
Synthèse : soit I le milieu de $[BC]$, on trace la droite (IM) (elle existe puisque M n'appartient pas à la droite (BC)).
Sur cette droite, on place le point A tel que $\overrightarrow{IA} = 3\overrightarrow{IM}$. Le point M vérifie alors $\overrightarrow{AM} = \frac{2}{3}\overrightarrow{AI}$, M est donc le centre de gravité du triangle ABC ainsi construit.



1.2.

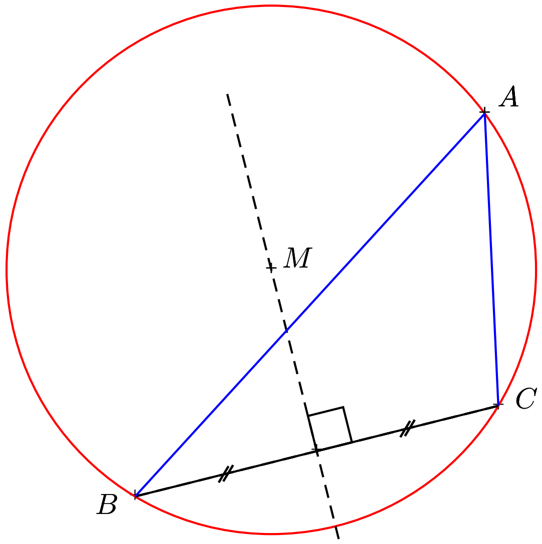
M est le centre du cercle circonscrit au triangle ABC.

Propriété 2
Les trois médiatrices d'un triangle sont concourantes en un point O qui est le **centre du cercle circonscrit** à ce triangle.



Analyse : supposons qu'un tel triangle existe, alors M appartient à la médiatrice du segment $[BC]$ et tout point A du cercle de centre M passant par B vérifie $MA = MB = MC$.
Donc, on peut construire une infinité de points A.

Synthèse : si M n'est pas sur la médiatrice de $[BC]$, la construction n'est pas possible.
Si M est sur la médiatrice de $[BC]$, on trace le cercle de centre M passant par B et on place sur ce cercle un point A (distinct de B et de C). Le point M vérifie alors $MA = MB = MC$ et M est le centre du cercle circonscrit au triangle ABC.

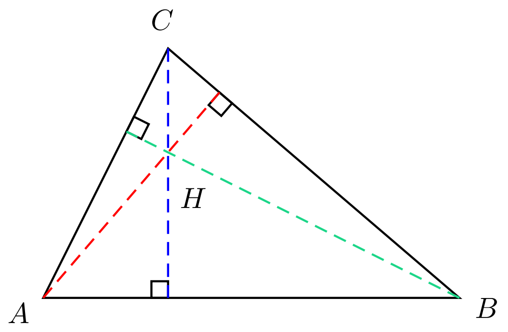


1.3.

M est l'orthocentre du triangle ABC.

Propriété 3

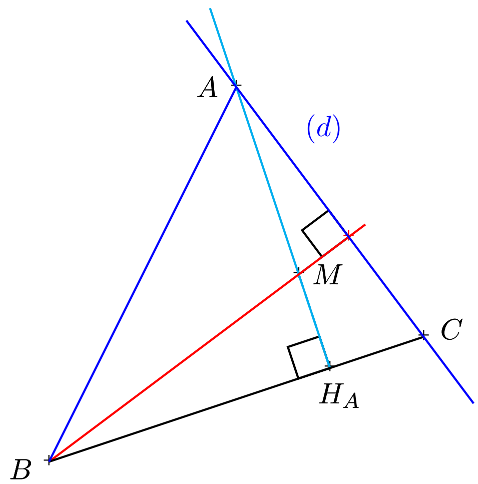
Les trois hauteurs d'un triangle sont concourantes en un point H appelé **l'orthocentre** du triangle.



Analyse : si M est l'orthocentre du triangle ABC, notons H_A le pied de la hauteur issue de A sur (BC). Le point A appartient à le droite (MH_A) et on a également (BM) perpendiculaire à (AC) et (CM) perpendiculaire à (AB).

Synthèse : on place le projeté orthogonal H_A de M sur la droite (BC) : (MH_A) est perpendiculaire à (BC). On trace la droite (d) perpendiculaire à (BM) passant par C, elle coupe la droite (MH_A) au point A. On a ainsi : (BM) perpendiculaire à (AC) ; (AM) perpendiculaire à (BC). M est donc l'orthocentre du triangle ABC. Le point A ainsi déterminé est unique.

Remarquons que si le triangle MBC est rectangle en M, la construction fonctionne encore, mais dans ce cas, on a M et A confondus.



Remarque culturelle.

La construction est toujours possible grâce à la propriété dite du « quadrilatère orthocentrique » : si le point H est l'orthocentre du triangle ABC, alors chaque triangle formé à partir de trois des quatre points A, B, C, H admet le quatrième point comme orthocentre.

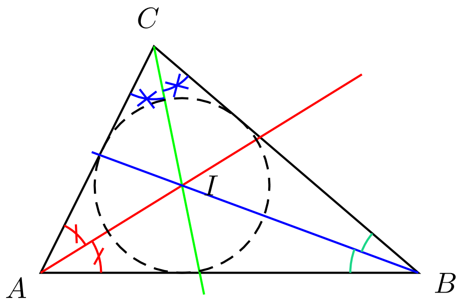
Donc, ici, construire A c'est chercher l'orthocentre du triangle BMC. On peut aller voir une animation faite par Yves Martin avec DGPad, où l'on peut déplacer l'orthocentre sur l'un des sommets du triangle et apercevoir en bleu le triangle dont il est l'orthocentre : `http://goo.gl/QPgzOe`

1.4.

M est le centre du cercle inscrit au triangle ABC.

Propriété 4

Les trois bissectrices d'un triangle sont concourantes en un point I qui est le **centre du cercle inscrit** dans le triangle.

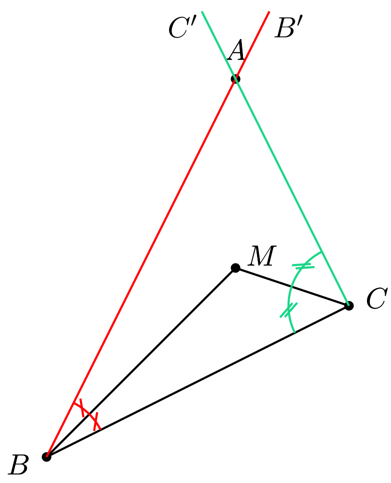


Analyse : si M est le centre du cercle inscrit au triangle ABC, on a en particulier les inégalités d'angles géométriques suivantes : $0° < \widehat{CBM} = \widehat{MBA} < 90°$ et $0° < \widehat{BCM} = \widehat{MCA} < 90°$.

De plus, on a
$$\begin{aligned}\widehat{BMC} &= 180° - \left(\widehat{MCB} + \widehat{CBM}\right)\\ &= 180° - \left(\frac{1}{2}\widehat{ACB} + \frac{1}{2}\widehat{CBA}\right)\\ &= 180° - \frac{1}{2}\left(180° - \widehat{BAC}\right)\\ &= 90° + \frac{1}{2}\widehat{BAC}.\end{aligned}$$

Or, $0° < \widehat{BAC} < 180°$ donc, $90° < \widehat{BMC} < 180°$.

Synthèse : on trace la droite (BB') symétrique de la droite (BC) par rapport à la droite (BM).
Puis la droite (CC') symétrique de la droite (CB) par rapport à la droite (CM).
Les droites (BB') et (CC') sont sécantes en un point A tel que M soit à l'intérieur du triangle ABC lorsque $90° < \widehat{BMC} < 180°$.
Ceci est réalisé lorsque M est strictement à l'intérieur du cercle de diamètre $[BC]$.
Dans ce cas, (BM) est la bissectrice intérieure de l'angle $\widehat{CBA}$ et (CM) est la bissectrice intérieure de l'angle $\widehat{BCA}$. Les bissectrices se coupent au centre du cercle inscrit au triangle ABC. Le point A ainsi construit est unique.
Si M est à l'extérieur ou sur le cercle de diamètre $[BC]$, il n'y a pas de construction possible.



Autre remarque culturelle.
Concernant la dernière phrase, on peut préciser un peu : si M est sur le cercle de diamètre $[BC]$, les droites construites sont parallèles donc il n'y a pas de solution (sauf en géométrie projective !).
Par contre, si M est à l'extérieur du cercle, les deux droites sont sécantes mais l'intersection ne peut pas être le centre du cercle inscrit, il s'agit du centre d'un des cercles exinscrits, celui opposé au point A.

Question 2.

Quel type de raisonnement a été utilisé dans les questions précédentes ?
Donner un autre exemple de problème pouvant être résolu à l'aide de cette méthode.

Il s'agit d'un raisonnement par analyse-synthèse : il permet de déterminer l'ensemble des objets vérifiant des propriétés données. on suppose que l'ensemble n'est pas vide et on cherche des conditions nécessaires vérifiées par ses éléments. Il se décompose en deux phases.

- la phase d'analyse : on suppose que l'ensemble n'est pas vide et on cherche des conditions nécessaires vérifiées par ses éléments ;
- la phase de synthèse : on étudie réciproquement si les objets qui satisfont aux conditions nécessaires dégagées par l'analyse vérifient ou non les propriétés données.

Ce raisonnement s'utilise généralement lorsque l'on veut déterminer l'ensemble des éléments d'un ensemble E qui satisfont une certaine propriété.

Nous avons ici un exemple d'utilisation en géométrie, on peut donner un exemple en analyse :

Montrer que toute fonction f définie sur un sous-ensemble $\mathcal{D}$ symétrique par rapport à 0 de $\mathbb{R}$ s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire.

Analyse : supposons que $f = g + h$ où $g : \mathcal{D} \to \mathbb{R}$ est une fonction paire et $h : \mathcal{D} \to \mathbb{R}$ une fonction impaire, on a :

$$\begin{aligned}\forall x \in \mathcal{D}, \frac{f(x)+f(-x)}{2} &= \frac{g(x)+h(x)+g(-x)+h(-x)}{2} \\ &= \frac{g(x)+\cancel{h(x)}+g(x)-\cancel{h(x)}}{2} \\ &= g(x)\end{aligned}$$

De la même manière, on a $\dfrac{f(x)-f(-x)}{2} = h(x)$.

Ainsi, si cette décomposition existe, elle est unique.

Synthèse : soient g et h les fonctions définies sur $\mathcal{D}$ par : $g(x) = \dfrac{f(x)+f(-x)}{2}$ et $h(x) = \dfrac{f(x)-f(-x)}{2}$.

On a $\forall x \in \mathcal{D}$:

- $g(x)+h(x) = \dfrac{f(x)+\cancel{f(-x)}}{2} + \dfrac{f(x)-\cancel{f(-x)}}{2} = f(x)$;
- g est paire puisque $g(-x) = \dfrac{f(-x)+f(x)}{2} = g(x)$;
- h est paire puisque $h(-x) = \dfrac{f(-x)-f(x)}{2} = -h(x)$.

D'où l'existence de la décomposition !

PROBLÈME N°2

Quelques nombres irrationnels

Bref aperçu historique...

Etymologiquement, un nombre *irrationnel* est un nombre que l'on ne peut pas compter, c'est à dire que l'on ne peut pas écrire car le nombre de décimales qui le constitue est infini et ces décimales se suivent sans suite logique. On l'oppose au nombre *rationnel*, quotient de deux entiers dont l'écriture décimale peut être infinie mais dans ce cas nécessairement périodique.

Des nombres irrationnels célèbres sont $\sqrt{2}$, π, ϕ, e.

Les premières traces de nombres irrationnels remontent aux civilisations babyloniennes (environ 1800 avant J.C) : une tablette de l'époque représente un carré et ses diagonales, et indique une approximation de $\sqrt{2}$ avec quatre chiffres sexagésimaux significatifs (six chiffres décimaux).

L'ensemble des nombres rationnels est noté $\mathbb{Q}$.
On rappelle que tout nombre rationnel non nul peut s'écrire sous la forme $\dfrac{p}{q}$, où p et q sont des entiers relatifs premiers entre eux. Un nombre réel est dit irrationnel s'il n'appartient pas à $\mathbb{Q}$.
Dans ce problème, on se propose de démontrer l'irrationalité de quelques nombres réels.
Les deux parties de ce problème sont indépendantes.

Partie A : quelques exemples de nombres irrationnels

Question 1.

Soit n un entier naturel. Démontrer que si $\sqrt{n}$ n'est pas entier, alors il est irrationnel.

Soit n un entier naturel tel que $\sqrt{n}$ ne soit pas entier. Supposons alors que $\sqrt{n}$ soit rationnel :
$\exists\,(p,q) \in \mathbb{N} \times \mathbb{N}^*$ tels que $\sqrt{n} = \dfrac{p}{q}$ avec $\text{pgcd}(p,q) = 1$.

Or, $\left\{\begin{array}{l} \sqrt{n} = \dfrac{p}{q} \\ \text{pgcd}(p,q) = 1 \end{array}\right. \implies \left\{\begin{array}{l} n = \dfrac{p^2}{q^2} \in \mathbb{N} \\ \text{pgcd}(p^2,q^2) = 1 \end{array}\right.$

d'où $q^2 = 1$ et $n = p^2$, et donc $\sqrt{n} = p$ est un entier, ce qui est contraire à notre hypothèse.

Conclusion : $\boxed{\textit{Si } \sqrt{n} \textit{ n'est pas entier, alors il est irrationnel.}}$

Question 2.

En déduire que si p désigne un nombre premier, alors $\sqrt{p}$ est irrationnel.

Si p est un nombre premier, $p \geq 2$, supposons que $\sqrt{p} = N$ soir un entier, alors $p = N^2$ n'est pas un nombre premier puisqu'il est divisible par $1, N$ et p.

D'où $\sqrt{p}$ n'est pas entier, et d'après la question 1, $\sqrt{p}$ est irrationnel.

Conclusion : $\boxed{\textit{Si } p \textit{ est un nombre premier, alors } \sqrt{p} \textit{ est irrationnel.}}$

Question 3.

Démontrer que le nombre $\dfrac{\ln 2}{\ln 3}$ est irrationnel.

Supposons que $\dfrac{\ln 2}{\ln 3} > 0$ soit rationnel, $\exists\,(p,q) \in \mathbb{N}^* \times \mathbb{N}^*$ tel que $\dfrac{\ln 2}{\ln 3} = \dfrac{p}{q}$ avec $\mathrm{pgcd}(p,q) = 1$.

Or, $\dfrac{\ln 2}{\ln 3} = \dfrac{p}{q} \iff q\ln 2 = p \ln 3 \iff \ln 2^q = \ln 3^p \iff 2^q = 3^p$.

Cette dernière égalité n'est possible que si $p = q = 0$ (sinon, on aurait 2 divise 3 et 3 divise 2 en vertu du théorème de Gauss). Or $(p,q) \in \mathbb{N}^* \times \mathbb{N}^*$, on aboutit donc à une contradiction.

Conclusion : $\boxed{\dfrac{\ln 2}{\ln 3} \text{ est irrationnel.}}$

Question 4.

Préciser le(s) type(s) de raisonnement utilisé(s) dans les questions 1, 2 et 3. Donner des exemples de mise en œuvre de tels raisonnements au niveau collège.

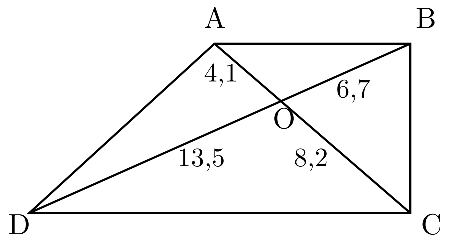
On utilise dans les questions précédentes un raisonnement par l'absurde :

Propriété 1
Pour démontrer qu'une proposition P est vraie, on peut supposer que P est fausse puis en déduire une contradiction.

Une assertion est soit vraie, soit fausse ; elle ne peut être les deux à la fois. Montrer qu'une assertion P est vraie est donc équivalent à montrer que l'assertion (non P) est fausse.

Au niveau du collège, on peut donner les exemples suivants :

- *Nombres et calculs, sixième :* les fractions $\dfrac{3}{5}$ et $\dfrac{6}{9}$ sont-elles égales ?
- *Géométrie, cinquième :* peut-on construire un triangle dont les côtés mesurent 3 cm, 4 cm et 8 cm ?
- *Géométrie, quatrième :* le quadrilatère ABCD suivant est-il un trapèze ?



- *Nombres et calculs, troisième :* $\sqrt{2}$ peut-il s'écrire sous la forme d'une fraction du type $\dfrac{a}{b}$ avec a et b deux nombres entiers premiers entre eux ?

Question 5.

On rappelle que $\mathrm{e} = \displaystyle\sum_{k=0}^{+\infty} \frac{1}{k!}$. On se propose de démontrer que le nombre e est un nombre irrationnel.

Pour cela, on fait l'hypothèse qu'il existe p et q, entiers naturels non nuls, tels que $\mathrm{e} = \dfrac{p}{q}$.

Pour tout entier naturel n, on pose : $u_n = \displaystyle\sum_{k=0}^{n} \frac{1}{k!}$ et $v_n = u_n + \dfrac{1}{n \times n!}$.

5.1. Démontrer que les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont adjacentes, puis montrer que : $u_q < \mathrm{e} < v_q$.

Définition 2 : suites adjacentes
Deux suites réelles sont dites adjacentes si l'une des suites est croissante, l'autre est décroissante et si la différence des deux converge vers 0.

i) Montrons que la suite $(u_n)_{n\in\mathbb{N}}$ est strictement croissante :

$\forall\, n \in \mathbb{N},\ u_{n+1} - u_n = \displaystyle\sum_{k=0}^{n+1} \frac{1}{k!} - \sum_{k=0}^{n} \frac{1}{k!} = \frac{1}{(n+1)!}.$

Or, $\forall\, n \in \mathbb{N}, \dfrac{1}{(n+1)!} > 0$, donc la suite $(u_n)_{n\in\mathbb{N}}$ est strictement croissante.

ii) Montrons que la suite $(v_n)_{n\in\mathbb{N}^*}$ est strictement décroissante :

$$\begin{aligned}
\forall\, n \in \mathbb{N}^*,\ v_{n+1} - v_n &= u_{n+1} - u_n + \frac{1}{(n+1)(n+1)!} - \frac{1}{n \times n!}\\
&= \frac{1}{(n+1)!} + \frac{1}{(n+1)(n+1)!} - \frac{1}{n \times n!}\\
&= \frac{n(n+1) + n - (n+1)^2}{n(n+1)(n+1)!}\\
&= \frac{\cancel{n^2} + \cancel{n} + \cancel{n} - \cancel{n^2} - \cancel{2n} - 1}{n(n+1)(n+1)!}\\
&= \frac{-1}{n(n+1)(n+1)!}
\end{aligned}$$

Or, $\forall\, n \in \mathbb{N}^*, \dfrac{-1}{n(n+1)(n+1)!} < 0$, donc la suite $(v_n)_{n\in\mathbb{N}^*}$ est strictement décroissante.

iii) Montrons que la différence des deux suites converge vers 0 :

$\forall\, n \in \mathbb{N}^*,\ v_n - u_n = \dfrac{1}{n \times n!}$ donc, $\displaystyle\lim_{n\to+\infty} \frac{1}{n \times n!} = 0$

Donc, la différence des deux suites converge vers 0.

☆ *Conclusion :* $\boxed{\textit{Les suites } (u_n)_{n\in\mathbb{N}} \textit{ et } (v_n)_{n\in\mathbb{N}^*} \textit{ sont adjacentes.}}$

iv) Montrons l'inégalité :

Théorème 3 : des suites adjacentes
Deux suites adjacentessont convergentes et ont la même limite.

D'après le théorème des suites adjacentes, les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}^*}$ convergent vers $\displaystyle\lim_{n\to+\infty} u_n = \mathrm{e}$. De plus, la suite $(u_n)_{n\in\mathbb{N}}$ est strictement croissante majorée par e, la suite $(v_n)_{n\in\mathbb{N}^*}$ est strictement décroissante minorée par e d'où $\forall\, q \in \mathbb{N}^*$, $u_q < \mathrm{e}$ et $v_q > \mathrm{e}$.

☆ *Conclusion :* $\boxed{\forall\, q \in \mathbb{N}^*,\ u_q < \ e\ < v_q.}$

5.2.

Conclure en multipliant les termes de cet encadrement par $q\,! \times q$.

$$\forall\, q \in \mathbb{N}^*,\ \sum_{k=0}^{q} \frac{1}{k!} < \ \mathrm{e}\ < \sum_{k=0}^{q} \frac{1}{k!} + \frac{1}{q!q} \iff \sum_{k=0}^{q} \frac{q!q}{k!} < \mathrm{e}q!q < \sum_{k=0}^{q} \frac{q!q}{k!} + \frac{q!q}{q!q}$$
$$\iff q \times \sum_{k=0}^{q} \frac{q!}{k!} < \frac{p}{\cancel{q}} q!\cancel{q} < q \times \sum_{k=0}^{q} \frac{q!}{k!} + 1.$$

Or, $\forall\, 0 \leq k \leq q$, $\dfrac{q!}{k!} = q \times (q-1) \times \cdots \times (q-k+1) \in \mathbb{N}^*$. Donc, $\displaystyle\sum_{k=0}^{q} \frac{q!}{k!}$ est un nombre entier naturel.

Si l'on note $N = q \times \displaystyle\sum_{k=0}^{q} \frac{q!}{k!}$, on obtient alors l'encadrement : $N < p \times q! < N+1$, ce qui est impossible puisque $p \times q!$ est un entier.
On en conclut à la fausseté de notre hypothèse.

Conclusion : $\boxed{\textit{e est un nombre irrationnel.}}$

Partie B : une preuve de l'irrationalité de π

On se propose ici de démontrer que le nombre π est un nombre irrationnel. Pour cela, on fait l'hypothèse qu'il existe a et b, entiers naturels non nuls, tels que $\pi = \dfrac{a}{b}$.
Étant donnés un entier naturel non nul n et un réel x, on pose :

$$P_n(x) = \frac{x^n(a-bx)^n}{n!} \quad \text{et} \quad P_0(x) = 1$$

Étant donné un entier naturel n, on pose :

$$I_n = \int_0^{\pi} P_n(x) \sin x \,\mathrm{d}x$$

Question 1.

1.1.

Pour un entier naturel n non nul, exprimer la dérivée de P_n en fonction de P_{n-1}.

$\forall\, n \in \mathbb{N}^*, \forall\, x \in \mathbb{R},\ P_n(x) = \dfrac{1}{n!} \times x^n \times (a-bx)^n$.
Or, les fonctions polynomiales $x \mapsto x^n$ et $x \mapsto (a-bx)^n$ sont dérivables sur $\mathbb{R}$ donc P_n est dérivable sur $\mathbb{R}$, de dérivée :

$$P_n'(x) = \frac{nx^{n-1} \times (a-bx)^n + x^n \times (-nb)(a-bx)^{n-1}}{n!}$$
$$P_n'(x) = \frac{x^{n-1}(a-bx)^{n-1} \times (a-bx-bx)}{(n-1)!}.$$

Conclusion : $\boxed{\forall\, n \in \mathbb{N}^*, \forall\, x \in \mathbb{R},\ P_n'(x) = (a-2bx)P_{n-1}(x).}$

1.2.

Calculer $\sup\limits_{x\in[0,\pi]} |P_n(x)|$ en fonction de a, b et n.

Pour $n = 0, P_0(x) = 1$ et $\sup\limits_{x\in[0,\pi]} |P_0(x)| = 1$.

Pour $n \geq 1$, sachant que $\pi = \dfrac{a}{b}$ avec a et b des nombres entiers naturels non nuls, alors
$0 \leq x \leq \pi \iff 0 \leq x \leq \dfrac{a}{b} \iff 0 \leq bx \leq a \iff a - bx \geq 0$.

On a alors, $\forall n \geq 1, \forall x \in [0, \pi]$, $x^{n-1}, (a-bx)^{n-1}$ et $(n-1)!$ positifs.
Le signe de $P_n'(x)$ dépend donc uniquement du signe de $(a - 2bx)$.

Or, $a - 2bx \geq 0 \iff x \leq \dfrac{a}{2b}$ d'où le tableau de variation de P_n :

x	0		$\frac{a}{2b}$		π
$P_n'(x)$		$+$	0	$-$	
P_n	0	$\nearrow$	$\dfrac{\left(\frac{a^2}{4b}\right)^n}{n!}$	$\searrow$	0

- $P_n(0) = 0$;
- $P_n\left(\frac{a}{2b}\right) = \dfrac{\left(\frac{a}{2b}\right)^n \left(a - \cancel{b} \times \frac{a}{2\cancel{b}}\right)^n}{n!} = \dfrac{\left(\frac{a^2}{4b}\right)^n}{n!}$;
- $P_n(\pi) = 0$.

Remarque : l'expression obtenue comme valeur maximale de P_n est encore valable pour $n = 0$.

Conclusion : $\boxed{\forall n \in \mathbb{N},\ \sup\limits_{x\in[0,\pi]} |P_n(x)| = \dfrac{\left(\frac{a^2}{4b}\right)^n}{n!}}$.

1.3.

Démontrer que : $\forall n \in \mathbb{N} \quad \forall x \in \mathbb{R} \quad P_n\left(\dfrac{a}{b} - x\right) = P_n(x)$.

$$\begin{aligned}
\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, \quad P_n\left(\frac{a}{b} - x\right) &= \frac{\left(\frac{a}{b} - x\right)^n \left[a - b\left(\frac{a}{b} - x\right)\right]^n}{n!} \\
&= \frac{\left(\frac{a}{b} - x\right)^n (\cancel{a} - \cancel{a} + bx)^n}{n!} \\
&= \frac{\left[\left(\frac{a}{b} - x\right) bx\right]^n}{n!} \\
&= \frac{[(a - bx)\, x]^n}{n!} \\
&= \frac{(a - bx)^n x^n}{n!} \\
&= P_n(x).
\end{aligned}$$

Conclusion : $\boxed{\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, P_n\left(\dfrac{a}{b} - x\right) = P_n(x).}$

1.4.

Démontrer que : $\forall n \in \mathbb{N} \quad I_n > 0$.

Les fonctions $x \mapsto P_n(x)$ et $x \mapsto \sin(x)$ sont continues et positives sur l'intervalle $[0, \pi]$.
La fonction produit $x \mapsto P_n(x)\sin(x)$ est donc continue, positive, non identiquement nulle sur $[0, \pi]$, et intégrable sur $[0, \pi]$. De plus, sa somme est positive et non nulle.

Conclusion : $\boxed{\forall n \in \mathbb{N},\, I_n > 0.}$

1.5.

Après avoir justifié que la suite de terme général $\dfrac{\pi}{n!}\left(\dfrac{a^2}{4b}\right)^n$ tend vers 0, démontrer la convergence de la suite $(I_n)_{n\in\mathbb{N}}$ et déterminer sa limite.

i) Montrons que la suite tend vers 0 :

Théorème 4 : Exponentielle
Le développement en série entière de la fonction exponentielle est : $\forall x \in \mathbb{R}, \exp(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{n!}$.

On a donc $\exp\left(\dfrac{a^2}{4b}\right) = \displaystyle\sum_{n=0}^{+\infty} \frac{\left(\frac{a^2}{4b}\right)^n}{n!}$ et $\pi\exp\left(\dfrac{a^2}{4b}\right) = \displaystyle\sum_{n=0}^{+\infty} \frac{\pi}{n!}\left(\frac{a^2}{4b}\right)^n$.

La série de terme général $\dfrac{\pi}{n!}\left(\dfrac{a^2}{4b}\right)^n$ converge vers $\pi\exp\left(\dfrac{a^2}{4b}\right)$, son terme général tend donc vers 0.

Conclusion : $\boxed{\textit{La suite de terme général } \dfrac{\pi}{n!}\left(\dfrac{a^2}{4b}\right)^n \textit{ tend vers 0.}}$

ii) Montrons la convergence de la suite $(I_n)_{n\in\mathbb{N}}$ et sa convergence vers 0 :

$\forall x \in [0, \pi], 0 \leq \sin(x) \leq 1$ et $0 \leq P_n(x) \leq \dfrac{1}{n!}\left(\dfrac{a^2}{4b}\right)^n$ d'après la question 1.2.

D'ou : $0 \leq P_n(x)\sin(x) \leq \dfrac{1}{n!}\left(\dfrac{a^2}{4b}\right)^n$, et par intégration d'une fonction continue positive sur $[0, \pi]$, on a

$0 \leq I_n \leq \pi \times \dfrac{1}{n!}\left(\dfrac{a^2}{4b}\right)^n$ avec $\displaystyle\lim_{n\to+\infty} \frac{\pi}{n!}\left(\frac{a^2}{4b}\right)^n = 0$.

Théorème 5 : d'encadrement
Soient u, v et w trois suites réelles telles que :

- $\exists N \in \mathbb{N}\, \forall n \in N\, (n \geq N \Longrightarrow u_n \leq v_n \leq w_n)$;
- (u_n) et (w_n) convergent vers la même limite l ;

alors la suite (v_n) converge et sa limite est l.

Conclusion : $\boxed{\textit{La suite } (I_n)_{n\in\mathbb{N}} \textit{ converge et sa limite vaut 0.}}$

Question 2.

Pour tout entier naturel k, la dérivée d'ordre k de P_n est notée $P_n^{(k)}$. Par définition, $P_n^{(0)} = P_n$.
En distinguant les trois cas suivants, démontrer que $P_n^{(k)}(0)$ et $P_n^{(k)}\left(\dfrac{a}{b}\right)$ sont des entiers relatifs :
2.1. $0 \leq k \leq n-1$
2.2. $n \leq k \leq 2n$
2.3. $k \geq 2n+1$
Pour le cas 2.2, on pourra utiliser la relation entre $P_n^{(k)}(0)$ et le coefficient de x^k dans $P_n(x)$.

i) Relation entre $P_n^{(k)}(0)$ et $P_n^{(k)}\left(\dfrac{a}{b}\right)$:

P_n est une fonction polynomiale indéfiniment dérivable, dont la dérivée est nulle à partir d'un certain rang.
Or, d'après la question 1.3, $\forall\, n \in \mathbb{N}\,,\ \forall\, x \in \mathbb{R}\,,\ P_n\left(\dfrac{a}{b} - x\right) = P_n(x)$.
En dérivant k fois, on obtient $P_n^{(k)}\left(\dfrac{a}{b} - x\right) = (-1)^k P_n^{(k)}(x)$.
Soit, pour $x = 0$: $P_n^{(k)}\left(\dfrac{a}{b}\right) = (-1)^k P_n^{(k)}(0)$.

Il suffit donc de montrer que $P_n^{(k)}(0)$ est un entier relatif pour en déduire que $P_n^{(k)}\left(\dfrac{a}{b}\right)$ est également un entier relatif.

ii) Question 2.1. : étude du cas $0 \leq k \leq n-1$:

$P_n(x) = x^n \times \dfrac{(a-bx)^n}{n!}$ donc $(X-0)^n$ divise P_n et 0 et une racine d'ordre n de P_n.
Ainsi, 0 est une racine de la dérivée k-ème de P_n pour $0 \leq k \leq n-1$.
D'où $P_n^{(k)}(0) = P_n^{(k)}\left(\dfrac{a}{b}\right) = 0$ et ces deux nombres sont des entiers relatifs.

iii) Question 2.2. : étude du cas $n \leq k \leq 2n$:

Théorème 6 : Formule de Taylor pour les polynômes
Soit P un polynôme de $\mathbb{C}[X]$ de degré inférieur ou égal à n, et $x_0 \in \mathbb{C}$, alors :

$$P = \sum_{k=0}^{n} \frac{P^{(k)}(x_0)}{k!}(X - x_0)^k$$

$$\begin{aligned}
P_n(x) &= \frac{x^n}{n!} \times (a-bx)^n = \frac{x^n}{n!} \times \sum_{k=0}^{n} \binom{n}{k} (-bx)^k\, a^{n-k} \\
&= \sum_{k=0}^{n} \frac{1}{n!} \times \binom{n}{k} (-b)^k\, a^{n-k}\, x^{n+k} \\
&= \sum_{k=n}^{2n} \frac{1}{n!} \times \binom{n}{k-n} (-b)^{k-n}\, a^{2n-k}\, x^k.
\end{aligned}$$

Par identification avec la formule de Taylor, on obtient :
$\forall n \leq k \leq 2n,\ \dfrac{P^{(k)}(0)}{k!} = \dfrac{1}{n!} \times \dbinom{n}{k-n} (-b)^{k-n}\, a^{2n-k}$

Soit : $P^{(k)}(0) = \dfrac{k!}{n!} \times \dbinom{n}{k-n} (-b)^{k-n}\, a^{2n-k}$. Or,

- $\dfrac{k!}{n!}$ est un entier puisque $k \geq n$;
- le coefficient binomial $\dbinom{n}{k-n}$ est un nombre entier ;
- $k-n$ et $2n-k$ sont positifs et a et b sont des entiers naturels non nuls, donc $(-b)^{k-n}\, a^{2n-k} \in \mathbb{Z}$.

D'où : $\forall n \leq k \leq 2n$, $P^{(k)}(0)$, et donc $P_n^{(k)}\left(\dfrac{a}{b}\right)$ sont des entiers relatifs.

Remarque : on retrouve également avec la formule de Taylor le résultat de la question 2.1.

iv) Question 2.3. : étude du cas $k \geq 2n+1$:

P_n est un polynôme de degré $2n$, sa dérivée k-ème est donc nulle pour tout $k \geq 2n+1$.
D'où $P_n^{(k)}(0) = P_n^{(k)}\left(\dfrac{a}{b}\right) = 0$ sont des entiers relatifs.

Conclusion : $\boxed{\forall k \in \mathbb{N},\ P_n^{(k)}(0) \text{ et } P_n^{(k)}\left(\dfrac{a}{b}\right) \text{ sont des entiers relatifs.}}$

Question 3.

3.1.

Démontrer que pour tout entier naturel n, I_n est un entier relatif. On pourra procéder par intégrations par parties successives.

i) Formule de récurrence.

On procède à une première intégration par parties avec : $\begin{cases} u(x) = P_n(x) & v'(x) = \sin(x) \\ u'(x) = P_n'(x) & v(x) = -\cos(x) \end{cases}$, $(u, v) \in \mathcal{C}^1$.

$$\begin{aligned} I_n &= \int_0^\pi P_n(x)\sin(x)\mathrm{d}x = [-P_n(x)\cos(x)]_0^\pi + \int_0^\pi P_n^{'}(x)\cos(x)\mathrm{d}x \\ &= -P_n(\pi)\cos(\pi) + P_n(0)\cos(0) + \int_0^\pi P_n^{'}(x)\cos(x)\mathrm{d}x \\ &= P_n(\pi) + P_n(0) + \int_0^\pi P_n^{'}(x)\cos(x)\mathrm{d}x \end{aligned}$$

Puis à une seconde intégration par parties avec : $\begin{cases} u(x) = P_n^{'}(x) & v'(x) = \cos(x) \\ u'(x) = P_n^{''}(x) & v(x) = \sin(x) \end{cases}$, $(u, v) \in \mathcal{C}^1 \times \mathcal{C}^1$.

$$\begin{aligned} I_n &= P_n(\pi) + P_n(0) + \left[P_n^{'}(x)\sin(x)\right]_0^\pi - \int_0^\pi P_n^{''}(x)\sin(x)\mathrm{d}x \\ &= P_n(\pi) + P_n(0) + \cancel{P_n^{'}(\pi)\sin(\pi)} - \cancel{P_n^{'}(0)\sin(0)} - \int_0^\pi P_n^{''}(x)\sin(x)\mathrm{d}x \end{aligned}$$

En réitérant ces deux intégrations, montrons par récurrence que l'on a :

$$\mathcal{H}_N \ : \ I_n = \sum_{k=0}^{N}\left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1}\int_0^\pi P_n^{(2N+2)}(x)\sin(x)\mathrm{d}x$$

- **Initialisation :** Pour $N = 0$, on obtient $I_n = (-1)^0 P_n^{(0)}(\pi) + P_n^{(0)}(0) + (-1)^1\displaystyle\int_0^\pi P_n^{(2)}(x)\sin(x)\mathrm{d}x$

  Soit $I_n = P_n(\pi) + P_n(0) - \displaystyle\int_0^\pi P_n^{''}(x)\sin(x)\mathrm{d}x$, ce qui est vrai.

- **Hérédité :** supposons l'hypothèse vraie au rang N, on effectue deux intégration par parties successives :

$$\begin{aligned} I_n &= \sum_{k=0}^{N}\left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1}\left(\left[-P_n^{(2N+2)}(x)\cos(x)\right]_0^\pi + \int_0^\pi P_n^{(2N+3)}(x)\cos(x)\mathrm{d}x\right) \\ &= \sum_{k=0}^{N}\left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1}\left(P_n^{(2N+2)}(\pi) + P_n^{(2N+2)}(0) + \int_0^\pi P_n^{(2N+3)}(x)\cos(x)\mathrm{d}x\right) \\ &= \sum_{k=0}^{N+1}\left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1}\left(\left[P_n^{(2N+3)}(x)\sin(x)\right]_0^\pi - \int_0^\pi P_n^{(2N+4)}(x)\sin(x)\mathrm{d}x\right) \end{aligned}$$

$$= \sum_{k=0}^{N+1} \left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1} \left(\cancel{P_n^{(2N+3)}(\pi)\sin(\pi)} - \cancel{P_n^{(2N+3)}(0)\sin(0)} - \int_0^\pi P_n^{(2N+4)}(x)\sin(x)\mathrm{d}x\right)$$

$$= \sum_{k=0}^{N+1} \left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+2} \int_0^\pi P_n^{(2N+4)}(x)\sin(x)\mathrm{d}x$$

L'hypothèse reste donc encore vraie au rang $N+1$.

On en conclut donc que $\mathcal{H}_N$ est vraie pour tout $N \geq 0$.

ii) Appartenance de I_n à $\mathbb{Z}$.

On vient de montrer que I_n peut s'écrire, $\forall n \in \mathbb{N}, \forall N \in \mathbb{N}$;
$$I_n = \sum_{k=0}^{N} \left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{N+1} \int_0^\pi P_n^{(2N+2)}(x)\sin(x)\mathrm{d}x.$$
Cette égalité appliquée au cas particulier $n = N$ donne alors
$$I_n = \sum_{k=0}^{n} \left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right] + (-1)^{n+1} \int_0^\pi P_n^{(2n+2)}(x)\sin(x)\mathrm{d}x$$

Or, P_n est un polynôme de degré $2n$ donc, $P_n^{(2n+2)} = 0$. On obtient alors
$$I_n = \sum_{k=0}^{n} \left[(-1)^k P_n^{(2k)}(\pi) + P_n^{(2k)}(0)\right].$$

D'après la question 2, $\forall\, k \in \mathbb{N}$, $P_n^{(k)}(0)$ et $P_n^{(k)}(\pi)$ sont des entiers relatifs, il en est de même pour $P_n^{(2k)}(0)$ et $P_n^{(2k)}(\pi)$ et nous avons donc une somme d'entiers relatifs.

☆ *Conclusion :* $\boxed{\forall n \in \mathbb{N},\ I_n \textit{ est un entier relatif.}}$

3.2.

Conclure quant à l'hypothèse $\pi = \dfrac{a}{b}$.

D'après les questions 1.4 et 3.1, I_n est un entier strictement positif, donc $\forall n \in \mathbb{N}$, $I_n \geq 1$.

D'après la question 1.5, la suite (I_n) tend vers 0, donc :
$\forall \epsilon > 0$, $\exists N \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq N \Longrightarrow |I_n| < \epsilon)$.
Pour $\epsilon = \frac{1}{2}$, $\exists N_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq N_0 \Longrightarrow |I_n| < \frac{1}{2})$, ce qui contredit le fait que $I_n \geq 1$.

☆ *Conclusion :* $\boxed{\textit{Il n'existe pas d'entiers } a \textit{ et } b \textit{ tels que } \pi = \dfrac{a}{b}, \textit{ et donc } \pi \textit{ n'est pas rationnel.}}$

Problème n°3

Statistiques et probabilités

Partie A : deux indicateurs de dispersion

En 1801, un astronome italien, Piazzi découvre une nouvelle planète Cérès, qu'il perd bientôt de vue. Le problème posé alors aux scientifiques est le suivant : comment, à partir d'une série de résultats d'observations effectuées par différents astronomes, choisir une valeur qui se rapproche le plus possible de la « vraie position » et prédire ainsi le futur passage de Cérès. Deux options s'affrontent : celle de Laplace, qui propose de minimiser les valeurs absolues des écarts et celle de Gauss et Legendre, qui proposent de minimiser les carrés des écarts.

Dans cette partie, n désigne un entier naturel non nul et $(x_1, \dots, x_n)$, un n-uplet de réels.
On définit sur $\mathbb{R}$ les deux fonctions G et L par :

$$G(x) = \sum_{i=1}^{n} (x - x_i)^2$$

$$L(x) = \sum_{i=1}^{n} |x - x_i|$$

Bref aperçu historique...

La méthode élaborée par *Gauss* et *Legendre* est couramment appelée la méthode des moindres carrés.
Legendre publie sa méthode en 1805 dans ses « Nouvelles méthodes pour la détermination des orbites des comètes », *Gauss* exposera la sienne de manière plus élaborée et justifiée par la théorie des probabilités en 1809 dans sa « Théorie du mouvement des corps célestes ».
Une polémique s'installe concernant la paternité de la méthode. Legendre a la primauté du texte écrit, tandis que *Gauss* peut faire état de témoignages d'astronomes montrant qu'il avait déjà établit le principe dès 1795.

Question 1.

Minimisation de G.

1.1.

En écrivant $G(x)$ sous la forme d'un trinôme du second degré, démontrer que la fonction G admet un minimum sur $\mathbb{R}$ et indiquer pour quelle valeur de x il est atteint.

$$\begin{aligned} G(x) &= \sum_{i=1}^{n} (x^2 - 2 \times x \times x_i + x_i^2) \\ &= nx^2 - 2x \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} x_i^2. \end{aligned}$$

Propriété 1

La courbe représentative de la fonction trinôme $f : x \mapsto ax^2 + bx + c$ où $a \neq 0$ est une parabole dont le sommet est atteint au point d'abscisse $x_0 = -\dfrac{b}{2a}$.

- Si $a > 0$, la parabole est tournée vers le haut et f atteint un minimum en x_0 ;
- si $a < 0$, la parabole est tournée vers le bas et f atteint un maximum en x_0.

G est une fonction polynomiale du second degré dont la représentation graphique est une parabole tournée vers le haut puisque n est positif.

Son minimum est atteint lorsque $x = -\dfrac{-2\sum_{i=1}^{n} x_i}{2 \times n} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

☆ *Conclusion :* $\boxed{\text{La fonction } G \text{ admet un minimum atteint pour } x_0 = \frac{1}{n}\sum_{i=1}^{n} x_i.}$

1.2.

Que représente d'un point de vue statistique la valeur de x trouvée à la question 1.1. ?

☆ $\boxed{x_0 \text{ correspond à la moyenne arithmétique des valeurs } x_i \text{ du } n\text{-uplet.}}$

1.3.

On se place dans $\mathbb{R}^n$, espace affine euclidien muni du produit scalaire usuel. Comment peut-on interpréter la fonction G en terme de distance euclidienne ? Comment peut-on alors retrouver le résultat de la question 1.1. ?

On se place dans un repère orthonormé de $\mathbb{R}^n$, muni du produit scalaire usuel.
On note $M(x, x, \dots, x)$ et $A(x_1, x_2, \dots, x_n)$ deux points de ce repère.

$G(x) = \sum_{i=1}^{n}(x - x_i)^2$ correspond donc à la distance AM^2.

On considère maintenant la droite d passant par l'origine de vecteur directeur $\overrightarrow{u}(1, 1, \dots, 1)$.
On souhaite minimiser la fonction G, ceci revient à trouver la distance minimale entre le point A et la droite d.

Pour tout point M de d, il existe un unique réel t tel que $\overrightarrow{OM} = t\overrightarrow{u}$.
Soit M_0 le point de d pour lequel la distance AM^2 est minimale et $\phi : \mathbb{R} \mapsto \mathbb{R}$ la fonction définie par $\phi(t) = \overrightarrow{AM}^2$.

$$\begin{aligned}
\phi(t) &= (\overrightarrow{AO} + \overrightarrow{OM})^2 \\
&= (t\overrightarrow{u} - \overrightarrow{OA})^2 \\
&= ||\overrightarrow{u}||^2 t^2 - 2\overrightarrow{u}.\overrightarrow{OA}\, t + ||\overrightarrow{OA}||^2.
\end{aligned}$$

ϕ est une fonction polynomiale du second degré dont la représentation graphique est une parabole tournée vers le haut puisque $||\overrightarrow{u}||^2$ est positif.

Son minimum est atteint au point M_0 tel que $t_0 = -\dfrac{-2\overrightarrow{u}.\overrightarrow{OA}}{2||\overrightarrow{u}||^2} = \dfrac{\overrightarrow{u}.\overrightarrow{OA}}{||\overrightarrow{u}||^2}$.

Dans notre cas, on obtient $t_0 = \dfrac{1 \times x_1 + 1 \times x_2 + \cdots + 1 \times x_n}{1 + 1 + \cdots + 1} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

On a alors $\overrightarrow{OM_0} = \dfrac{1}{n}\sum_{i=1}^{n} x_i \overrightarrow{u} \iff (x_0\,, x_0\,, \dots, x_0) = \left(\dfrac{1}{n}\sum_{i=1}^{n} x_i\,,\ \dfrac{1}{n}\sum_{i=1}^{n} x_i\,,\ \dots\,,\ \dfrac{1}{n}\sum_{i=1}^{n} x_i\right)$.

☆ *Conclusion :* $\boxed{G \text{ représente la distance au carré entre les points } M(x, x, \dots, x) \text{ et } A(x_1, x_2, \dots, x_n), \text{ celle-ci est rendue minimale pour } x_0 = \frac{1}{n}\sum_{i=1}^{n} x_i.}$

Question 2.

Minimisation de L.

On supposera dans cette question que la série est ordonnée, c'est-à-dire que : $x_1 \leq x_2 \leq \ldots \leq x_n$.

2.1

Démontrer que la fonction L admet un minimum m sur $\mathbb{R}$ et indiquer pour quelle(s) valeur(s) de x il est atteint.

On distinguera les cas n pair et n impair.

La fonction L est une fonction continue, affine par morceaux.
La série étant ordonnée, on peut étudier trois cas :

- $\forall x \in]-\infty; x_1[\,, \forall i \in [1;n]\,, |x - x_i| = x_i - x$ et $\displaystyle L(x) = \sum_{i=1}^{n}(x_i - x) = -nx + \sum_{i=1}^{n} x_i.$

  Sur cet intervalle, la fonction est affine, strictement décroissante.

- $\forall j \in [1; n-1]\,, \forall x \in [x_j; x_{j+1}]$
  $$\begin{aligned} L(x) &= \sum_{i=1}^{j} |x - x_i| + \sum_{i=j+1}^{n} |x - x_i| \\ &= \sum_{i=1}^{j}(x - x_i) - \sum_{i=j+1}^{n}(x - x_i) \\ &= jx - \sum_{i=1}^{j} x_i - (n-j)x + \sum_{i=j+1}^{n} x_i. \\ &= (2j - n)x - \sum_{i=1}^{j} x_i + \sum_{i=j+1}^{n} x_i. \end{aligned}$$
  Sur chacun de ces intervalles, la fonction est strictement décroissante si $2j - n < 0$, donc si $j < \dfrac{n}{2}$; constante si $j = \dfrac{n}{2}$ et strictement croissante si $j > \dfrac{n}{2}$.

- $\forall x \in ,]x_n; +\infty[\,, \forall i \in [1;n]\,, |x - x_i| = x - x_i$ et $\displaystyle L(x) = \sum_{i=1}^{n}(x - x_i) = nx - \sum_{i=1}^{n} x_i.$

  Sur cet intervalle, la fonction est strictement croissante.

On obtient donc deux cas en fonction de la parité de n :

- Si $n = 2k, k \in \mathbb{N}^*$ est pair, la fonction est strictement décroissante sur $]-\infty; x_k]$, constante sur $[x_k; x_{k+1}]$ et strictement croissante sur $[x_{k+1}; +\infty[$.
- Si $n = 2k+1, k \in \mathbb{N}^*$ est impair, la fonction est strictement décroissante sur $]-\infty; x_{k+1}]$ et strictement croissante sur $[x_{k+1}; +\infty[$.

☆ *Conclusion :* *La fonction L admet un minimum m sur $\mathbb{R}$, atteint sur l'intervalle $[x_{\frac{n}{2}}; x_{\frac{n}{2}+1}]$ si n est pair, et en $x_{\frac{n+1}{2}}$ si n est impair.*

2.2.

Que représentent d'un point de vue statistique les valeurs de x trouvées à la question 2.1. ?

L'intervalle $[x_{\frac{n}{2}}; x_{\frac{n}{2}+1}]$ correspond à l'intervalle médian des intervalles $[x_1; x_2]; \dots; [x_{n-1}; x_n]$; la valeur $x_{\frac{n+1}{2}}$ est la médiane des valeurs $x_1; \dots; x_n$.

Question 3.

Dans les deux premières questions, on a déterminé les variations d'une fonction d'une variable réelle.

3.1

Donner la définition d'une fonction croissante sur $\mathbb{R}$ au niveau lycée.

Soit f une fonction définie sur un intervalle I inclus dans $\mathbb{R}$. On dit que f est croissante sur I si pour tous réel a et b de I, si $a < b$, alors $f(a) \leq f(b)$.

3.2.

Donner plusieurs méthodes permettant de déterminer les variations d'une fonction au niveau lycée.

Soit f une fonction définie sur $\mathcal{D} \subset \mathbb{R}$. On peut, selon le niveau, utiliser les méthodes suivantes :

1. Pour tous réels $a < b$ de $I \subset \mathcal{D}$, on compare $f(a)$ et $f(b)$.
   Si $f(a) < f(b)$, f est strictement croissante sur I ; si $f(a) = f(b)$, f est constante sur I et si $f(a) > f(b)$, f est strictement décroissante sur I. Si aucun des trois cas n'est réalisé, on ne peut pas conclure.
2. Pour tous réels $a < b$ de $I \subset \mathcal{D}$, on étudie le signe de $f(a) - f(b)$.
   Si $f(a) - f(b) < 0$, f est strictement croissante sur I ; si $f(a) - f(b) = 0$, f est constante sur I et si $f(a) - f(b) > 0$, f est strictement décroissante sur I. Si aucun des trois cas n'est réalisé, on ne peut pas conclure.
3. On décompose la fonction en fonctions de référence dont on connait le sens de variation.
   La composée de deux fonctions toutes les deux croissantes ou toutes les deux décroissantes est croissante ; la composition de deux fonctions, l'une croissante et l'autre décroissante, est décroissante.
4. Si f est dérivable sur $I \subset \mathcal{D}$, on étudie le signe de la dérivée de f sur I.
   Si $f'(x) \geq 0$, la fonction est croissante sur I ; si $f'(x) = 0$, la fonction est constante sur I et si $f'(x) \leq 0$, la fonction est décroissante sur I.

Par exemple, on souhaite étudier ls variations de la fonction f définie sur $[1\,;4]$ par $f(x) = \dfrac{1}{x^2}$.

1. Soit $1 \leq a < b \leq 4$, alors $0 < 1 \leq a^2 < b^2 \leq 16$, puis $\dfrac{1}{16} \leq \dfrac{1}{b^2} < \dfrac{1}{a^2} \leq 1$, donc $f(b) < f(a)$.
   La fonction f est donc décroissante sur $[1\,;4]$.
2. Soit $1 \leq a < b \leq 4$, alors $f(a) - f(b) = \dfrac{1}{a^2} - \dfrac{1}{b^2} = \dfrac{b^2 - a^2}{a^2 b^2} = \dfrac{(b-a)(b+a)}{(ab)^2} > 0$ sur $[1\,;4]$.
   La fonction f est donc décroissante sur $[1\,;4]$.
3. La fonction f se décompose de la manière suivante : $x \xrightarrow{g} x^2 \xrightarrow{h} \dfrac{1}{x^2}$.
   La fonction $g : x \mapsto x^2$ est croissante sur $\mathbb{R}^+$, donc sur $[1\,;4]$ et la fonction $h : x \mapsto \dfrac{1}{x}$ est décroissante sur $\mathbb{R}^+_*$, donc sur $[1\,;4]$. La fonction f est donc décroissante sur $[1\,;4]$, comme composée d'une fonction croissante et d'une fonction décroissante sur $[1\,;4]$.
4. f est continue, dérivable sur $[1\,;4]$, de dérivée $f'(x) = -\dfrac{2}{x^3}$. f' est négative sur $[1\,;4]$ donc, f est décroissante sur $[1\,;4]$.

Le 7 décembre 1801, Cérès sera observée à l'endroit prévu par les calculs de Gauss. Il prolongera ce travail en établissant, grâce à la théorie des probabilités, que la répartition des erreurs suit une loi normale.

Partie B : théorie de l'information, le cas discret

La théorie de l'information est un modèle mathématique créé par Claude Shannon en 1948, qui vise à quantifier mathématiquement la notion d'incertitude. Elle a depuis connu des développements aussi bien en statistique qu'en physique théorique ou en théorie du codage.

On se place dans cette partie dans un espace probabilisé $(\Omega, \mathcal{A}, P)$.
Étant donné un entier naturel non nul n, on considère un système complet d'événements $A = \{A_1, \ldots, A_,\}$ de probabilités respectives $(p_1, \ldots, p_n)$ toutes non nulles.
On définit l'entropie de ce système par le nombre :

$$H(A) = -\sum_{k=1}^{n} p_k \ln p_k$$

Ce nombre quantifie l'incertitude, tandis que son opposé quantifie la quantité d'information. L'entropie doit être maximale lorsqu'aucune hypothèse ne peut être privilégiée.

Question 1.

Deux exemples.
On se place ici dans le cas $n = 4$. Quatre chevaux sont au départ d'une course, et on note A_i l'événement : *Le cheval numéro i remporte la course.* Calculer dans chacun des cas suivants l'entropie du système.

1.1.
$p_1 = p_2 = p_3 = p_4$.

Nous sommes ici dans un cas d'équiprobabilité puisque $p_1 = p_2 = p_3 = p_4$. Donc, $\forall i \in [1; 4], p_i = \dfrac{1}{4}$.
On a alors :

$H(A) = -4 \times \dfrac{1}{4} \ln \dfrac{1}{4} = \ln 4 = 2 \ln 2.$

Conclusion : $\boxed{\text{Dans le cas où } p_1 = p_2 = p_3 = p_4, \text{ l'entropie vaut } H(A) = 2\ln 2.}$

1.2.
$p_1 = \dfrac{1}{8}, p_2 = \dfrac{1}{8}, p_3 = \dfrac{1}{4}, p_4 = \dfrac{1}{2}$.

$$\begin{aligned} H(A) &= -\frac{1}{8}\ln\frac{1}{8} - \frac{1}{8}\ln\frac{1}{8} - \frac{1}{4}\ln\frac{1}{4} - \frac{1}{2}\ln\frac{1}{2} \\ &= \frac{1}{4}\ln 8 + \frac{1}{4}\ln 4 + \frac{1}{2}\ln 2 \\ &= \frac{3}{4}\ln 2 + \frac{2}{4}\ln 2 + \frac{2}{4}\ln 2 = \frac{7}{4}\ln 2. \end{aligned}$$

Conclusion : $\boxed{\text{Dans le cas où } p_1 = \frac{1}{8}, p_2 = \frac{1}{8}, p_3 = \frac{1}{4}, p_4 = \frac{1}{2}, \text{ l'entropie vaut } H(A) = \frac{7}{4}\ln 2.}$

On va à présent établir la propriété générale suivante : l'entropie est maximale lorsqu'aucune hypothèse ne peut être privilégiée, c'est-à-dire lorsqu'il y a équiprobabilité.

Question 2.

Cas $n = 2$.
On considère un système complet $A = \{A_1, A_2\}$. On pose $p_1 = p$ et $p_2 = 1 - p$.
Démontrer que l'entropie est maximale lorsque les deux événements A_1 et A_2 sont équiprobables.

Dans ce cas, $H(A) = -p \ln p - (1-p)\ln(1-p)$.

Posons, pour tout $p \in]0;1[$, $f(p) = -p\ln p - (1-p)\ln(1-p)$.
Cette fonction est continue, dérivable sur $]0;1[$ de dérivée :

$$\begin{aligned} f'(p) &= -1 \times \ln p - p \times \frac{1}{p} + 1 \times \ln(1-p) - (1-p) \times \frac{-1}{1-p} \\ &= -\ln p - \cancel{1} + \ln(1-p) + \cancel{1} \\ &= \ln\left(\frac{1-p}{p}\right) = \ln\left(\frac{1}{p} - 1\right). \end{aligned}$$

Or, $\ln\left(\frac{1}{p}-1\right) \leq 0 \iff 0 < \frac{1}{p} - 1 \leq 1$

$$\begin{aligned} &\iff 1 < \frac{1}{p} \leq 2 \\ &\iff \frac{1}{2} \leq p < 1 \end{aligned}$$

D'où le tableau de variations :

p	0		$\frac{1}{2}$		1
$f'(p)$		$+$	0	$-$	
f		$\nearrow$	$\ln 2$	$\searrow$	

Le maximum est atteint pour $p = \frac{1}{2}$, il vaut : $f\left(\frac{1}{2}\right) = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2} = \ln 2$.
De plus, si $p = \frac{1}{2}$, alors $1 - p = \frac{1}{2}$ et $p_1 = p_2$.

Conclusion : *L'entropie est maximale lorsque les deux événement A_1 et A_2 sont équiprobables.*

Question 3.

Cas général.

3.1

Un résultat préliminaire : l'inégalité de Jensen.
Soit f une fonction à valeurs réelles définie sur un intervalle I. On dit que f est convexe sur I si :

$$\forall (x,y) \in I^2, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

On considère une fonction f convexe sur $I, (x_1, \dots, x_n) \in I^n, (\lambda_1, \dots, \lambda_n) \in \mathbb{R}_+^n$, avec $\sum_{k=1}^{n} \lambda_k = 1$.

Démontrer que :

$$f\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} \lambda_k f(x_k)$$

On pourra procéder par récurrence sur n, en remarquant que si $\lambda_n \neq 1$:

$$\sum_{k=1}^{n} \lambda_k x_k = \lambda_n x_n + (1-\lambda_n)\left(\sum_{k=1}^{n-1} \frac{\lambda_k}{1-\lambda_n} x_k\right)$$

Bref aperçu historique...

L'inégalité de *Jensen* est due au mathématicien et ingénieur danois *Ludwig William Valdemar Jensen*, (8 mai 1859 à Nakskov - 5 mars 1925 à Copenhague).
Il en donna la preuve en 1906. On peut l'écrire de deux manières : discrète comme dans le présent sujet, ou intégrale sous la forme suivante :
Soient g une fonction continue de $[a;b]$ dans $\mathbb{R}$ avec $-\infty \leq a < b \leq +\infty$ et φ une fonction continue convexe de $\mathbb{R}$ dans $\mathbb{R}$. Alors,

$$\varphi\left(\frac{1}{b-a}\int_a^b g(x)\,\mathrm{d}x\right) \leq \frac{1}{b-a}\int_a^b \varphi(g(x))\,\mathrm{d}x.$$

Soit $\mathcal{H}_n$ l'hypothèse : f convexe sur I et $\forall(x_1,\ldots,x_n)\in I^n, \forall(\lambda_1,\ldots,\lambda_n)\in\mathbb{R}_+^n$ tels que $\sum_{k=1}^{n}\lambda_k = 1$;

$$f\left(\sum_{k=1}^{n}\lambda_k x_k\right) \leq \sum_{k=1}^{n}\lambda_k f(x_k)$$

• Initialisation :
Si $n=2$, $\lambda_2 = 1-\lambda_1$, on obtient la définition de la convexité. Donc, $\mathcal{H}_2$ est vraie.

• Hérédité :
Supposons $\mathcal{H}_n$ vraie. On a également les hypothèses suivantes : f est une fonction convexe sur I, $(x_1,\ldots,x_{n+1})\in I^{n+1}$ et $(\lambda_1,\ldots,\lambda_{n+1})\in\mathbb{R}_+^{n+1}$ tels que $\sum_{k=1}^{n+1}\lambda_k = 1$.

Il existe i et j distincts de $[1;n+1]$ tels que la somme $\lambda_i+\lambda_j$ n'est pas nulle (sinon, la somme des λ_k serait nulle).

On a alors : $f\left(\sum_{k=1}^{n+1}\lambda_k x_k\right) = f\left((\lambda_i+\lambda_j)\times\frac{\lambda_i x_i+\lambda_j x_j}{\lambda_i+\lambda_j} + \sum_{\substack{k=1\\k\neq i;k\neq j}}^{n+1}\lambda_k x_k\right)$

Or, $\frac{\lambda_i x_i+\lambda_j x_j}{\lambda_i+\lambda_j}\in I$ en tant que barycentre de deux élément de I affectés de coefficients positifs et $\lambda_i+\lambda_j\in\mathbb{R}_+$. On peut donc appliquer l'hypothèse de récurrence $\mathcal{H}_n$:

$f\left(\sum_{k=1}^{n+1}\lambda_k x_k\right) \leq (\lambda_i+\lambda_j)f\left(\frac{\lambda_i x_i+\lambda_j x_j}{\lambda_i+\lambda_j}\right) + \sum_{\substack{k=1\\k\neq i;k\neq j}}^{n+1}\lambda_k f(x_k)$.

De plus, par convexité de la fonction f :

$f\left(\frac{\lambda_i x_i+\lambda_j x_j}{\lambda_i+\lambda_j}\right) = f\left(\frac{\lambda_i}{\lambda_i+\lambda_j}x_i + \frac{\lambda_j}{\lambda_i+\lambda_j}x_j\right) \leq \frac{\lambda_i}{\lambda_i+\lambda_j}f(x_i) + \frac{\lambda_j}{\lambda_i+\lambda_j}f(x_j)$

D'où : $f\left(\sum_{k=1}^{n+1}\lambda_k x_k\right) \leq \cancel{(\lambda_i+\lambda_j)}\frac{\lambda_i}{\cancel{\lambda_i+\lambda_j}}f(x_i) + \cancel{(\lambda_i+\lambda_j)}\frac{\lambda_j}{\cancel{\lambda_i+\lambda_j}}f(x_j) + \sum_{\substack{k=1\\k\neq i;k\neq j}}^{n+1}\lambda_k f(x_k)$

$\leq \sum_{k=1}^{n+1}\lambda_k f(x_k)$. Donc, $\mathcal{H}_{n+1}$ est encore vraie.

Conclusion : $\boxed{\textit{Avec les hypothèses de l'énonce, on a bien } f\left(\sum_{k=1}^{n}\lambda_k x_k\right) \leq \sum_{k=1}^{n}\lambda_k f(x_k).}$

3.2.

On admet le théorème suivant :
si f est deux fois dérivable sur I, f est convexe sur I si et seulement si f'' est positive sur I.
Démontrer que la fonction $x \mapsto x \ln x$ est convexe sur $]0,1[$.

La fonction f définie sur $]0;1[$ par $f(x) = x \ln x$ est continue et deux fois dérivable sur $]0;1[$ comme produit de deux fonction continues et deux fois dérivables sur cet intervalle.

$\forall x \in]0;1[\,, f'(x) = \ln x + x \times \dfrac{1}{x} = \ln x + 1.$

$\forall x \in]0;1[\,, f''(x) = \dfrac{1}{x}.$

Or, $\forall x \in]0;1[\,, \dfrac{1}{x} > 1$ donc f" est positive.

Conclusion : $\boxed{\textit{La fonction } x \mapsto x \ln x \textit{ est convexe sur }]0;1[.}$

3.3.

Démontrer que $H(A) \leq \ln n$. Conclure.

$H(A) = -\sum_{k=1}^{n} p_k \ln p_k$ avec $\sum_{k=1}^{n} p_k = 1$ puisque $A = \{A_1, \ldots, A_,\}$ est un système complet d'événements.

De plus, $\forall k \in [1;n]\,, p_k \in]0,1[$ puisque les probabilités sont toutes non nulles.

D'après la question 3.2 précédente, le fonction $f : x \mapsto x \ln x$ est convexe sur]0 ;1[.

On pose $\forall k \in [1;n]\,, \lambda_k = \dfrac{1}{n} > 0$ de sorte que $\sum_{k=1}^{n} \lambda_k = 1$.

D'après la question 3.1 : $f\left(\sum_{k=1}^{n} \dfrac{1}{n} p_k\right) \leq \sum_{k=1}^{n} \dfrac{1}{n} f(p_k).$

$$\iff f\left(\frac{1}{n}\sum_{k=1}^{n} p_k\right) \leq \frac{1}{n}\sum_{k=1}^{n} f(p_k).$$

$$\iff f\left(\frac{1}{n}\right) \leq \frac{1}{n}\sum_{k=1}^{n} f(p_k).$$

$$\iff \cancel{\frac{1}{n}} \ln \frac{1}{n} \leq \cancel{\frac{1}{n}} \sum_{k=1}^{n} p_k \ln p_k.$$

$$\iff -\ln n \leq \sum_{k=1}^{n} p_k \ln p_k.$$

$$\iff \ln n \geq -\sum_{k=1}^{n} p_k \ln p_k = H(A).$$

Conclusion : $\boxed{H(A) \leq \ln n.}$

De plus, si l'on se place dans le cas particulier d'équiprobabilité, on a :

$$H(A) = -\frac{1}{n}\sum_{k=1}^{n} \ln \frac{1}{n} = -\frac{1}{\cancel{n}} \times \cancel{n} \times (-\ln n) = \ln n.$$

On en déduit que $H(A) \geq \ln n$, et donc d'après la conclusion précédente que $H(A) = \ln n$.

Conclusion : $\boxed{\textit{L'entropie est maximale lorsqu'il y a équiprobabilité. Ce maximum vaut } \ln n.}$