# DEVOIR SURVEILLÉ # 9
## le Vendredi 03/04/2020, 14h ⟶ 17h30

**Consignes** La qualité de la rédaction, la clarté et la précision des raisonnements, entreront pour une part importante dans l'appréciation de la copie. Les abréviations, sigles ou phrases nominales sont à proscrire. La numérotation des exercices (et des questions) doit être respectée et mise en évidence. Les résultats doivent être encadrés proprement.
Chaque candidat est responsable de la vérification de son sujet d'épreuve : pagination et impression de chaque page. Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il convient de le signaler sur la copie et de poursuivre la composition en expliquant les raisons des initiatives qui ont été prises.
L'usage de la calculatrice est interdit.

*Le sujet est composé de quatre parties indépendantes.*
*Les questions d'Informatique peuvent être traitées indépendamment des questions de mathématiques qui les précèdent.*

**Problème 1** ***Autour de la loi de Student*** Le but de ce problème est d'introduire et de discuter la loi de *Student*, notamment d'établir un intervalle de confiance utilisant cette loi. Le contexte est le suivant : soit $X \hookrightarrow \mathcal{N}(\mu,\sigma^2)$ avec $\mu \in \mathbf{R}$ et $\sigma > 0$. On note $(X_1,\dots,X_n)$ un $n$-échantillon de même loi que $X$, *i.e.* une famille de variables aléatoires réelles indépendantes et de même loi. On pose :

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i,\quad \sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X_n})^2,\quad \sigma_n^{2,\text{cor}} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X_n})^2,\quad T_n = \sqrt{n}\frac{\overline{X_n}-\mu}{\sigma_n^{\text{cor}}}.$$

respectivement, la moyenne empirique, la variance empirique, la variance empirique corrigée et la *statistique de Student* d'ordre $n$. On admettra dans le problème que le théorème central limite est également vérifié avec $\sigma_n^{2,\text{cor}}$ au lieu de $\sigma_n^2$[3].

### *Partie I— Résultats préliminaires.*

**1)** ***(Rappels de cours)***

**1.1.** Si $G_1,\dots G_n$ sont des variables aléatoires réelles indépendantes et de loi $\mathcal{N}(\mu_i,\sigma_i^2)$ avec $\mu_i \in \mathbf{R}$ et $\sigma_i > 0$, quelle est la loi de $a_1G_1 + \cdots a_nG_n$ pour tout $(a_1,\dots,a_n) \in \mathbf{R}^n$ ?

**1.2.** Si $G$ est de loi $\mathcal{N}(0,1)$ et $a \neq 0$, montrer que $aX+b \hookrightarrow \mathcal{N}(b,a^2)$ avec $(a,b) \in \mathbf{R}^2$. *On demande donc de redémontrer la propriété du cours.*

Les deux résultats ainsi rappelés seront utilisés tout au long du problème.

**2)** ***(Étude de la variance empirique non biaisée)***

**2.1.** Donner l'expression de $\mathbf{E}\left(X_i^2\right)$ en fonction de $\mu$ et $\sigma^2$ pour tout $i \in [\![1\,,\,n]\!]$.

**2.2.** Calculer l'espérance et la variance de $\overline{X_n}$. En déduire $\mathbf{E}\left(\overline{X_n}^2\right)$.

[3] C'est une conséquence d'une propriété de la convergence en loi appelée « lemme de Slutksy »

**2.3.** Justifier que : $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X_n}^2$. En déduire l'expression de l'espérance de $\sigma_n^2$.

**2.4.** En déduire l'espérance de $\sigma_n^{2,\text{cor}}$.

**2.5.** Pensez-vous que les statisticiens préfèrent utiliser $\sigma_n^{2,\text{cor}}$ ou $\sigma_n^2$ comme estimateur de $\sigma$ ?

**3)** Soit $G \hookrightarrow \mathcal{N}(0,1)$. On note $\Phi$ la fonction de répartition de la loi $\mathcal{N}(0,1)$ et $F_{G^2}$ la fonction de répartition de $G^2$.

**3.1.** Exprimer $F_{G^2}$ en fonction de $\Phi$.

**3.2.** En déduire qu'une densité de $G^2$ est donnée par :

$$\forall t \in \mathbf{R},\quad f_{G^2}(t) = \frac{1}{\sqrt{2\pi t}} e^{-t/2} \mathbb{1}_{\mathbf{R}^{+\star}}(t) = \begin{cases} \frac{1}{\sqrt{2\pi t}} e^{-t/2} & \text{si } t > 0, \\ 0 & \text{sinon.} \end{cases}.$$

## *Partie II— Statistique de Student, intervalle de confiance.*

**4)** Soit $n \in \mathbf{N}^\star$, on pose $G_n = \frac{X_n - \mu}{\sigma}$, et pour $n \geqslant 2$, $\overline{G_n} = \frac{1}{n}\sum_{i=1}^{n} G_i$.

**4.1.** Donner la loi de $G_n$ et de $\overline{G_n}$.

**4.2.** Montrer que : $T_n = \sqrt{n}\frac{\overline{G_n}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(G_i - \overline{G_n})^2}}$.

**4.3.** La loi de $T_n$ dépend-elle de $\mu$ et $\sigma$ ?

**4.4.** D'après le cours, de quelle loi usuelle la loi de $T_n$ doit-elle être proche lorsque $n$ est grand ? On donnera, avec précision, le théorème utilisé.

**5)** ***(Intervalle de confiance de théorème central limite en Python)*** Dans cette partie, on rappelle que la commande permettant de calculer $\Phi^{-1}(x)$ pour $x \in \mathbf{R}$, $\Phi$ étant la fonction de répartition de la loi $\mathcal{N}(0,1)$, est donnée ci-dessous.

```python
from scipy.stats import norm
norm.ppf(x)
```

**5.1.** Rappeler l'intervalle de confiance du cours pour la moyenne inconnue du $n$-échantillon $(X_1,\ldots,X_n)$, de niveau $1-\alpha$ avec $\alpha \in [0,1]$ en fonction notamment de $\sigma_n$. On expliquera soigneusement les notations et pour quel nombre d'observations celui-ci est acceptable.

Dans la suite de la question, on supposera qu'une simulation de l'échantillon $(X_1,\ldots,X_n)$ est donné sous forme d'une liste `D`.

**5.2.** Écrire une fonction Python `Stat_Moy` prenant en argument une liste `D` de réels, et qui retourne la moyenne de ces réels.

**5.3.** Écrire une fonction Python `Stat_Var` prenant en argument une liste `D` de réels, et qui renvoie une liste dont la première coordonnée est la variance, et la seconde est la variance corrigée.

**5.4.** Écrire une fonction Python `ICMoyenneTest` qui, étant donnés une liste de réels `D` et un float `mu`, renvoie :

❶ sous forme d'une liste l'intervalle de confiance rappelé précédemment pour la moyenne associé à la série de données contenue dans `D`.

❷ Une seconde variable booléenne renvoyant `True` si la valeur `mu` satisfait le test d'adéquation à la moyenne de niveau 95%.

**5.5.** $\boxed{\text{>\_ Python}}$ Écrire une fonction Python `Stat_Stud` prenant en argument une liste `D` de réels et `mu` un réel, qui renvoie la statistique $T_n$ de Student.

### *Partie III— Expression matricielle de la projection orthogonale sur un vecteur.*

On rappelle que $\mathbf{R}^n$ peut être muni de son produit scalaire usuel défini par :

$$\langle X|Y\rangle = {}^{\mathbf{T}}XY = \sum_{i=1}^{n} x_i.y_i$$

pour tout $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{R}^n$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbf{R}^n$. On note de plus $\|X\| = \sqrt{\sum_{i=1}^{n} x_i^2}$ la norme euclidienne d'un vecteur $X$ de $\mathbf{R}^n$ et $\mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, $F = \mathrm{Vect}(\mathbb{1})$. On rappelle que $F^{\perp}$ est l'ensemble des vecteurs de $\mathbf{R}^n$ orthogonaux à ceux de $F$. On note $M$ la matrice de format $n \times n$ dont tous les coefficients valent $1/n$, *i.e.*

$$M = \frac{1}{n}\left(\mathbb{1} \mid \cdots \mid \mathbb{1}\right) = \frac{1}{n}\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

**6)** Justifier qu'il existe $P \in \mathscr{M}_n(\mathbf{R})$ telle que : ${}^{\mathbf{T}}PP = I_n$ et ${}^{\mathbf{T}}PMP$ soit diagonale. Que signifie la condition ${}^{\mathbf{T}}PP = I_n$ sur les colonnes de $P$ ? *Dans la suite on va proposer une construction explicite de la matrice $P$.*

On admettra dans cette partie que toute famille libre de $\mathbf{R}^n$ peut être orthonormalisée. On notera également pour $M \in \mathfrak{M}_n(\mathbf{R})$ :

$$\mathrm{Ker}\, M = \left\{ (x_1, \dots, x_n) \in \mathbf{R}^n, \quad M\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \right\}, \quad \mathrm{Im}\, M = \left\{ (y_1, \dots, y_n) \in \mathbf{R}^n, \quad \exists X \in \mathfrak{M}_{n,1}(\mathbf{R}), \quad MX = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\}.$$

Ce sont donc des sous-ensembles de $\mathbf{R}^n$.

**7)** ***(Étude de la valeur propre zéro de $M$)***

**7.1.** Montrer que : $F^{\perp} = \{X = (x_1, \dots, x_n) \in \mathbf{R}^n, \quad x_1 + \dots + x_n = 0\}$.

**7.2.** En déduire que : $F^{\perp} = \mathrm{Ker}\, M$.

**7.3.** Que vaut le rang de $M$ ? En déduire la dimension de l'espace propre associé à la valeur propre 0.

**7.4.** Considérons $\varphi \left| \begin{array}{ccc} \mathbf{R}^n & \longrightarrow & \mathbf{R} \\ (x_1, \dots, x_n) & \longmapsto & x_1 + \cdots + x_n \end{array} \right.$.

Justifier que $\varphi$ est une forme linéaire. Retrouver la valeur de $\dim F^{\perp} = \dim \mathrm{Ker}\, M$ en utilisant $\varphi$.

**8)** ***(Étude des éventuelles autres valeurs propres de $M$)***

**8.1.** Justifier sans calculs que $\mathbb{1}$ est vecteur propre de $M$. Pour quelle valeur propre ? Quelle est la dimension de l'espace propre associé ? Y-a-t-il d'autres valeurs propres ?

**8.2.** Montrer que : $\quad \operatorname{Im} M = F$.

**9)** ***(Conclusion)*** Justifier alors que $P$ peut être choisie de la forme $P = \begin{pmatrix} P_1 \mid \cdots \mid P_n \end{pmatrix}$ où $P_1, \ldots, P_n \in$ $\mathfrak{M}_{n,1}(\mathbf{R})$, $\quad P_n = \dfrac{1}{\sqrt{n}} \mathbb{1} \quad$ et $\quad {}^{\mathbf{T}}PMP = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1 \end{pmatrix}$.

Dans la suite, on notera $\mathscr{P} = ({}^{\mathbf{T}}P_1, \ldots, {}^{\mathbf{T}}P_n)$ la famille de $\mathbf{R}^n$ des colonnes de $P$ transposées en lignes : on montre sans difficultés que puisque $(P_1, \ldots, P_n)$ est une base de $\mathfrak{M}_{n,1}(\mathbf{R})$, alors $\mathscr{P}$ est une base de $\mathbf{R}^n$. On notera également ${}^{\mathbf{T}}F \underset{\textbf{(défi.)}}{=} \operatorname{Vect}({}^{\mathbf{T}}\mathbb{1}) = \operatorname{Vect}((1, \ldots, 1))$.

**10)** On notera $p \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$ l'unique endomorphisme vérifiant $\underset{\mathscr{B}^{\mathbf{can}}}{\mathscr{M}\mathrm{at}}(p) = M$. Montrer que $M$ est, relativement à la base canonique $\mathscr{B}^{\mathbf{can}}$ de $\mathbf{R}^n$, la matrice de la projection orthogonale sur ${}^{\mathbf{T}}F$.
*Indication : On étudiera $p(x)$ et $p(x) - x$ pour $x \in \boldsymbol{R}^n$.*

**11)** Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ la matrice des coordonnées dans la base canonique d'un certain vecteur $z$ de $\mathbf{R}^n$. On note alors $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ la matrice de ce même vecteur mais cette fois-ci dans la base $\mathscr{P}$ de $\mathbf{R}^n$.

**11.1.** Rappeler la relation entre $P, X, Y$.

**11.2.** Quelle est la dernière ligne de $P^{-1}$ ? En déduire que la dernière coordonnée de $Y$ a pour expression : $\quad y_n = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^{n} x_i$.

**11.3.** Justifier que pour tout $k \in [\![1\,,\,n-1]\!]$, $\quad y_k = {}^{\mathbf{T}}P_k X$.

**11.4.** Montrer que $\quad x_1^2 + \cdots + x_n^2 = y_1^2 + \cdots + y_{n-1}^2 + y_n^2$. *Indication : On pourra calculer ${}^{\mathbf{T}}Y.Y$ en fonction de $X$ à l'aide de la question* ***11.1.***.

**12)** On reprend les notations de la question précédente et on note de plus $\quad \overline{x_n} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i$.

**12.1.** ***(König-Huygens empirique)*** Justifier que $\quad \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} (x_i - \overline{x_n})^2 = \dfrac{1}{n} \sum_{i=1}^{n} x_i^2 - \overline{x_n}^2$.

**12.2.** Montrer que :

$$y_1^2 + \cdots + y_{n-1}^2 = \sum_{i=1}^{n} \left(x_i - \overline{x_n}\right)^2.$$

*Indication : On pourra écrire $y_1^2 + \cdots + y_{n-1}^2 = \left(y_1^2 + \cdots + y_n^2\right) - y_n^2$.*

## *Partie IV— Vecteurs gaussiens.*

Les échantillons de lois gaußiennes interviennent dans de nombreux problèmes en inférence statistique. Ces échantillons ont donc été étudiés à part entière, on les appelle des *vecteurs gaußiens* ou *échantillons gaußiens*.

Soient $n \in \mathbf{N}^\star$ et $\mathscr{G} = \begin{pmatrix} G_1 \\ \vdots, G_n \end{pmatrix}$ un $n$-échantillon gaussien de loi commune $\mathscr{N}(0,1)$. On rappelle que $P = (P_1 \mid \cdots \mid P_n)$ désigne la matrice construite dans la partie précédente, écrite en colonnes. Pour $k \in [\![1\,,\,n]\!]$, on pose

$$H_k = {}^{\mathbf{T}}P_k\mathscr{G}.$$

On admet dans la suite que les $H_k$ sont mutuellement indépendantes.

**13)** Montrer que $H_k$ suit une loi gaußienne et calculer les paramètres. On pose dans la suite $n \geqslant 2$,

$$\overline{G_n} = \frac{1}{n}\sum_{i=1}^{n} G_i \quad \text{et} \quad W_n = \sum_{i=1}^{n}(G_i - \overline{G_n})^2.$$

**14)** Montrer, en utilisant les résultats de la Partie III, que :

$$W_n = \sum_{i=1}^{n-1} H_i^2, \quad \overline{G_n} = \frac{1}{\sqrt{n}}H_n.$$

**15)** En déduire, à l'aide de la question précédente et du lemme des coalitions, que $\overline{G_n}$ et $W_n$ sont indépendantes.

**16)** En reprenant les notations de la partie 1, exprimer $T_n$ en fonction de $H_n$ et $W_n$ et en déduire, en supposant que $T_n$ admet une espérance, la valeur de l'espérance de $T_n$.

## *Partie V— Loi de $W_n$ ou loi du khi-deux.*

On a vu dans la partie précédente que $W_n = \sum_{i=1}^{n-1} H_i^2$ est la somme de $n-1$ carrés de variables gaußiennes centrées réduites indépendantes. On note dans la suite $d = n-1$ et on s'intéresse à la somme de $d$ gaußiennes notée

$$Y_d = \sum_{i=1}^{d} H_i^2.$$

Cette loi est appelée *loi du Khi-deux à $d$ degrés de liberté*. On rappelle que le cas $d = 1$ a été traité dans la première partie.

**17)** Supposons que $d = 2$. On note $f_i$ la densité de $H_i^2$ pour $i \in \{1,2\}$.

**17.1.** Rappeler la formule du produit de convolution, donnant une densité de $f_{Y_2}$, en fonction de $f_1, f_2$.

**17.2.** En déduire une densité de $Y_2$. Le résultat pourra faire intervenir l'intégrale de l'inverse d'une racine carrée.

**18)** On admet de manière générale pour $d \geqslant 2$ qu'une densité du Khi-deux à $d$ degrés de liberté est donnée par la fonction suivante :

$$f_{Y_d}(t) = \frac{1}{2^{d/2}\Gamma(d/2)} t^{d/2-1}e^{-t/2}\mathbb{1}_{\mathbf{R}^+}(t), \quad \Gamma(d/2) = \int_0^{\infty} t^{d/2-1}e^{-t}\,\mathrm{d}t.$$

Justifier la convergence de l'intégrale $\Gamma(d/2)$ pour $d \geqslant 2$.

## *Partie VI— Intervalle de confiance et test.*

À partir de la loi de $W_n$, il est possible d'obtenir la loi de $T_n$ et donc de tracer les fonctions ci-dessous.

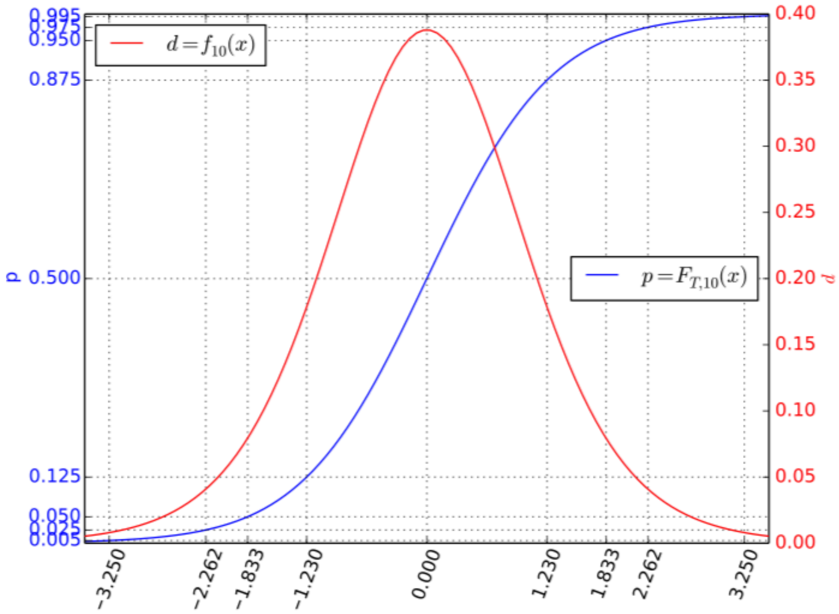


FIG. DSIC.1.3 : Densité $f_{10}$ et fonction de répartition de $F_{T,10}$ de $T_{10}$

Le graphique ci-contre montre la fonction de répartition de $T_n$ pour $n = 10$. L'objectif est ici de l'utiliser afin de mettre en place un test d'adéquation à la moyenne utilisant la statistique $T$, et une autre zone de rejet que celle du cours.
***Remarque* — la graduation de droite correspond à l'échelle de la densité, celle de gauche à la fonction de répartition.**

**19)19.1.** À votre avis, quelle est la parité de la densité ?

**19.2.** En admettant la conjecture faite précédemment, trouver une relation entre $\mathbf{P}(|T_{10}| < u)$ et $F_{T,10}(u)$ pour tout $u \in \mathbf{R}$.

**20)** ***(Recherche de quantiles)***

**20.1.** Déterminer un $u_1$ tel que : $\mathbf{P}(|T_{10}| < u_1) = 95\%$.

**20.2.** Déterminer un $u_2$ tel que : $\mathbf{P}(|T_{10}| < u_2) = 99\%$.

**21)** En déduire un intervalle de confiance non asymptotique, de niveau $0,05$ ou $0,01$ pour la moyenne $\mu$ inconnue.

**22)** En déduire un test statistique de niveau $0,05$ ou $0,01$ permettant de tester l'hypothèse $\mathscr{H}_0$ « $\mu = \mu_0$ » avec $\mu_0 \in \mathbf{R}$.

**23)** ***(Comparaison avec l'intervalle de confiance obtenu via le théorème central limite)*** On rappelle que, si $\Phi$ désigne la fonction de répartition de la loi $\mathscr{N}(0,1)$, nous avons : $\Phi^{-1}(1 - 0,05/2) = 1.96$.
Déterminer un intervalle de confiance asymptotique pour la moyenne de niveau $\alpha = 0,05$. Quelle est la principale différence avec l'intervalle précédent ? Comparer leur longueur lorsque $\alpha = 0,05$. Lequel est plus précis ?

## Correction

### Solution (Problème 1)

#### *Partie I— Résultats préliminaires.*

**1)** **1.1.** Puisque les variables aléatoires $G_i$ sont indépendantes, le cours nous dit que $a_1G_1+\cdots+a_nG_n$ suit encore une loi normale et de paramètres l'espérance et la variance. Ainsi, puisque $\mathbf{E}\left(a_1G_1+\cdots+a_nG_n\right)=a_1\mu_1+\cdots+a_n\mu_n=0$ et $\mathbf{Var}\left(a_1G_1+\cdots+a_nG_n\right)=a_1^2\mathbf{Var}\left(G_1\right)+\cdots+a_n^2\mathbf{Var}\left(G_n\right)=\sum_{i=1}^n a_i^2\sigma_i^2$ par indépendance et propriété de la variance, donc

$$\boxed{a_1G_1+\cdots+a_nG_n\hookrightarrow\mathcal{N}\left(\sum_{i=1}^n a_i\mu_i,\sum_{i=1}^n a_i^2\sigma_i^2\right).}$$

**1.2.** Soit $X\hookrightarrow\mathcal{N}(0,1)$. Alors calculons la fonction de répartition de $aX+b$. Comme $(aX+b)(\Omega)=\mathbf{R}$, soit $t\in\mathbf{R}$,

$$\begin{aligned}\mathbf{P}\left(aX+b\leqslant t\right)&=\mathbf{P}\left(aX+b\leqslant t-b\right)\\&=\begin{cases}\mathbf{P}\left(X\leqslant\dfrac{t-b}{a}\right)=\Phi\left(\dfrac{t-b}{a}\right)&\text{si } a>0,\\\mathbf{P}\left(X\geqslant\dfrac{t-b}{a}\right)=1-\Phi\left(\dfrac{t-b}{a}\right)&\text{si } a<0,\end{cases}\end{aligned}$$

où $\Phi$ désigne la fonction de répartition de la loi $\mathcal{N}(0,1)$.

— Si $a>0$; la fonction $t\longmapsto\Phi\left(\dfrac{t-b}{a}\right)$ est continue et $\mathscr{C}^1$ sur $\mathbf{R}$ — mais $\mathscr{C}^1$ sauf en un nombre fini de points suffirait — donc $aX+b$ est une variable aléatoire réelle à densité et une densité est donnée par sa dérivée $t\longmapsto\dfrac{1}{a}\varphi\left(\dfrac{t-b}{a}\right)=\dfrac{1}{a\sqrt{2\pi}}\mathrm{e}^{-\frac{\left(\frac{t-b}{a}\right)^2}{2}}=\dfrac{1}{a\sqrt{2\pi}}\mathrm{e}^{-\frac{(t-b)^2}{2a^2}}$, $\varphi$ désignant la densité de la $\mathcal{N}(0,1)$.

— Si $a<0$; on montre de-même que la fonction de répartition obtenue précédemment est celle d'une variable aléatoire réelle à densité, et qu'une densité est donnée par $t\longmapsto-\dfrac{1}{a}\varphi\left(\dfrac{t-b}{a}\right)=-\dfrac{1}{a\sqrt{2\pi}}\mathrm{e}^{-\frac{(t-b)^2}{2a^2}}$.

*In fine*, nous avons obtenu la densité suivante :

$$t\longmapsto\frac{1}{|a|\sqrt{2\pi}}\mathrm{e}^{-\frac{(t-b)^2}{2a^2}}.$$

C'est la densité d'une loi $\mathcal{N}(b,a^2)$. Donc $\boxed{aX+b\hookrightarrow\mathcal{N}(b,a^2).}$

**2)** **2.1.** Par définition de la variance nous avons $\boxed{\mathbf{E}\left(X_i^2\right)=\mu^2+\sigma^2.}$

**2.2.** Par linéarité de l'espérance, $\mathbf{E}\left(\overline{X_n}\right)=\dfrac{1}{n}\sum_{i=1}^n\mathbf{E}\left(X_i\right)=\dfrac{1}{n}\sum_{i=1}^n\mu=\boxed{\mu.}$ Par indépendance et propriété de la variance, nous avons

$$\mathbf{Var}\left(\overline{X_n}\right)=\frac{1}{n^2}\sum_{i=1}^n\mathbf{Var}\left(X_i\right)=\frac{1}{n^2}\sum_{i=1}^n\sigma^2=\boxed{\frac{\sigma^2}{n}}.$$

Alors :

$$\mathbf{E}\left(\overline{X_n}^2\right)=\mathbf{Var}\left(\overline{X_n}\right)+\mathbf{E}\left(\overline{X_n}\right)=\boxed{\frac{\sigma^2}{n}+\mu^2}.$$

**2.3.** On développe le carré :

$$\begin{aligned}
\sigma_n^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\overline{X_n} + \overline{X_n}^2)\\
&= \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - 2\left(\sum_{i=1}^{n} X_i\right)\overline{X_n} + n\overline{X_n}^2\right)\\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\overline{X_n}^2 + \overline{X_n}^2\\
&= \boxed{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X_n}^2.}
\end{aligned}$$

On déduit alors par linéarité de l'espérance

$$\mathbf{E}\left(\sigma_n^2\right) = \frac{1}{n}(n\sigma^2 + n\mu^2) - \frac{\sigma^2}{n} + \mu^2 = \boxed{\frac{n-1}{n}\sigma^2}.$$

**2.4.**

**2.5.** Ils préfèrent utiliser l'estimateur corrigé car il est non biaisé.

**3) 3.1.** Nous avons $G^2(\Omega) = \mathbf{R}^+$, soit de plus $x \in \mathbf{R}^+$, alors

$$\begin{aligned}
\mathbf{P}\left(G^2 \leqslant x\right) &= \mathbf{P}\left(-\sqrt{x} \leqslant G \leqslant \sqrt{x}\right)\\
&= \Phi(\sqrt{x}) - \Phi(-\sqrt{x})\\
&= \Phi(\sqrt{x}) - (1 - \Phi(\sqrt{x}))\\
&= 2\Phi(\sqrt{x}) - 1
\end{aligned}$$

Donc

$$\boxed{F_G^2(x) = \begin{cases} 2\Phi(\sqrt{x}) - 1 & \text{si } x \geqslant 0,\\ 0 & \text{sinon.}\end{cases}}$$

**3.2.** La fonction $\Phi$ est de classe $\mathscr{C}^\infty$ d'après le cours, donc par composée $F_{G^2}$ est une fonction continue et $\mathscr{C}^1$. Donc $G^2$ est à densité et une densité est donnée par la dérivée

$$x \longmapsto \frac{1}{\sqrt{x}}\varphi(\sqrt{x})\mathbb{1}_{\mathbf{R}^{+\star}(x)}.$$

Donc :

$$\boxed{f_{G^2}(t) = \begin{cases} \dfrac{1}{\sqrt{2\pi t}}\mathrm{e}^{-t/2} & \text{si } t > 0,\\ 0 & \text{sinon.}\end{cases}}$$

## *Partie II— Statistique de Student, intervalle de confiance.*

**4) 4.1.** D'après les résultats rappelés plus haut, comme $G_n$ est affine en $X_n$, elle suit une loi normale de paramètres espérance et variance : $\boxed{G_n \hookrightarrow \mathcal{N}(0,1)}$.

Par propriétés de stabilité de la loi normale (les $X_i$ sont indépendantes), nous savons aussi que $\overline{G_n}$ suit également une loi normale, plus précisément, comme

$$\mathbf{E}\left(\overline{G_n}\right) = \mathbf{E}\left(X_1\right) = \mu, \quad \mathbf{Var}\left(\overline{G_n}\right) = \frac{1}{n}\mathbf{Var}\left(X_1\right) = \frac{1}{n}\sigma^2.$$

D'où : $\boxed{\overline{G_n} \hookrightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)}$.

**4.2.** Calculons $T_n$ en remplaçant $X_n$ par son expression en fonction des $G_n$.
On a $X_n = \sigma G_n + \mu$ pour tout entier $n$. Ainsi :

$$\begin{aligned}
T_n &= \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2}} \\
&= \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n}(\sigma G_i + \mu) - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2}} \\
&= \sqrt{n}\frac{\sigma\overline{G_n} + \mu - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\sigma G_i + \mu - \sigma\overline{G_n} - \mu)^2}} \\
&= \sqrt{n}\frac{\sigma\overline{G_n}}{\sqrt{\frac{\sigma^2}{n-1}\sum_{i=1}^{n}(G_i - \overline{G_n})^2}} \\
&= \boxed{\sqrt{n}\frac{\overline{G_n}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(G_i - \overline{G_n})^2}}}.
\end{aligned}$$

**4.3.** Cette nouvelle expression ne fait plus apparaître $\mu$ et $\sigma^2$, et la loi des $G_i$ pour tout $i$, et celle $\overline{G_n}$ ne dépendent ni de $\mu$ ni de $\sigma$. Donc $\boxed{\text{la loi de } T_n \text{ ne dépend ni de } \mu \text{ ni de } \sigma.}$

**4.4.** Reprenant l'expression première, nous savons que d'après le théorème central limite (seconde version, qui s'applique car les $X_i$ sont i.i.d. et admettent un moment d'ordre deux), la loi de $T_n$ est proche de la $\mathcal{N}(0,1)$.

**5)**

**5.1.** L'intervalle du cours, asymptotique de niveau $\alpha$ pour l'espérance qui découle du théorème central limite, est le suivant :

$$\boxed{\left[\overline{X_n} - \Phi^{-1}(1-\alpha/2)\frac{\sigma_n}{\sqrt{n}}, \overline{X_n} + \Phi^{-1}(1-\alpha/2)\frac{\sigma_n}{\sqrt{n}}\right]}.$$

**5.2.**

```python
def Stat_Moy(D):
    S = 0
    for x in D:
        S += x
    return S/len(D)
```

**5.3.**

```python
def Stat_Var(D):
    S = 0
    for x in D:
        S += x**2
    V = S/len(D)-Stat_Moy(D)**2
    return [V, (n/(n-1))*V]
```

**5.4.**

```python
from math import sqrt
from scipy.stats import norm
def ICMOyenneTest(D,mu):
    u = norm.ppf(1-0.05/2)
    a = Stat_Moy(D)-u*(sqrt(Stat_Var(D)[0]))/(sqrt(len(D)))
    b = Stat_Moy(D)+u*(sqrt(Stat_Var(D)[0]))/(sqrt(len(D)))
    return [a,b, (mu<b) & (mu>a)]
```

**5.5.**

```python
from math import sqrt
def StatStud(D, mu):
    return sqrt(len(D))*(Stat_Moy(D)-mu)/sqrt(Stat_Var(D)[1])
```

***Partie III— Expression matricielle de la projection orthogonale sur un vecteur..***

**6)** La matrice $P$ est clairement symétrique. D'après le théorème spectral, il existe une matrice $P$ orthogonale — *i.e.* correspondant à un changement de base entre la base canonique et une base orthonormée — telle que ${}^{\mathsf{T}}PMP$ soit diagonale.

**7)** **7.1.** Nous avons, puisque $F = \operatorname{Vect} \mathbb{1}$ :

$$X \in F^\perp \iff \forall \lambda \in \mathbf{R},\quad \langle X | \lambda \mathbb{1} \rangle = 0 = \lambda \langle X | \mathbb{1} \rangle \iff X \perp \mathbb{1} \iff X = (x_1, \dots, x_n),\quad \sum_{i=1}^n x_i 1 = \sum_{i=1}^n x_i = 0.$$

En conclusion, nous avons bien :

$$\boxed{F^\perp = \{X = (x_1, \dots, x_n) \in \mathbf{R}^n,\quad x_1 + \dots + x_n = 0\}.}$$

**7.2.** Puisque $X \in \operatorname{Ker} M$ si et seulement si $MX = 0 = \dfrac{1}{n}(x_1 + \cdots + x_n)$ en notant de nouveau $X = (x_1, \dots, x_n)$, nous déduisons immédiatement que $\boxed{\operatorname{Ker} M = F^\perp.}$

**7.3.** Toutes les colonnes de $M$ sont identiques, et $\mathbb{1} \neq 0$ donc $\boxed{\operatorname{Rg} M = 1}$. D'après le théorème du rang, $\dim \operatorname{Ker} M = \dim E_0(M) = n - 1$, donc $\boxed{E_0(M) \text{ est un hyperplan de } \mathbf{R}^n.}$

**7.4.** L'application $\varphi$ est linéaire. En effet, soit $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{R}^n$ et $\lambda, \mu \in \mathbf{R}$, alors :

$$\begin{aligned}\varphi(\lambda(x_1, \dots, x_n) + \mu(y_1, \dots, y_n)) &= \varphi(\lambda x_1 + \mu y_1, \dots, \lambda x_n + \mu y_n) \\ &= \sum_{i=1}^n (\lambda x_i + \mu y_i) = \lambda \varphi(x_1, \dots, x_n) + \mu \varphi(y_1, \dots, y_n).\end{aligned}$$

Donc $\boxed{\varphi \text{ est une application linéaire.}}$ D'après le théorème du rang,

$$\dim \operatorname{Ker} \varphi + \operatorname{Rg} \varphi = \dim \mathbf{R}^n = n.$$

Or, puisque $\operatorname{Im}(\varphi) \subset \mathbf{R}$ et $\varphi$ est non nulle, l'image est de dimension 1. Donc $\operatorname{Im}(\varphi) = \mathbf{R}$, elle est donc de rang 1. Et on obtient finalement

$$\boxed{\dim \operatorname{Ker} \varphi = n - 1 = \dim \operatorname{Ker}(M) = \dim F^\perp.}$$

**8)** **8.1.** La somme sur une ligne étant constante égale à $\dfrac{1}{n} n = 1$, nous avons alors $M\mathbb{1} = \mathbb{1}$. Donc $\boxed{\mathbb{1} \text{ est un vecteur propre associé à la valeur propre } 1.}$
Comme 0 est une valeur propre d'espace propre associé de dimension $n - 1$, et que la somme des dimensions des espaces propres est inférieure ou égale à la dimension de l'espace ambiant (ici $n$), nous avons forcément

$$\dim E_1(M) \leqslant 1.$$

De plus, par ce qui précède, $\dim E_1(M) \geqslant 1$. On déduit alors que $\boxed{\dim E_1(M) = 1}$. Il n'y a pas d'autres valeurs propres.

**8.2.** Nous avons déjà établi que $\dim \operatorname{Im} F = 1$ et $\dim F = 1$ donc il suffit d'établir une inclusion sur les deux. Soit $Y = (y_1, \dots, y_n) \in \operatorname{Im} F$. Alors il existe $(x_1, \dots, x_n) \in F$ tel que

$$\forall i \in [\![1\,,\,n]\!], \quad y_i = \frac{1}{n}(x_1 + \dots, x_n) = \lambda \mathbb{1} \quad \text{avec} \quad \lambda = \frac{1}{n}(x_1 + \dots, x_n).$$

Donc $\operatorname{Im} M \subset F$ et il y a égalité des dimensions donc $\boxed{\operatorname{Im} M = F}$.

**9)** On choisit pour $(P_1, \dots, P_{n-1})$ une base orthonormée (existence admise en BCPST) de $E_0(M)$, le dernier espace propre étant $E_1(M) = \operatorname{Im} F$, on pourrait prendre $P_n = \mathbb{1}$, mais ce vecteur n'est pas de norme un. On choisit donc sa version normalisée : $P_n = \dfrac{\mathbb{1}}{\|\mathbb{1}\|} = \boxed{\dfrac{\mathbb{1}}{\sqrt{n}}}$.

Il reste à montrer que la famille ainsi construite est bien orthonormée pour avoir la relation $P^{-1} = {}^{\mathbf{T}}P$.
Les vecteurs colonnes $P_i$ sont tous de norme 1 par construction. Soit $(i, j) \in [\![1, n]\!]^2$. Montrons que $\left\langle P_i \middle| P_j \right\rangle = 0$ si $i \neq j$.

— 1er cas : si $i = n$ et $j \in [\![1\,,\,n-1]\!]$. Alors comme nous avons montré que $F^\perp = \operatorname{Ker} M = E_0(M)$, $E_0(M)$ est orthogonal à $E_1(M)$. Donc $\left\langle P_i \middle| P_j \right\rangle = 0$.

— 2ème cas. Si $i \neq j$ avec $(i, j) \in [\![1\,,\,n-1]\!]^2$, $\left\langle P_i \middle| P_j \right\rangle = 0$ puisque $(P_1, \dots, P_{n-1})$ est, par construction, une base orthonormée de $\operatorname{Ker} M = E_0(M)$.

En conclusion, la base ainsi construite est orthonormée, donc la matrice de passage $P$ vérifie la relation ${}^{\mathbf{T}}P = P^{-1}$. C'est terminé.

**10)** Il faut montrer que $p(x) \in {}^{\mathbf{T}}F$ et $p(x) - x$ est dans l'orthogonal de ${}^{\mathbf{T}}F = \operatorname{Vect}((1, \dots, 1))$ *i.e.* que $\langle p(x) - x | (1, \dots, 1) \rangle = 0$, en faisant intervenir $M$.

— $p(x) \in F$ : en effet, en notant $X = \underset{\mathscr{B}^{\text{can}}}{\mathscr{M}\text{at}}(x)$, on a

$$\underset{\mathscr{B}^{\text{can}}}{\mathscr{M}\text{at}}\big(p(x)\big) = MX = \begin{pmatrix} \frac{1}{n}(x_1 + \cdots + x_n) \\ \vdots \\ \frac{1}{n}(x_1 + \cdots + x_n) \end{pmatrix} \in \operatorname{Vect}(\mathbb{1}) = F.$$

Or, $\underset{\mathscr{B}^{\text{can}}}{\mathscr{M}\text{at}}\big(p(x)\big) = {}^{\mathbf{T}}\big(p(x)\big)$ (en effet, la matrice d'un vecteur dans la base canonique est simplement les coordonnées du vecteur mais écrits en colonne, donc la transposée), et on obtient alors ${}^{\mathbf{T}}\big(p(x)\big) \in \operatorname{Vect}(\mathbb{1})$, *i.e.* $\boxed{p(x) \in \operatorname{Vect}((1, \dots, 1)) = {}^{\mathbf{T}}F}$.

— $x - p(x) \in F^\perp$ ; montrons que $\langle p(x) - x | (1, \dots, 1) \rangle = 0$. Or, puisque la base canonique est une base orthonormée de $\mathbf{R}^n$, le cours nous dit que :

$$\begin{aligned}
\langle p(x) - x | (1, \dots, 1) \rangle &= {}^{\mathbf{T}}\left( \underset{\mathscr{B}^{\text{can}}}{\mathscr{M}\text{at}}(p(x) - x) \right) \underset{\mathscr{B}^{\text{can}}}{\mathscr{M}\text{at}}((1, \dots, 1)) \\
&= {}^{\mathbf{T}}(MX - X)\mathbb{1} = ({}^{\mathbf{T}}X\,{}^{\mathbf{T}}M - {}^{\mathbf{T}}X)\mathbb{1} \\
&= {}^{\mathbf{T}}X \underbrace{M\mathbb{1}}_{=\mathbb{1}} - {}^{\mathbf{T}}X\mathbb{1}\mathbb{1} \qquad M \text{ symétrique,} \\
&= {}^{\mathbf{T}}X\mathbb{1} - {}^{\mathbf{T}}X\mathbb{1} = 0.
\end{aligned}$$

Donc, par définition d'une projection orthoognale : $\boxed{M \text{ est la matrice de la projection orthogonale sur } {}^{\mathbf{T}}F \text{ dans } \mathscr{B}^{\text{can}}}$.

**11)11.1.** On sait que $\mathrm{P}^{\mathscr{B}^{\text{can}} \to \mathscr{P}} = P$ puisqu'il s'agit de la matrice des coordonnées **en colonne** de la famille $\mathscr{P}$.
Le cours nous donne alors $\boxed{Y = P^{-1}X.}$

**11.2.** Comme $P^{-1} = {}^{\mathbf{T}}P$, les coefficients de la dernière ligne de $P^{-1}$ sont ceux de la dernière colonne de $P$, donc elle vaut : $\boxed{\dfrac{1}{\sqrt{n}}\begin{pmatrix} 1 & \dots & 1 \end{pmatrix}}$.

La dernière ligne du produit matriciel $Y = {}^{\mathbf{T}}PX$ donne donc : $\boxed{y_n = \dfrac{1}{\sqrt{n}}\begin{pmatrix} 1 & \dots & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \dfrac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i.}$

**11.3.** Les $y_k$ sont les coordonnées de $z$ dans la base $\mathscr{P}$. Or, cette base est orthonormée, donc d'après le cours les composantes sont données par les produits scalaires selon cette base :

$$y_k = \langle {}^{\mathbf{T}}P_k | z \rangle.$$

Or, puisque $\mathscr{B}^{\mathbf{can}}$ est une base orthonormée, on en déduit d'après le cours que :

$$\langle {}^{\mathbf{T}}P_k | z \rangle = {}^{\mathbf{T}}\Bigg( \underbrace{\underset{\mathscr{B}^{\mathbf{can}}}{\mathscr{M}\mathrm{at}}\left({}^{\mathbf{T}}P_k\right)}_{=P_k} \Bigg) \underset{\mathscr{B}^{\mathbf{can}}}{\mathscr{M}\mathrm{at}}(z) = \boxed{{}^{\mathbf{T}}P_k Y}.$$

**11.4.** Reprenons la relation $Y = {}^{\mathbf{T}}PX$ déjà établie. Alors :

$${}^{\mathbf{T}}YY = {}^{\mathbf{T}}\left({}^{\mathbf{T}}PX\right){}^{\mathbf{T}}PX = {}^{\mathbf{T}}XP{}^{\mathbf{T}}PX = {}^{\mathbf{T}}XX.$$

Écrivant cette égalité avec les coordonnées des deux vecteurs, il vient :

$$\boxed{x_1^2 + \cdots + x_n^2 = y_1^2 + \cdots + y_{n-1}^2 + y_n^2}.$$

**12) 12.1.** On développe le carré.

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x_n})^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\overline{x_n} + \overline{x_n}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x_n}\frac{2}{n}\sum_{i=1}^{n}x_i + \overline{x_n}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\overline{x_n}^2 + \overline{x_n}^2 \\
&= \boxed{\frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x_n}^2.}
\end{aligned}$$

**12.2.** Utilisant l'indication, il vient :

$$\begin{aligned}
y_1^2 + \cdots + y_{n-1}^2 &= \left(y_1^2 + \cdots + y_n^2\right) - y_n^2 \\
&= x_1^2 + \cdots + x_n^2 - y_n^2 \\
&= \sum_{i=1}^{n}x_i^2 - y_n^2 = \sum_{i=1}^{n}(x_i - \overline{x_n})^2 + n\overline{x_n}^2 - y_n^2 \\
&= \sum_{i=1}^{n}(x_i - \overline{x_n})^2 + n\overline{x_n}^2 - \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i\right)^2 = \boxed{\sum_{i=1}^{n}(x_i - \overline{x_n})^2}.
\end{aligned}$$

## *Partie IV— Vecteurs Gaussiens.*

**13)** En effectuant le produit matriciel, notant $P_k = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $a_i \in \mathbf{R}$ pour tout $i$, on constate que $H_k = \sum_{i=1}^{n} a_i G_i$, c'est donc une combinaison linéaire de lois normales indépendantes ; donc d'après une partie précédente elle suit encore une loi normale. Plus précisément

$$H_k \hookrightarrow \mathscr{N}\left(0, \sum_{i=1}^{n} a_i^2\right).$$

Mais on peut même préciser encore un peu plus : $\sum_{i=1}^{n} a_i^2 = {}^{\mathbf{T}}(P_k)\,P_k = \left[{}^{\mathbf{T}}P.P\right]_{k,k} = 1$ — autre argument : on a vu dans le cours que les colonnes d'une matrice orthogonale forme une base orthonormée, on peut donc l'affirmer de suite. Donc $\boxed{H_k \hookrightarrow \mathscr{N}(0,1).}$

**14)** On applique les résultats précédents à $X = \begin{pmatrix} G_1 \\ \vdots \\ G_n \end{pmatrix}$ ainsi que $Y = P^{-1}X = {}^{\mathbf{T}}PX = \begin{pmatrix} H_1 \\ \vdots \\ H_n \end{pmatrix}$. Alors :

$$\begin{aligned} W_n &= \sum_{i=1}^{n} (G_i - \overline{G_n})^2 \\ &= \sum_{i=1}^{n-1} y_i^2 = \boxed{\sum_{i=1}^{n-1} H_i^2}. \end{aligned}$$

Et toujours d'après la partie précédente, la dernière coordonnée de $Y$, qui est $H_n$, s'écrit

$$H_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i = \boxed{\sqrt{n}\overline{G_n}}.$$

**15)** On sait que $\overline{G_n}$ est une fonction de $H_n$, par ailleurs $W_n$ est une fonction de $H_1, \ldots, H_{n-1}$, or les $H_i$ sont supposées mutuellement indépendantes. Donc par lemme des coalitions, $\boxed{\overline{G_n} \text{ et } W_n \text{ sont indépendantes.}}$

**16)** On a $T_n = \dfrac{\sqrt{n}\overline{G_n}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(G_i - \overline{G_n})^2}} = \boxed{\sqrt{n}\dfrac{\overline{G_n}}{\sqrt{\frac{1}{n-1}W_n}}}$. Sous réserve d'existence (admise d'après l'énoncé),

$$\mathbf{E}\left(T_n\right) = \sqrt{n-1}\sqrt{n}\mathbf{E}\left(\overline{G_n}\right)\mathbf{E}\left(\sqrt{\frac{1}{W_n}}\right) \quad \boxed{= 0},$$

puisque les $G_i$ sont centrées.

## *Partie V— Loi de $W_n$ ou loi du khi-deux.*

**17)17.1.** $f_{Y_2}(z) = \displaystyle\int_{-\infty}^{+\infty} f_1(t) f_2(z-t)\,\mathrm{d}t$ pour tout $z \in \mathbf{R}$.

**17.2.** D'après la partie de préliminaires, pour tous $t, z \in \mathbf{R}$,

$$\begin{aligned} f_1(t) &= \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-t^2/2} \mathbb{1}_{]0,\infty[}(t), \\ f_2(z-t) &= \frac{1}{\sqrt{2\pi(z-t)}} \mathrm{e}^{-(z-t)^2/2} \mathbb{1}_{[0,\infty[}(z-t). \end{aligned}$$

Donc :

$$\begin{aligned} f_{Y_2}(z) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-t/2} \mathbb{1}_{]0,\infty[}(t) \frac{1}{\sqrt{2\pi(z-t)}} \mathrm{e}^{-(z-t)/2} \mathbb{1}_{[0,\infty[}(z-t)\,\mathrm{d}t, \\ &= \frac{1}{2\pi} \int_0^{+\infty} \frac{1}{\sqrt{t(z-t)}} \mathrm{e}^{-t/2-(z-t)/2} \mathbb{1}_{[0,\infty[}(z-t)\,\mathrm{d}t \\ &= \boxed{\begin{cases} 0 & \text{si } z < 0, \\ \dfrac{\mathrm{e}^{-z/2}}{2\pi} \displaystyle\int_0^z \frac{1}{\sqrt{t(z-t)}}\,\mathrm{d}t & \text{sinon.} \end{cases}} \end{aligned}$$

**18)** On étudie $\displaystyle\int_0^{\infty} t^{d/2-1}\mathrm{e}^{-t}\,\mathrm{d}t$. L'énoncé précise $d \geqslant 2$, la fonction intégrée est donc continue sur $[0,\infty[$. L'intégrale est donc seulement imporpre en $+\infty$. Or, puisque $t^2 t^{d/2-1}\mathrm{e}^{-t} \underset{t\to\infty}{\longrightarrow} 0$, on déduit que pour $t$ assez grand $\quad t^{d/2-1}\mathrm{e}^{-t} \leqslant \dfrac{1}{t^2}$.

— L'intégrale $\displaystyle\int_0^1 t^{d/2-1}\mathrm{e}^{-t}\,\mathrm{d}t$ converge puisque la fonction intégrée est continue sur le segment $[0,1]$.

— L'intégrale $\int_1^{\infty} t^{d/2-1}\mathrm{e}^{-t}\,\mathrm{d}t$ converge aussi par comparaison puisque les fonctions $t \longmapsto t^{d/2-1}\mathrm{e}^{-t}$ et $t \longmapsto \frac{1}{t^2}$ sont positives, et que $\int_1^{\infty}\frac{1}{t^2}\,\mathrm{d}t$ converge.

En conclusion $\boxed{\text{l'intégrale converge}}$.

## *Partie VI— Intervalle de confiance et test.*

**19) 19.1.** La densité semble paire.

**19.2.** Ainsi $\mathbf{P}\left(|T_{10}|<u\right) = \mathbf{P}\left(-u<T_{10}<u\right) = F_{10}(u) - \mathbf{P}\left(T_{10}\leqslant -u\right)$. Or,

$$\mathbf{P}\left(T_{10}\leqslant -u\right) = \int_{-\infty}^{-u} f_{10}(x)\,\mathrm{d}x = \int_{\infty}^{u} f_{10}(-x)(-\,\mathrm{d}x) = \int_{u}^{\infty} f_{10}(x)\,\mathrm{d}x = 1-\mathbf{P}\left(T_{10}\leqslant u\right) = 1-F_{10}(u).$$

À la dernière étape nous avons réalisé le changement de variable $x \leftarrow -x$ et utilisé la parité de la densité. On obtient

$$\mathbf{P}\left(|T_{10}|<u\right) = \boxed{2F_{10}(u)-1.}$$

**20) 20.1.** D'après les calculs précédents, nous avons par lecture graphique :

$$\mathbf{P}\left(|T_{10}|<u\right) = 0{,}95 \iff F_{10}(u) = 0{,}975 \iff \boxed{u = 2{,}26}.$$

**20.2.** De-même on obtient :

$$\mathbf{P}\left(|T_{10}|<u\right) = 0{,}99 \iff F_{10}(u) = 0{,}995 \iff \boxed{u = 3{,}25}.$$

**21)** Par manipulation d'encadrement dans l'évènement $|T_{10}|<u$, on obtient le même type d'intervalle de confiance que dans le cours, mais avec une variance corrigée au lieu d'une variance :

— au niveau 0,95 : $I_1 = \left[\overline{X_{10}} - 2{,}262.\frac{{\sigma_{10}}^{\mathbf{cor}}}{\sqrt{10}} ; \overline{X_{10}} + 2{,}262.\frac{{\sigma_{10}}^{\mathbf{cor}}}{\sqrt{10}}\right]$,

— au niveau 0,99 : $I_2 = \left[\overline{X_{10}} - 3{,}25.\frac{{\sigma_{10}}^{\mathbf{cor}}}{\sqrt{10}} ; \overline{X_{10}} + 3{,}25.\frac{{\sigma_{10}}^{\mathbf{cor}}}{\sqrt{10}}\right]$.

**22)** Pour un 10-échantillon, on peut considérer la prise de décision suivante :

❶ Si $\mu_0 \in I_1$ (ou $I_2$), alors on ne rejette pas l'hypothèse « $\mu = \mu_0$ ».

❷ Sinon, on rejette l'hypothèse « $\mu = \mu_0$ ».

**23)** Pour rappel, l'intervalle de confiance de niveau 0,95 provenant du théorème central limite est

$$\boxed{\left[\overline{X_{10}} - \Phi^{-1}(1-\alpha/2)\frac{\sigma_{10}}{\sqrt{10}}, \overline{X_{10}} + \Phi^{-1}(1-\alpha/2)\frac{\sigma_{10}}{\sqrt{10}}\right]}.$$

Pour savoir lequel est le plus précis il s'agit de comparer les quantités $\sigma_{10}1{,}96$ et $\sigma_{10}^{\mathbf{cor}}2{,}26 = \frac{9}{10}\sigma_{10}2{,}26$. Or, $9/10 * 2.26 = 2.034 > 1{,}96$ donc l'intervalle donné par le théorème central limite est plus précis, mais il est asymptotique. Donc pour des très petites valeurs ($n = 10$ par exemple), il faut plutôt privilégier celui avec la variance corrigée.