MATHÉMATIQUES APPROFONDIES - Maths 2 HEC-ESCP 2025

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major Prépa

L'énoncé supposait, pour toutes les questions en langage `Python`*, les bibliothèques usuelles déjà importées sous leurs raccourcis habituels.*

```python
import numpy as np
import numpy.random as rd
import matplotlib.pyplot as plt
```

Première partie

Soit $(X_i)_{i\in\mathbb{N}^*}$ une famille de variables aléatoires mutuellement indépendantes, chacune de loi normale $\mathcal{N}(0,1)$.

L'énoncé supposait que $X_i(\omega) \neq 0$ pour tout $i \in \mathbb{N}^*$ et tout $\omega \in \Omega$, et définit, pour tout entier $n \in \mathbb{N}^*$, $S_n = \displaystyle\sum_{i=1}^{n} X_i^2$.

La loi de S_n est appelée **loi du khi-deux** à n degrés de liberté, et on note $S_n \hookrightarrow \chi^2(n)$.

On pose aussi, pour tout entier $n \in \mathbb{N}^*$, $W_n = \dfrac{1}{2} S_n$.

1. (a) D'après le cours sur la loi normale, chaque variable aléatoire X_i ($i \in \mathbb{N}^*$) admet une espérance qui vaut 0 et une variance qui vaut 1, donc un moment d'ordre 2 : $\mathbf{E}(X_i^2) = 1$ aussi (conséquence de la formule de Koenig-Huygens).

   La variable aléatoire S_n admet donc une espérance comme somme finie de variables qui en admettent une, et par linéarité de l'espérance :

$$\mathbf{E}(S_n) = \sum_{i=1}^{n} \mathbf{E}(X_i^2) = n.$$

   L'énoncé admettait ici que S_n possède un moment d'ordre 2.

   (b) On simule facilement S_n comme somme des carrés de n simulations indépendantes de la loi normale centrée, réduite. L'annexe `Python` du sujet rappelle opportunément la commande adaptée.

```python
def simul(n):
    X = rd.normal(0,1,n)
    return np.sum(X**2)
```

   Rappelons ici que la puissance 2 appliquée au vecteur `X` s'entend terme à terme à chaque élément du vecteur. On pouvait bien sûr construire cette somme avec une boucle `for`.

   (c) Le script fourni par l'énoncé calcule des réalisations de $\overline{S_n^2} = \dfrac{1}{N}\displaystyle\sum_{k=0}^{N-1} (S_n^2)_k$, moyenne empirique que $N = 50000$ simulations indépendantes de la loi du khi-deux à n degrés de libertés, et renvoie via la fonction f, la valeur de $\overline{S_n^2} - n^2 = \overline{S_n^2} - \mathbf{E}(S_n)^2$.

   En considérant que S_n^2 admet un moment d'ordre 2, ce qui revient à dire que S_n admet un moment d'ordre 4 (ce qui est vrai car c'est la somme de n variables indépendantes de loi normale qui admettent des moments de tous ordres) : alors $\overline{S_n^2}$ est un estimateur sans biais et convergent de $\mathbf{E}(S_n^2)$, et le graphique affiche ici des estimations de $\mathbf{E}(S_n^2) - \mathbf{E}(S_n)^2 = \mathbf{V}(S_n)$.

Il semble donc, au vu des valeurs prises, que l'on puisse conjecturer : $\boxed{\mathbf{V}(S_n) = 2n}$.

2. (a) La variable aléatoire $W_1 = \frac{1}{2}S_1 = \frac{1}{2}X_1^2$ est à valeurs positives, donc sa fonction de répartition F_{W_1} vérifie déjà : $\forall x < 0,\ F_{W_1}(x) = 0$.

Pour tout $x \geqslant 0$:

$$F_{W_1}(x) = \mathbf{P}(W_1 \leqslant x) = \mathbf{P}(X_1^2 \leqslant 2x) \overset{(*)}{=} \mathbf{P}(|X_1| \leqslant \sqrt{2x})$$

$$= \mathbf{P}(-\sqrt{2x} \leqslant X_1 \leqslant \sqrt{2x}) \overset{(**)}{=} \Phi(\sqrt{2x}) - \Phi(-\sqrt{2x}) \overset{(***)}{=} 2\Phi(\sqrt{2x}) - 1$$

$(*)$ car la fonction carrée est continue strictement croissante, bijective de $\mathbb{R}^+$ dans lui-même.
$(**)$ Car $X_1 \hookrightarrow \mathcal{N}(0,1)$ est à densité.
$(***)$ d'après une propriété fondamentale de Φ : $\forall x \in \mathbb{R},\ \Phi(-x) = 1 - \Phi(x)$.

La fonction $F_{W_1} :\ x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 2\Phi(\sqrt{2x}) - 1 & \text{si } x \geqslant 0 \end{cases}$ est alors de classe $\mathcal{C}^1$, donc continue, sur $]-\infty\,;0[$ comme fonction constante sur cet intervalle, et sur $]0\,;+\infty[$ car c'est le cas de la composée de $x \mapsto \sqrt{2x} \in \mathcal{C}^1(\mathbb{R}_+^*, \mathbb{R}_+^*)$ et de $\Phi \in \mathcal{C}^1(\mathbb{R}),]0\,;1[)$.

De plus, les fonctions $\sqrt{\cdot}$ et Φ étant continues en 0 :

$$\lim \lim_{x \to 0^+} 2\Phi(\sqrt{2x}) - 1 = 2\Phi(0) - 1 = 2 \cdot \tfrac{1}{2} - 1 = 0 = F_{W_1}(0) = \lim_{x \to 0^-} F_{W_1}(x),$$

donc F_{W_1} est continue en 0. Ainsi, F_{W_1} est continue sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf peut-être en 0, donc en un nombre fini de points : W_1 est bien une variable à densité, et une densité f_{W_1} de W_1 est obtenue par dérivation de F_{W_1}, sauf en 0 où on choisit une valeur arbitraire positive.

La fonction f_{W_1} définie par :

$$\forall x \leqslant 0,\ f_{W_1}(x) = 0 \text{ et } \forall x > 0,\ f_{W_1}(x) = 2 \cdot \frac{2}{2\sqrt{2x}}\varphi(\sqrt{2x}) = \sqrt{\frac{2}{x}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{2x})^2}{2}} = \frac{1}{\sqrt{\pi}} \cdot \frac{e^{-x}}{\sqrt{x}},$$

est une densité de W_1 (on a choisi arbitrairement $f_{W_1}(0) = 0$).

(b) La densité de W_1 précédente vérifie :

$$\int_{-\infty}^{+\infty} f_{W_1}(x)\mathrm{d}x = 1 \iff \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \frac{1}{\sqrt{x}} e^{-x}\mathrm{d}x = 1.$$

Or pour tout $x > 0$, $\dfrac{1}{\sqrt{x}} = x^{-\frac{1}{2}} = x^{\frac{1}{2}-1}$, donc la dernière égalité ci-dessus se réécrit :

$$\int_0^{+\infty} x^{\frac{1}{2}-1} e^{-x}\mathrm{d}x = \sqrt{\pi} \iff \boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}},$$

l'intégrale étant bien convergente puisque $\frac{1}{2} > 0$.

(c) Le résultat précédent implique surtout, que $W_1 = \frac{1}{2}X_1^2$ suit en fait la loi $\gamma\left(\dfrac{1}{2}\right)$.

En repérant alors que $W_n = \frac{1}{2}S_n = \frac{1}{2}X_1^2 + \cdots + \frac{1}{2}X_n^2$: le lemme des coalitions assure que l'indépendance mutuelle des $(X_i)_{i \in \mathbb{N}^*}$ entraîne celle des $\left(\frac{1}{2}X_i^2\right)_{i \in \mathbb{N}^*}$, et chaque variable $\frac{1}{2}X_i^2$ suit la même loi $\gamma\left(\dfrac{1}{2}\right)$ que $W_1 = \frac{1}{2}X_1^2$: le théorème de stabilité de la loi gamma par somme indépendantes, assure alors que W_n suit la loi gamma dont le paramètre est la somme des paramètres des $\frac{1}{2}X_i^2$, c'est-à-dire que $W_n \hookrightarrow \gamma\left(\dfrac{n}{2}\right)$.

(d) Le cours sur la loi gamma donne alors : $\mathbf{E}(W_n) = \dfrac{n}{2} \iff \dfrac{1}{2}\mathbf{E}(S_n) = \dfrac{n}{2} \iff \mathbf{E}(S_n) = n$, et aussi :

$$\mathbf{V}(W_n) = \frac{n}{2} \iff \frac{1}{4}\mathbf{V}(S_n) = \frac{n}{2} \iff \boxed{\mathbf{V}(S_n) = 2n}$$

ce qui correspond bien à la conjecture qu'on avait faite, suite à l'observation des résultats du script `Python` de la question 1.

3. (a) Soit un entier $n \geqslant 3$. La fonction inverse est continue sur $]0\,;+\infty[$ qui contient presque-sûrement les valeurs prises par W_n, donc d'après le théorème de transfert, $\mathbf{E}\Big(\dfrac{1}{W_n}\Big)$ existe si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{x} f_{W_n}(x)\mathrm{d}x$ est absolument convergente, où $f_{W_n} : x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{1}{\Gamma(n/2)} x^{\frac{n}{2}-1}e^{-x} & \text{si } x > 0 \end{cases}$ est une densité de $W_n \hookrightarrow \gamma(n/2)$.

Comme la fonction $x \mapsto \dfrac{1}{x} f_{W_n}(x)$ est nulle sur $]-\infty\,;0[$ et positive sur $]0\,;+\infty[$, cela revient à étudier la convergence simple de :

$$\frac{1}{\Gamma(n/2)} \int_0^{+\infty} \frac{1}{x} x^{\frac{n}{2}-1}e^{-x}\mathrm{d}x = \frac{1}{\Gamma(n/2)} \int_0^{+\infty} x^{(\frac{n}{2}-1)-1}e^{-x}\mathrm{d}x.$$

On reconnaît l'intégrale $\Gamma\big(\frac{n}{2}-1\big)$, bien convergente puisque $n \geqslant 3 \Longrightarrow \frac{n}{2} - 1 \geqslant \frac{1}{2} > 0$.

Ainsi $\dfrac{1}{W_n}$ admet bien une espérance, qui vaut effectivement :

$$\mathbf{E}\Big(\frac{1}{W_n}\Big) = \frac{\Gamma\big(\frac{n}{2}-1\big)}{\Gamma\big(\frac{n}{2}\big)}.$$

Puisque $S_n = 2W_n$, alors $\dfrac{1}{S_n} = \dfrac{1}{2}\cdot\dfrac{1}{W_n}$ admet bien une espérance, et par linéarité de cette dernière :

$$\mathbf{E}\Big(\frac{1}{S_n}\Big) = \frac{1}{2}\cdot\mathbf{E}\Big(\frac{1}{W_n}\Big) = \frac{\Gamma\big(\frac{n}{2}-1\big)}{2\Gamma\big(\frac{n}{2}\big)} = \frac{\Gamma\big(\frac{n}{2}-1\big)}{2\big(\frac{n}{2}-1\big)\Gamma\big(\frac{n}{2}-1\big)} = \frac{1}{n-2}$$

d'après la relation fondamentale : $\Gamma(x+1) = x\Gamma(x)$ pour tout $x > 0$ rappelée par l'énoncé, ici utilisée avec $\frac{n}{2} - 1 > 0$ ($n \geqslant 3$).

(b) Soit $n \geqslant 3$: la variable aléatoire $\dfrac{1}{\sqrt{S_n}}$ est presque sûrement bien définie, et le résultat précédent affirme que $\dfrac{1}{S_n} = \Big(\dfrac{1}{\sqrt{S_n}}\Big)^2$ admet une espérance.

Cela signifie que $\dfrac{1}{\sqrt{S_n}}$ admet un moment d'ordre 2 : on sait que cela implique que cette variable aléatoire admet aussi un moment d'ordre 1, c'est-à-dire une espérance.

Soit Y une variable aléatoire indépendante des $(X_k)_{k\geqslant 1}$ et de loi $\mathcal{N}(0,1)$. Pour tout $n \geqslant 1$, on pose :

$$T_n = \frac{Y}{\sqrt{S_n/n}}.$$

La loi de T_n est appelée **loi de** STUDENT de paramètre n, et on note $T_n \hookrightarrow \mathcal{T}(n)$.
L'énoncé admettait que T_n est une variable aléatoire à densité possédant une densité strictement positive et continue sur $\mathbb{R}$.

4. Soit $\alpha \in]0\,;1[$. La variable aléatoire $|T_n|$ est à valeurs positives, et pour tout $x \geqslant 0$,

$$F_{|T_n|}(x) = \mathbf{P}(|T_n| \leqslant x) = \mathbf{P}(-x \leqslant T_n \leqslant x) = F_{T_n}(x) - F_{T_n}(-x).$$

Comme T_n admet une densité continue sur $\mathbb{R}$, alors sa fonction de répartition est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc $F_{|T_n|}$ aussi d'après l'expression précédente, et en notant f_{T_n} la densité de T_n évoquée par l'énoncé, on a :

$$\forall x \geqslant 0, \quad F'_{|T_n|}(x) = F'_{T_n}(x) + F'_{T_n}(-x) = f_{T_n}(x) + f_{T_n}(-x) > 0$$

puisque f_{T_n} est strictement positive sur $\mathbb{R}$.

La fonction $F_{|T_n|}$ est donc continue, strictement croissante sur $[0\,;+\infty[$.

Comme $F_{|T_n|}(0) = \mathbf{P}(|T_n| \leqslant 0) = \mathbf{P}(T_n = 0) = 0$ car T_n est à densité, et $\lim\limits_{x\to+\infty} F_{|T_n|}(x) = 1$ (propriété d'une fonction de répartition), alors d'après le théorème éponyme, $F_{|T_n|}$ réalise une bijection de $[0\,;+\infty[$ dans $[0\,;1[$.

Puisque $1-\alpha \in]0\,;1[$, il existe un unique réel $t_{n,\alpha} \in [0\,;+\infty[$ tel que $\mathbf{P}(|T_n| \leqslant t_{n,\alpha}) = 1-\alpha$.

C'est aussi l'unique solution sur $\mathbb{R}$ puisque sur $]-\infty\,;0[$, $F_{|T_n|}$ est nulle, donc ne peut pas prendre la valeur $1-\alpha$.

5. (a) Soit un entier $n \geqslant 3$. On a vu en 3.(b) que $\dfrac{1}{\sqrt{S_n}}$ admet une espérance ; comme Y admet aussi une espérance (qui vaut 0), et puisque Y est indépendante des $(X_k)_{k\geqslant 1}$, donc de $\dfrac{1}{\sqrt{S_n}}$ d'après le lemme des coalitions, alors $T_n = \sqrt{n}\cdot Y \times \dfrac{1}{\sqrt{S_n}}$ admet une espérance qui vaut :

$$\mathbf{E}(T_n) = \sqrt{n}\cdot \underbrace{\mathbf{E}(Y)}_{=0} \times \mathbf{E}\Big(\frac{1}{\sqrt{S_n}}\Big) = 0.$$

(b) Selon le même argument : $T_n^2 = n\cdot Y^2 \times \dfrac{1}{S_n}$ admet une espérance comme produit d'un facteur constant et de deux variables aléatoires indépendantes qui admettent une espérance.

Cela signifie que T_n admet un moment d'ordre 2, qui vaut :

$$\mathbf{E}(T_n^2) = n\cdot \mathbf{E}(Y^2) \times \mathbf{E}\Big(\frac{1}{S_n}\Big) = \frac{n}{n-2}$$

puisque, comme on l'a vu au début de cette partie, $\mathbf{E}(Y^2) = \mathbf{V}(Y) = 1$.

Ce moment d'ordre 2 est aussi, d'après la formule de Koenig-Huygens, la variance de T_n puisque son espérance est nulle.

(c) Soit $n \geqslant 3$: $(T_n - Y)^2 = T_n^2 - 2YT_n + Y^2 = T_n^2 - 2\sqrt{n}\dfrac{Y^2}{\sqrt{S_n}} + Y^2$ est, d'après ce qui précède, la somme de trois variables aléatoires qui admettent une espérance, donc $\mathbf{E}\big((T_n-Y)^2\big)$ existe et vaut :

$$\begin{aligned}\mathbf{E}\big((T_n-Y)^2\big) &= \mathbf{E}(T_n^2) - 2\sqrt{n}\mathbf{E}(Y^2)\cdot\mathbf{E}\Big(\frac{1}{\sqrt{S_n}}\Big) + \mathbf{E}(Y^2) = \frac{n}{n-2} - 2\sqrt{n}\cdot\mathbf{E}\Big(\frac{1}{\sqrt{2W_n}}\Big) + 1\\ &= \frac{2n-2}{n-2} - \sqrt{2n}\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big)\end{aligned}$$

en sortant un facteur $\dfrac{1}{\sqrt{2}}$ de la dernière espérance, et en réduisant au même dénominateur la somme $\dfrac{n}{n-2} + 1 = \dfrac{2n-2}{n-2}$.

6. Pour tout entier $n \geqslant 2$, on pose : $\quad u_n = \mathbf{E}\Big(\dfrac{1}{\sqrt{W_n}}\Big)$.

(a) En remarquant que pour tout entier $n \geqslant 2$:

$$\forall \omega \in \Omega, \quad W_{n+1}(\omega) = \frac{1}{2}S_{n+1}(\omega) = W_n(\omega) + \frac{1}{2}X_{n+1}^2(\omega) \geqslant W_n(\omega) > 0$$

puisque l'énoncé suppose que $X_i(\omega) \neq 0 \Longrightarrow X_i^2(\omega) > 0$; la stricte croissance de $\sqrt{\cdot}$ sur $\mathbb{R}_+^*$ et la stricte décroissance de la fonction inverse sur cet intervalle donnent :

$$\forall \omega \in \Omega, \quad \sqrt{W_{n+1}(\omega)} \geqslant \sqrt{W_n(\omega)} > 0 \Longrightarrow \frac{1}{\sqrt{W_{n+1}(\omega)}} \leqslant \frac{1}{\sqrt{W_n(\omega)}},$$

ce qui signifie l'inégalité entre variables aléatoires : $\quad \forall n \geqslant 2, \ \dfrac{1}{\sqrt{W_{n+1}}} \leqslant \dfrac{1}{\sqrt{W_n}}$.

La propriété de *croissance* de l'espérance donne alors (les espérances étant bien définies) :

$$\forall n \geqslant 2, \quad \mathbf{E}\Big(\frac{1}{\sqrt{W_{n+1}}}\Big) \leqslant \mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big) \iff \forall n \geqslant 2, \quad u_{n+1} \leqslant u_n,$$

donc la suite $(u_n)_{n\geqslant 2}$ est bien décroissante.

Enfin, la variable W_2 suit la loi $\gamma(1)$ qui est aussi la loi exponentielle de paramètre 1.

D'après 3.(b), $\mathbf{E}\Big(\dfrac{1}{\sqrt{W_2}}\Big)$ existe et vaut, d'après le théorème de transfert :

$$\int_0^{+\infty} \frac{1}{\sqrt{x}}e^{-x}\mathrm{d}x = \int_0^{+\infty} x^{-\frac{1}{2}}e^{-x}\mathrm{d}x = \int_0^{+\infty} x^{\frac{1}{2}-1}e^{-x}\mathrm{d}x = \Gamma\Big(\frac{1}{2}\Big) = \sqrt{\pi} = u_2$$

d'après 2.(b).

(b) Soit un entier $n \geqslant 2$. De la même façon qu'à la question précédente, puisque $W_n \hookrightarrow \gamma\Big(\dfrac{n}{2}\Big)$ est (presque-sûrement) à valeurs dans $\mathbb{R}_+^*$ sur lequel $x \mapsto \dfrac{1}{\sqrt{x}}$ est continue et positive : d'après le théorème de transfert,

$$u_n = \mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big) = \frac{1}{\Gamma\big(\frac{n}{2}\big)}\int_0^{+\infty} \frac{1}{\sqrt{x}}x^{\frac{n}{2}-1}e^{-x}\mathrm{d}x = \frac{1}{\Gamma\big(\frac{n}{2}\big)}\int_0^{+\infty} x^{\frac{n-1}{2}-1}e^{-x}\mathrm{d}x = \frac{\Gamma\big(\frac{n}{2}-1\big)}{\Gamma\big(\frac{n}{2}\big)}.$$

Ainsi : $\quad u_{n+1}\cdot u_n = \dfrac{\Gamma\big(\frac{n}{2}\big) \times \Gamma\big(\frac{n-1}{2}\big)}{\Gamma\big(\frac{n+1}{2}\big) \times \Gamma\big(\frac{n}{2}\big)} = \dfrac{\Gamma\big(\frac{n-1}{2}\big)}{\Gamma\big(\frac{n+1}{2}\big)}$.

Et comme $\Gamma\big(\frac{n+1}{2}\big) = \Gamma\big(\frac{n-1}{2}+1\big) = \frac{n-1}{2}\Gamma\big(\frac{n-1}{2}\big)$, alors $u_{n+1}\cdot u_n = \dfrac{1}{\frac{n-1}{2}} = \dfrac{2}{n-1}$.

(c) La fonction `Python` ci-dessous prend en argument l'entier $n \geqslant 2$, et renvoie sous la forme d'un tableau `numpy` les $n-1$ valeurs $u_2, u_3, \ldots, u_n$.

On tiendra compte du fait que la numérotation des éléments dans un tableau commence à l'indice 0 : u_2 sera l'élément numéro 0, etc..., il y a donc ci-dessous un décalage de deux unités entre l'indice du terme u_k calculé, et son numéro de place dans le vecteur.

```python
def suite_u(n):
    u = np.zeros(n-1)
    u[0] = np.sqrt(np.pi)
    for k in range(1,n-1):
        u[k] = 2/(k*u[k-1])
    return u
```

Pour tout $k \in [\![1\,;n-2]\!]$:

$$u[k] = u_{k+2} \overset{6.(b)}{=} \frac{2}{k \cdot u_{k+1}} = \frac{2}{k \cdot \texttt{u[k-1]}}.$$

(d) Le programme de l'énoncé calcule alors les 79 premiers termes $u_2, \ldots, u_{80}$ de la suite $(u_n)_{n\geqslant 2}$, les entiers $k \in [\![2\,;80]\!]$ dans le vecteur `V`, et affiche les points de coordonnées $(k, ku_k^2)_{2\leqslant k\leqslant 80}$, donc les 79 permiers termes de la suite $(n \cdot u_n^2)_{n\geqslant 2}$.

Le comportement du nuage de points permet de conjecturer :

$$\lim_{n\to+\infty} n \cdot u_n^2 = 2,$$

ce qui s'écrit aussi : $\displaystyle\lim_{n\to+\infty} \frac{n}{2}u_n^2 = 1$, dont on déduit l'équivalent

$$u_n^2 \underset{n\to+\infty}{\sim} \frac{2}{n} \iff u_n \underset{n\to+\infty}{\sim} \sqrt{\frac{2}{n}}$$

puisque $u_n > 0$.

(e) On peut reprendre ici un schéma de raisonnement classique (utilisé notamment pour l'étude des intégrales de Wallis) : puisque la suite $(u_n)_{n\geqslant 2}$ est décroissante, alors pour tout entier $n \geqslant 3$, on a

$$u_{n+1} \leqslant u_n \leqslant u_{n-1} \overset{\times u_n \geqslant 0}{\Longrightarrow} u_{n+1}{\cdot}u_n \leqslant u_n^2 \leqslant u_n{\cdot}u_{n-1} \Longrightarrow \frac{2}{n-1} \leqslant u_n^2 \leqslant \frac{2}{n-2} \overset{\times n \geqslant 0}{\Longrightarrow} \frac{2}{1-\frac{1}{n}} \leqslant n{\cdot}u_n^2 \leqslant \frac{2}{1-\frac{2}{n}}.$$

Puisque $\displaystyle\lim_{n\to+\infty} \frac{2}{1-\frac{1}{n}} = 2 = \frac{2}{1-\frac{2}{n}}$, le théorème d'encadrement permet de conclure :

$$\lim_{n\to+\infty} n \cdot u_n^2 = 2 \Longrightarrow u_n \underset{n\to+\infty}{\sim} \sqrt{\frac{2}{n}}$$

en reprenant les étapes du raisonnement mené à la question précédente.

7. Soit $\varepsilon > 0$ fixé : on cherche à montrer que $\mathbf{P}\big(|T_n - Y| > \varepsilon\big)$ tend vers 0 lorsque n tend vers $+\infty$.
Pour cela, on remarque que $\mathbf{P}\big(|T_n - Y| > \varepsilon\big) = \mathbf{P}\big((T_n - Y)^2 > \varepsilon^2\big)$ car la fonction carrée est continue strictement croissante, bijective de $\mathbb{R}^+$ dans lui-même.
La variable aléatoire $(T_n - Y)^2$ étant à valeurs positives et admettant une espérance, on peut écrire pour elle l'**inégalité de Markov**, qui donne :

$$0 \leqslant \mathbf{P}\big((T_n - Y)^2 > \varepsilon^2\big) \leqslant \frac{\mathbf{E}\big((T_n - Y)^2\big)}{\varepsilon^2} = \frac{1}{\varepsilon^2}\left(\frac{2n-2}{n-2} - \sqrt{2n}\,\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big)\right)$$

d'après 5.(c).

Or : $2n-2 \underset{n\to+\infty}{\sim} 2n$ et $n-2 \underset{n\to+\infty}{\sim} n$, donc $\displaystyle\frac{2n-2}{n-2} \underset{n\to+\infty}{\sim} 2$, ce qui implique que $\displaystyle\lim_{n\to+\infty} \frac{2n-2}{n-2} = 2$.

Par ailleurs : $\displaystyle\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big) \underset{n\to+\infty}{\sim} \sqrt{\frac{2}{n}} \Longrightarrow \sqrt{2n}\,\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big) \underset{n\to+\infty}{\sim} 2$, donc $\displaystyle\lim_{n\to+\infty} \sqrt{2n}\,\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big) = 2$,

mais alors $\displaystyle\lim_{n\to+\infty} \frac{1}{\varepsilon^2}\left(\frac{2n-2}{n-2} - \sqrt{2n}\,\mathbf{E}\Big(\frac{1}{\sqrt{W_n}}\Big)\right) = \frac{2-2}{\varepsilon^2} = 0.$

Le théorème d'encadrement assure pour finir que :

$$\lim_{n\to+\infty} \mathbf{P}\big((T_n - Y)^2 > \varepsilon^2\big) = 0 \iff \lim_{n\to+\infty} \mathbf{P}\big(|T_n - Y| > \varepsilon\big) = 0,$$

et ce pour tout $\varepsilon > 0$: on a bien démontré que la suite $(T_n)_{n\geqslant 2}$ converge en probabilité vers Y.

Deuxième partie

Dans cette partie, on démontre le résultat suivant :

Théorème : Soient X, Y deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0,1)$.

Les variables aléatoires $\dfrac{X+Y}{\sqrt{2}}$ et $\dfrac{X-Y}{\sqrt{2}}$ sont alors indépendantes et de loi normale $\mathcal{N}(0,1)$.

Pour la preuve, l'énoncé fait appel au résultat suivant, à utiliser sans preuve :

Théorème : Soient X, Y deux variables aléatoires indépendantes, de densités respectives f et g. Soit $\mathcal{A} \subset \mathbb{R}^2$ une partie fermée. Alors :

$$\mathbf{P}\big((X,Y) \in \mathcal{A}\big) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \mathbb{1}_{\mathcal{A}}(x,y) f(x) \mathrm{d}x \right) g(y) \mathrm{d}y,$$

où $\mathbb{1}_{\mathcal{A}}$ est la fonction indicatrice de la partie $\mathcal{A}$.

Soient a, b deux réels fixés jusqu'à la fin de cette partie. On considère la partie $\mathcal{A} \subset \mathbb{R}^2$ définie par :

$$\mathcal{A} = \big\{(x,y) \in \mathbb{R}^2 \big|\ x+y \leqslant a \text{ et } x-y \leqslant b\big\}.$$

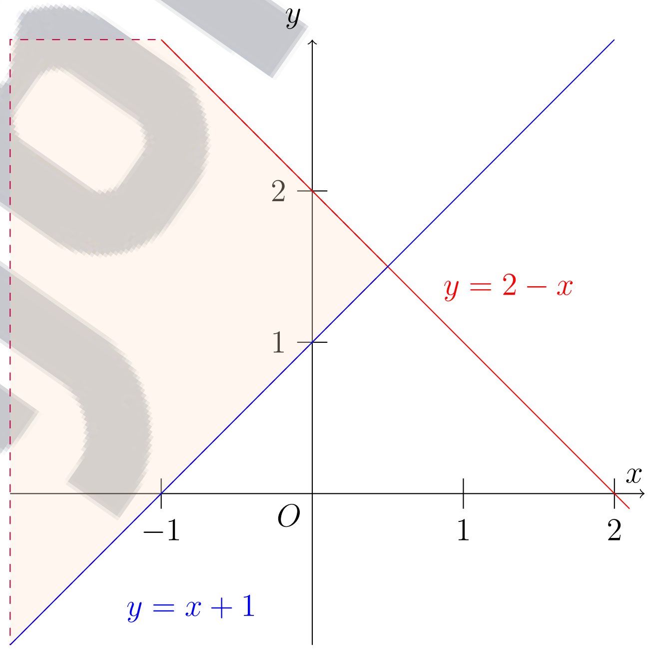
On pose également $c = \dfrac{a+b}{2}$, $d = \dfrac{a-b}{2}$.

Soient X et Y deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0,1)$.

On pose $W = \dfrac{X+Y}{\sqrt{2}}$ et $Z = \dfrac{X-Y}{\sqrt{2}}$.

8. Lorsque $a = 2$ et $b = -1$: $\mathcal{A} = \big\{(x,y) \in \mathbb{R}^2 \big|\ y \leqslant 2-x \text{ et } y \geqslant x+1\big\}$ est l'ensemble des points du plan $\mathbb{R}^2$ qui sont situés au-dessous de la droite D_1 d'équation $y = 2-x$, et au-dessus de la droite D_2 d'équation $y = x+1$.

   L'ensemble $\mathcal{A}$ est la partie en couleur ci-dessous (les pointillés indiquent que la partie s'étend à l'infini au-delà du cadre du dessin).



9. Soit y un réel tel que $y \geqslant d$.

   (a) Comme $x+y \leqslant a \iff x \leqslant a-y$ et puisque $x \leqslant a-y \Longrightarrow x-y \leqslant a-2y \Longrightarrow x-y \leqslant a-2d = b$, alors on a bien $(x,y) \in \mathcal{A}$ si et seulement si $x \in]-\infty\,; a-y]$.

(b) On a alors : $\mathbb{1}_{\mathcal{A}}(x,y)=1 \iff (x,y)\in\mathcal{A} \iff x\in]-\infty\,;a-y]$, sinon l'indicatrice vaut 0, donc :

$$\int_{-\infty}^{+\infty}\mathbb{1}_{\mathcal{A}}(x,y)\varphi(x)\mathrm{d}x=\int_{-\infty}^{a-y}\varphi(x)\mathrm{d}x=\Phi(a-y)$$

par définition de la fonction de répartition Φ de la loi $\mathcal{N}(0,1)$.

10. Si $y\leqslant d$, alors $x-y\leqslant b \iff x\leqslant y+b$, et puisque $x\leqslant b+y \Longrightarrow x+y\leqslant b+2y\leqslant b+2d=a$, alors $(x,y)\in\mathcal{A}$ si et seulement si $x\in]-\infty\,;b+y]$, de sorte que :

$$\int_{-\infty}^{+\infty}\mathbb{1}_{\mathcal{A}}(x,y)\varphi(x)\mathrm{d}x=\int_{-\infty}^{b+y}\varphi(x)\mathrm{d}x=\Phi(b+y).$$

11. a) Les fonctions $(x,y)\mapsto x+y$ et $(x,y)\mapsto x-y$ sont polynômiales en les variables x et y, donc continues sur $\mathbb{R}^2$.

Les deux ensembles $\{(x,y)\in\mathbb{R}^2 \mid x+y\leqslant a\}$ et $\{(x,y)\in\mathbb{R}^2 \mid x-y\leqslant b\}$ sont donc des fermés élémentaires de $\mathbb{R}^2$, et $\mathcal{A}$ est alors une partie fermée de $\mathbb{R}^2$ comme intersection de ces deux fermés.

b) D'après le théorème admis au début de cette partie, puisque $\mathcal{A}$ est une partie fermée de $\mathbb{R}^2$, et puisque X et Y sont indépendantes, de même densité φ :

$$\begin{aligned}\mathbf{P}\big((X,Y)\in\mathcal{A}\big)&=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty}\mathbb{1}_{\mathcal{A}}(x,y)\varphi(x)\mathrm{d}x\right)\varphi(y)\mathrm{d}y\\&=\int_{-\infty}^{d}\left(\int_{-\infty}^{+\infty}\mathbb{1}_{\mathcal{A}}(x,y)\varphi(x)\mathrm{d}x\right)\varphi(y)\mathrm{d}y+\int_{d}^{+\infty}\left(\int_{-\infty}^{+\infty}\mathbb{1}_{\mathcal{A}}(x,y)\varphi(x)\mathrm{d}x\right)\varphi(y)\mathrm{d}y\\&=\int_{-\infty}^{d}\Phi(b+y)\varphi(y)\mathrm{d}y+\int_{d}^{+\infty}\Phi(a-y)\varphi(y)\mathrm{d}y\end{aligned}$$

D'après la relation de Chasles et les résultats des questions 9. et 10.

Dans la première intégrale, on réalise alors le changement de variable affine $z=d-y$ et dans la seconde, le changement de variable $z=y-d$ et on obtient :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big)=\int_{0}^{+\infty}\Phi(b+d-z)\varphi(d-z)\mathrm{d}z+\int_{0}^{+\infty}\Phi(a-d-z)\varphi(z+d)\mathrm{d}z.$$

En remarquant que $b+d=b+\frac{a-b}{2}=\frac{a+b}{2}=c$ et $a-d=a-\frac{a-b}{2}=\frac{a+b}{2}=c$ aussi, la linéarité de l'intégrale et le facteur commun $\Phi(b+d-z)=\Phi(a-d-z)=\Phi(c-z)$ donnent :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big)=\int_{0}^{+\infty}\big(\varphi(d+z)+\varphi(d-z)\big)\Phi(c-z)\mathrm{d}z.$$

12. Par définition de la fonction de répartition Φ : pour tout réel z,

$$\Phi(c-z)=\int_{-\infty}^{c-z}\varphi(s)\mathrm{d}s=\int_{-\infty}^{c}\varphi(t-z)\mathrm{d}t$$

avec le changement de variable affine (donc licite) : $t=s+z$.

Le résultat de la question précédente se réécrit donc :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big)=\int_{0}^{+\infty}\big(\varphi(d+z)+\varphi(d-z)\big)\left(\int_{-\infty}^{c}\varphi(t-z)\mathrm{d}t\right)\mathrm{d}z=\int_{0}^{+\infty}\left(\int_{-\infty}^{c}\big(\varphi(d+z)+\varphi(d-z)\big)\varphi(t-z)\mathrm{d}t\right)\mathrm{d}z.$$

L'énoncé admettait pour la suite que l'on peut changer l'ordre d'intégration dans la formule précédente, c'est-à-dire que l'on a :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big)=\int_{-\infty}^{c}\left(\int_{0}^{+\infty}\big(\varphi(d+z)+\varphi(d-z)\big)\varphi(t-z)\mathrm{d}z\right)\mathrm{d}t.$$

13. Puisque pour tout réel x, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, alors pour tous réels u et v : $\varphi(u)\varphi(v) = \dfrac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}$ et

$$\varphi\Big(\frac{u+v}{\sqrt{2}}\Big)\varphi\Big(\frac{u-v}{\sqrt{2}}\Big) = \frac{1}{2\pi}e^{-\frac{(u+v)^2}{4}-\frac{(u-v)^2}{4}} = \frac{1}{2\pi}e^{-\frac{u^2+2uv+v^2+u^2-2uv+v^2}{4}} = \frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}} = \varphi(u)\varphi(v).$$

14. D'après le résultat précédent, pour tout $z \in]-\infty\,; c]$:

$$\varphi(d+z)\varphi(t-z) = \varphi\Big(\frac{d+z+t-z}{\sqrt{2}}\Big)\varphi\Big(\frac{d+z-t+z}{\sqrt{2}}\Big) = \varphi\Big(\frac{d+t}{\sqrt{2}}\Big)\varphi\Big(\frac{2z+d-t}{\sqrt{2}}\Big)$$

et de même :

$$\varphi(d-z)\varphi(t-z) = \varphi\Big(\frac{d-z+t-z}{\sqrt{2}}\Big)\varphi\Big(\frac{d-z-t+z}{\sqrt{2}}\Big) = \varphi\Big(\frac{t+d-2z}{\sqrt{2}}\Big)\varphi\Big(\frac{d-t}{\sqrt{2}}\Big),$$

donc :

$$\begin{aligned}\mathbf{P}\big((X,Y)\in\mathcal{A}\big) &= \int_{-\infty}^{c}\left(\int_0^{+\infty}\varphi(d+z)\varphi(t-z)\mathrm{d}z + \int_0^{+\infty}\varphi(d-z)\varphi(t-z)\mathrm{d}z\right)\mathrm{d}t\\ &= \int_{-\infty}^{c}\left(\varphi\Big(\frac{d+t}{\sqrt{2}}\Big)\int_0^{+\infty}\varphi\Big(\frac{2z+d-t}{\sqrt{2}}\Big)\mathrm{d}z + \varphi\Big(\frac{d-t}{\sqrt{2}}\Big)\int_0^{+\infty}\varphi\Big(\frac{t+d-2z}{\sqrt{2}}\Big)\mathrm{d}z\right)\mathrm{d}t\end{aligned}$$

En réalisant alors le changement de variable affine $u = \frac{2z+d-t}{\sqrt{2}}$ dans la première intégrale "intérieure" et $v = \frac{t+d-2z}{\sqrt{2}}$ dans la seconde, on obtient :

$$\int_0^{+\infty}\varphi\Big(\frac{2z+d-t}{\sqrt{2}}\Big)\mathrm{d}z = \int_{\frac{d-t}{\sqrt{2}}}^{+\infty}\varphi(u)\frac{\mathrm{d}u}{\sqrt{2}} = \frac{1}{\sqrt{2}}\Big(1-\Phi\Big(\frac{d-t}{\sqrt{2}}\Big)\Big) = \frac{1}{\sqrt{2}}\Phi\Big(\frac{t-d}{\sqrt{2}}\Big),$$

et

$$\int_0^{+\infty}\varphi\Big(\frac{t+d-2z}{\sqrt{2}}\Big)\mathrm{d}z = \int_{-\infty}^{\frac{t+d}{\sqrt{2}}}\varphi(v)\frac{\mathrm{d}v}{\sqrt{2}} = \frac{1}{\sqrt{2}}\Phi\Big(\frac{t+d}{\sqrt{2}}\Big),$$

donc en effet :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big) = \frac{1}{\sqrt{2}}\int_{-\infty}^{c}\left(\varphi\Big(\frac{t+d}{\sqrt{2}}\Big)\Phi\Big(\frac{t-d}{\sqrt{2}}\Big) + \varphi\Big(\frac{t-d}{\sqrt{2}}\Big)\Phi\Big(\frac{c-d}{\sqrt{2}}\Big)\right)\mathrm{d}t.$$

15. Il faut ici remarquer que les fonctions $f : t \mapsto \Phi\Big(\dfrac{t+d}{\sqrt{2}}\Big)$ et $g : t \mapsto \Phi\Big(\dfrac{t-d}{\sqrt{2}}\Big)$ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, et que :

$$\forall t\in\mathbb{R},\quad f'(t) = \frac{1}{\sqrt{2}}\varphi\Big(\frac{t+d}{\sqrt{2}}\Big) \quad\text{et}\quad g'(t) = \frac{1}{\sqrt{2}}\varphi\Big(\frac{t-d}{\sqrt{2}}\Big),$$

Donc :

$$\forall t\in\mathbb{R},\quad \frac{1}{\sqrt{2}}\left(\varphi\Big(\frac{t+d}{\sqrt{2}}\Big)\Phi\Big(\frac{t-d}{\sqrt{2}}\Big) + \varphi\Big(\frac{t-d}{\sqrt{2}}\Big)\Phi\Big(\frac{t+d}{\sqrt{2}}\Big)\right) = f'(t)g(t) + f(t)g'(t) = \big[f\times g\big]'(t).$$

Mais alors, pour tout $A < c$:

$$\begin{aligned}\frac{1}{\sqrt{2}}\int_A^c\left(\varphi\Big(\frac{t+d}{\sqrt{2}}\Big)\Phi\Big(\frac{t-d}{\sqrt{2}}\Big) + \varphi\Big(\frac{t-d}{\sqrt{2}}\Big)\Phi\Big(\frac{c-d}{\sqrt{2}}\Big)\right)\mathrm{d}t &= \Big[f(t)g(t)\Big]_A^c\\ &= \Phi\Big(\frac{c+d}{\sqrt{2}}\Big)\Phi\Big(\frac{c-d}{\sqrt{2}}\Big) - \Phi\Big(\frac{A+d}{\sqrt{2}}\Big)\Phi\Big(\frac{A-d}{\sqrt{2}}\Big).\end{aligned}$$

Comme $\displaystyle\lim_{A\to-\infty}\frac{A+d}{\sqrt{2}} = -\infty = \lim_{A\to-\infty}\frac{A-d}{\sqrt{2}}$ et $\displaystyle\lim_{x\to-\infty}\Phi(x) = 0$, alors on a bien :

$$\mathbf{P}\big((X,Y)\in\mathcal{A}\big) = \Phi\Big(\frac{c+d}{\sqrt{2}}\Big)\Phi\Big(\frac{c-d}{\sqrt{2}}\Big).$$

16. Sachant que $c+d=a$ et $c-d=b$, et vue la définition de $\mathcal{A}$, le résultat précédent se réécrit :

$$\forall (a,b)\in\mathbb{R}^2,\quad \mathbf{P}\big([X+Y\leqslant a]\cap[X-Y\leqslant b]\big)=\Phi\Big(\frac{a}{\sqrt2}\Big)\Phi\Big(\frac{b}{\sqrt2}\Big)$$

$$\iff \forall (a',b')\in\mathbb{R}^2,\ \mathbf{P}\Big(\Big[\frac{X+Y}{\sqrt2}\leqslant a'\Big]\cap\Big[\frac{X-Y}{\sqrt2}\leqslant b'\Big]\Big)=\Phi(a')\Phi(b'),$$

où lorsque a et b parcourent $\mathbb{R}$, il en est de même de $a'=\dfrac{a}{\sqrt2}$ et $b'=\dfrac{b}{\sqrt2}$.

Il reste à dire que puisque X et Y sont indépendantes, de même loi normale $\mathcal{N}(0,1)$, alors $W=\dfrac{X+Y}{\sqrt2}$ suit encore une loi normale comme combinaison linéaire de X et Y, dont les paramètres sont son espérance et sa variance :

$$\mathbf{E}\Big(\frac{X+Y}{2}\Big)=\frac{\mathbf{E}(X)+\mathbf{E}(Y)}{\sqrt2}=0\quad\text{et}\quad \mathbf{V}\Big(\frac{X+Y}{\sqrt2}\Big)=\frac{\mathbf{V}(X)+\mathbf{V}(Y)}{2}=\frac{1+1}{2}=1,$$

donc $W=\dfrac{X+Y}{\sqrt2}$ suit encore la loi $\mathcal{N}(0,1)$, et on montre de même que $Z=\dfrac{X-Y}{\sqrt2}$ suit aussi la loi $\mathcal{N}(0,1)$.

Le résultat précédent s'écrit alors :

$$\forall (a',b')\in\mathbb{R}^2,\quad \mathbf{P}\big([W\leqslant a']\cap[Z\leqslant b']\big)=\mathbf{P}(W\leqslant a')\times\mathbf{P}(Z\leqslant b'),$$

ce qui prouve bien que W et Z sont indépendantes.

Troisième partie

Soit $n\in\mathbb{N}$ tel que $n\geqslant 2$. On munit l'espace vectoriel $\mathbb{R}^n$ du produit scalaire euclidien standard, qui sera noté $\langle\cdot,\cdot\rangle$. On considère une base orthonormée $(a_1,\dots,a_n)$ de $\mathbb{R}^n$ telle que $a_1=\dfrac{1}{\sqrt n}(1,\dots,1)$.

17. Soit $x=(x_1,\dots,x_n)\in\mathbb{R}^n$. On pose $\overline{x}=\dfrac{x_1+\cdots+x_n}{n}$.

   (a) D'après les formules dans la base orthonormée $(a_1,a_2,\dots,a_n)$:

$$x=\sum_{k=1}^n\langle x\,,a_k\rangle a_k\Longrightarrow\sum_{k=2}^n\langle x,\,,a_k\rangle a_k=x-\langle x\,,a_1\rangle a_1=x-\Big(\underbrace{\frac1n\sum_{k=1}^n x_k}_{=\overline{x}}\Big)\cdot(1,\dots,1)=(x_1-\overline{x},x_2-\overline{x},\dots,x_n-\overline{x}).$$

   Il aura fallu prendre garde ici au fait que la somme de la formule demandée, commence à l'indice $k=2$ pour faire les liens.

   (b) Toujours du fait que la base $(a_1,a_2,\dots,a_n)$ est orthonormée :

$$\Big|\Big|\sum_{k=2}^n\langle x\,,a_k\rangle a_k\Big|\Big|^2=\sum_{k=2}^n\langle x\,,a_k\rangle^2||a_k||^2=\sum_{k=2}^n\langle x\,,a_k\rangle^2,$$

   aussi égale à $||(x_1-\overline{x},\dots,x_n-\overline{x})||^2=\displaystyle\sum_{i=1}^n(x_i-\overline{x})^2$.

18. (a) Avec les vecteurs $a_1=\frac{1}{\sqrt2}(1,1)$ et $a_2=\frac{1}{\sqrt2}(1,-1)$:
   il est clair que $\langle a_1\,,a_2\rangle=\frac12\big(1-1\big)=0$, $||a_1||=\frac{1}{\sqrt2}||(1,1)||=1=||a_2||$: la famille (a_1,a_2) est donc une base orthonormée de $\mathbb{R}^2$, et de plus $\langle (X,Y)\,,a_1\rangle=\dfrac{X+Y}{\sqrt2}$ et $\langle (X,Y)\,,a_2\rangle=\dfrac{X-Y}{\sqrt2}$.

   Le théorème de Cochran assure donc que $\dfrac{X+Y}{\sqrt2}$ et $\dfrac{X-Y}{\sqrt2}$ sont indépendantes, et suivent chacune la loi $\mathcal{N}(0,1)$.

(b) Soit $(\beta_1,\beta_2)\in\mathbb{R}^2\setminus\{(0,0)\}$. On sait déjà d'après le cours, que si R_1 et R_2 sont indépendantes, alors $\mathrm{Cov}(R_1,R_2)=0$.

Pour prouver la réciproque, on commence par remarquer que, par bilinéarité de la covariance :

$$\begin{aligned}\mathrm{Cov}(R_1,R_2)&=\mathrm{Cov}(X_1+X_2,\beta_1X_1+\beta_2X_2)=\beta_1\mathbf{V}(X_1)+\beta_2\mathrm{Cov}(X_1,X_2)+\beta_1\mathrm{Cov}(X_2,X_1)+\beta_2\mathbf{V}(X_2)\\&=\beta_1+\beta_2\end{aligned}$$

puisque X_1 et X_2 sont indépendantes, donc $\mathrm{Cov}(X_1,X_2)=0$ et puisque $\mathbf{V}(X_1)=1=\mathbf{V}(X_2)$ vu que ces deux variables suivent la loi $\mathcal{N}(0,1)$.

Mais alors : $\quad\mathrm{Cov}(R_1,R_2)=0\iff\beta_1+\beta_2=0\iff\beta_2=-\beta_1$,

ce qui implique : $\quad R_1=X_1+X_2=\sqrt{2}\cdot\dfrac{X_1+X_2}{\sqrt{2}}$ et $R_2=\beta_1(X_1-X_2)=\sqrt{2}\beta_1\cdot\dfrac{X_1-X_2}{\sqrt{2}}$.

Or comme on l'a revu en 18.(a), les variables aléatoires $\dfrac{X_1+X_2}{\sqrt{2}}$ et $\dfrac{X_1-X_2}{\sqrt{2}}$ sont indépendantes d'après le théorème de Cochran ; le lemme des coalitions assure alors que R_1 et R_2 sont encore indépendantes.

On a donc prouvé l'équivalence demandée par double implication.

19. On introduit les variables aléatoires :

$$\overline{X}=\frac{X_1+\cdots+X_n}{n},\quad U=\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

(a) La variable aléatoire $\overline{X}$ est une combinaison linéaire de variables aléatoires mutuellement indépendantes, suivant des lois normales : le théorème de stabilité pour cette loi, assure que $\overline{X}$ suit encore une loi normale, donc les paramètres sont son espérance et sa variance. Or :

$$\mathbf{E}(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}\underbrace{\mathbf{E}(X_i)}_{=0}\quad\text{et}\quad\mathbf{V}(\overline{X})=\frac{1}{n^2}\sum_{i=1}^{n}\mathbf{V}(X_i)=\frac{1}{n^2}\times n=\frac{1}{n}$$

par linéarité de l'espérance et par indépendance des $(X_i)_{1\leqslant i\leqslant n}$.

On en déduit donc que $\overline{X}\hookrightarrow\mathcal{N}\big(0,\frac{1}{n}\big)$.

(b) Sachant que $a_1=\frac{1}{\sqrt{n}}(1,\ldots,1)$, alors $Y_1=\langle X,a_1\rangle=\frac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{n}X_i$ donc $\overline{X}=\dfrac{1}{\sqrt{n}}Y_1$.

Par ailleurs, d'après 17.(b) : $\quad\displaystyle\sum_{k=2}^{n}\langle X\,,a_k\rangle^2=\sum_{i=1}^{n}(X_i-\overline{X})^2$, c'est-à-dire : $\quad U=\displaystyle\sum_{k=2}^{n}Y_k^2$.

(c) Au vu des expressions précédentes, et puisque $Y_1,Y_2,\ldots,Y_n$ sont mutuellement indépendantes d'après le théorème de Cochran, alors $\overline{X}=\frac{1}{\sqrt{n}}Y_1$ et $U=Y_2^2+\cdots+Y_n^2$ sont bien indépendantes, d'après le lemme des coalitions.

De plus, U apparaît comme la somme des carrés de $(n-1)$ variables aléatoires mutuellement indépendantes, chacune de loi $\mathcal{N}(0,1)$: d'après la partie 1, U suit effectivement la loi $\chi^2(n-1)$, du khi-deux à $(n-1)$ degrés de liberté.

On suppose pour la fin de ce problème, que $Z=(Z_1,\ldots,Z_n)$ est un n-uplet de variables aléatoires mutuellement indépendantes, chacune de loi normale $\mathcal{N}(\mu,\sigma^2)$ avec $\sigma>0$ où (μ,σ^2) est inconnue.
On pose :

$$\overline{Z}=\frac{Z_1+\cdots+Z_n}{n},\quad V=\sum_{i=1}^{n}(Z_i-\overline{Z})^2.$$

20. (a) Pour $i\in[\![1\,;n]\!]$, on pose $X_i=\dfrac{Z_i-\mu}{\sigma}$, et $\overline{X}=\dfrac{X_1+\cdots+X_n}{n}$, $\quad U=\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2$.

Pour tout $i \in [\![1\,;n]\!]$, on a alors $Z_i = \sigma X_i + \mu$, donc :

$$\overline{Z} = \frac{\sigma X_1 + \mu + \cdots + \sigma X_n + \mu}{n} = \sigma \frac{X_1 + \cdots + X_n}{n} + \frac{n\mu}{n} = \sigma \overline{X} + \mu,$$

et :

$$V = \sum_{i=1}^{n} (Z_i - \overline{Z})^2 = \sum_{i=1}^{n} (\sigma X_i + \mu - \sigma \overline{X} - \mu)^2 = \sigma^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sigma^2 U.$$

(b) La variable aléatoire $\dfrac{1}{n}V = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ est un estimateur de σ^2 comme fonction de l'échantillon Z dont l'expression ne dépend pas explicitement de σ.
Puisque $V = \sigma^2 U$ où $U \hookrightarrow \chi^2(n-1)$, d'après les résultats de la partie 1 et par linéarité de l'espérance :

$$\mathbf{E}(V) = \frac{\sigma^2}{n-1}\mathbf{E}(U) = \frac{\sigma^2}{n-1} \times (n-1) = \sigma^2,$$

ce qui prouve bien que $\dfrac{1}{n-1}V$ est un estimateur sans biais de σ^2.

(c) D'après les calculs précédents : $\dfrac{\overline{Z} - \mu}{\sqrt{\frac{V}{n-1}}} = \dfrac{\sigma \overline{X}}{\sqrt{\frac{\sigma^2 U}{n(n-1)}}} = \dfrac{\sqrt{n} \cdot \overline{X}}{\sqrt{U/(n-1)}}$.

Or on a vu (q. 19.(a)) que $\overline{X} \hookrightarrow \mathcal{N}\big(0, \frac{1}{n}\big)$ donc $\sqrt{n} \cdot \overline{X} \hookrightarrow \mathcal{N}(0,1)$, et $U \hookrightarrow \chi^2(n-1)$.
De plus, $\overline{X}$ et U sont indépendantes d'après 19.(c), donc $\sqrt{n} \cdot \overline{X}$ et U aussi.
On est donc dans la situation décrite à la fin de la première page de l'énoncé, qui permet de conclure que $\dfrac{\overline{Z} - \mu}{\sqrt{\frac{V}{n(n-1)}}}$ suit la loi de Student $\mathcal{T}(n-1)$.

21. Soit $\alpha \in]0\,;1[$ et $t_{n-1,\alpha}$ le réel tel que $\mathbf{P}\big(|T| \leqslant t_{n-1,\alpha}\big) = 1 - \alpha$ lorsque $T \hookrightarrow \mathcal{T}(n-1)$.

La variable aléatoire $\dfrac{\overline{Z} - \mu}{\sqrt{\frac{V}{n(n-1)}}}$ suivant cette loi de Student, on peut donc écrire :

$$P\Bigg(\Bigg|\frac{\overline{Z} - \mu}{\sqrt{\frac{V}{n(n-1)}}}\Bigg| \leqslant t_{n-1}, \alpha\Bigg) \iff \mathbf{P}\Bigg(-t_{n-1,\alpha} \leqslant \frac{\overline{Z} - \mu}{\sqrt{\frac{V}{n(n-1)}}} \leqslant t_{n-1,\alpha}\Bigg) = 1 - \alpha$$

$$\iff \mathbf{P}\Bigg(-\sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha} \leqslant \overline{Z} - \mu \leqslant \sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha}\Bigg) = 1 - \alpha \quad \text{car} \sqrt{\tfrac{V}{n(n-1)}} \text{ est à valeurs dans } \mathbb{R}_+^*$$

$$\iff \mathbf{P}\Bigg(-\sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha} - \overline{Z} \leqslant -\mu \leqslant \sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha} - \overline{Z}\Bigg) = 1 - \alpha$$

$$\iff \mathbf{P}\Bigg(\overline{Z} - \sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha} \leqslant \mu \leqslant \overline{Z} + \sqrt{\frac{n(n-1)}{V}}t_{n-1,\alpha}\Bigg) = 1 - \alpha,$$

ce qui prouve que $\Bigg[\overline{Z} - \sqrt{\dfrac{n(n-1)}{V}}t_{n-1,\alpha}\,; \overline{Z} + \sqrt{\dfrac{n(n-1)}{V}}t_{n-1,\alpha}\Bigg]$ est un intervalle de confiance de μ au niveau de confiance $1 - \alpha$.

Remarque : comme l'énoncé le précise à la fin, on obtient bien ici, sous réserve de connaître le réel $t_{n-1,\alpha}$, un intervalle de confiance non asymptotique pour l'espérance μ de la loi parente, gaussienne, de l'échantillon Z.

★★★ FIN DU SUJET ★★★