

HEC-ESCP – Mathématiques II

Mathématiques approfondies

Corrigé du sujet adapté – 2025

Les questions et sous-questions conservent la numérotation de l'énoncé. Les notations $\mathbb{E}$ et $\mathbb{V}$ désignent l'espérance et la variance. La transposée d'une matrice A est notée ${}^tA$.

Partie I – Lois du khi-deux et convergence

Partie I – Moments, loi de Student et convergence en probabilité

Les X_i sont indépendantes de loi $\mathcal{N}(0, 1)$. On pose $S_n = \sum_{i=1}^n X_i^2$ et $W_n = S_n/2$. Par définition, la loi de cette somme de n carrés de normales centrées réduites indépendantes est la loi du khi-deux $\chi^2(n)$. On note φ et Φ la densité et la fonction de répartition normales centrées réduites définies dans l'énoncé.

Corrigé – Partie I.

1. Espérance et simulation de S_n .

- a) Puisque X_i possède un second moment, $\mathbb{E}(X_i^2) = \mathbb{V}(X_i) + \mathbb{E}(X_i)^2 = 1$. La somme finie S_n admet donc une espérance et, par linéarité,

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

- b) On simule n normales centrées réduites indépendantes, puis on additionne leurs carrés.

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

`simul(n)` simule une réalisation de S_n .

- c) Pour n fixé, la moyenne des N simulations de S_n^2 approche $\mathbb{E}(S_n^2)$, dont l'existence est admise. Cette moyenne empirique fournit une estimation de ce moment. Ainsi `f(n, N)` approche $\mathbb{E}(S_n^2) - n^2 = \mathbb{V}(S_n)$. Les points obtenus sont proches de la droite d'équation $y = 2n$:

On conjecture que $\mathbb{V}(S_n) = 2n$.

Cette observation sera démontrée à la question 2.d).

2. Loi de W_n et moments de S_n .

- a) Comme $W_1 = X_1^2/2$, sa fonction de répartition est nulle sur $] -\infty, 0]$. Pour $x > 0$, la continuité de la loi normale et la symétrie de sa densité donnent

$$\mathbb{P}(W_1 \leq x) = \mathbb{P}(-\sqrt{2x} \leq X_1 \leq \sqrt{2x}) = 2\Phi(\sqrt{2x}) - 1.$$

Cette expression tend vers 0 en 0^+ , donc la fonction de répartition est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur chacun des intervalles $] -\infty, 0[$ et $]0, +\infty[$. Pour $x > 0$, sa dérivée vaut

$$\frac{2\varphi(\sqrt{2x})}{\sqrt{2x}} = \frac{x^{-1/2}e^{-x}}{\sqrt{\pi}}.$$

On peut choisir la valeur 0 de la densité en 0 :

$$f_{W_1}(x) = \begin{cases} \frac{x^{-1/2}e^{-x}}{\sqrt{\pi}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

b) Puisqu'une densité s'intègre à 1,

$$1 = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} x^{-1/2}e^{-x} dx = \frac{\Gamma(1/2)}{\sqrt{\pi}}.$$

Ainsi

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

c) La loi gamma de forme $a > 0$, notée $\gamma(a)$, admet la densité $x^{a-1}e^{-x}/\Gamma(a)$ pour $x > 0$, nulle pour $x \leq 0$. Les questions a) et b) montrent que $W_1 \sim \gamma(1/2)$. Chacune des variables $X_i^2/2$ suit la même loi ; ces variables sont indépendantes car elles sont des fonctions des X_i indépendantes. La stabilité par somme des lois gamma de même échelle 1 donne

$$W_n = \sum_{i=1}^n \frac{X_i^2}{2} \sim \gamma(n/2).$$

En particulier, sa densité est

$$f_{W_n}(x) = \begin{cases} \frac{x^{n/2-1}e^{-x}}{\Gamma(n/2)}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

d) Pour une variable de loi $\gamma(a)$, les moments d'ordre 1 et 2 existent et valent a et $a(a+1)$; sa variance vaut donc a . Avec $a = n/2$ et $S_n = 2W_n$, on retrouve

$$\mathbb{E}(S_n) = 2\mathbb{E}(W_n) = n, \quad \mathbb{V}(S_n) = 4\mathbb{V}(W_n) = 2n.$$

La conjecture de 1.c) est vérifiée.

3. Moments négatifs de S_n .

a) Pour $n \geq 3$, $n/2 - 1 > 0$. La formule de transfert pour l'espérance d'une fonction de W_n donne

$$\mathbb{E}\left(\frac{1}{W_n}\right) = \frac{1}{\Gamma(n/2)} \int_0^{+\infty} x^{n/2-2}e^{-x} dx.$$

L'intégrale converge car elle est $\Gamma(n/2 - 1)$, dont le paramètre est strictement positif. La récurrence de la fonction Gamma donne

$$\mathbb{E}\left(\frac{1}{W_n}\right) = \frac{\Gamma(n/2 - 1)}{\Gamma(n/2)} = \frac{2}{n - 2}.$$

Puisque $S_n = 2W_n$,

$$\mathbb{E}\left(\frac{1}{S_n}\right) = \frac{1}{n-2}.$$

- b) Pour tout réel $s > 0$, $s^{-1/2} \leq 1 + s^{-1}$: si $s \geq 1$, le premier membre est au plus 1, et si $s < 1$, il est au plus s^{-1} . Les variables S_n sont strictement positives par l'hypothèse de l'énoncé. Pour $n \geq 3$, on obtient

$$0 < \frac{1}{\sqrt{S_n}} \leq 1 + \frac{1}{S_n}.$$

Le majorant admet une espérance finie d'après a), donc

$$\forall n \geq 3, \quad \mathbb{E}\left(\frac{1}{\sqrt{S_n}}\right) \text{ existe et est finie.}$$

4. Existence et unicité du quantile de Student.

La loi de Student $\mathcal{T}(n)$ est définie ici comme la loi de $T_n = Y/\sqrt{S_n/n}$, où $Y \sim \mathcal{N}(0, 1)$ est indépendante des variables X_i définissant S_n . L'énoncé admet pour T_n une densité f_n strictement positive et continue sur $\mathbb{R}$. Pour $t \geq 0$, posons

$$g_n(t) = \mathbb{P}(|T_n| \leq t) = \int_{-t}^t f_n(x) dx.$$

La fonction g_n est continue, $g_n(0) = 0$, et $g_n(t) \rightarrow 1$ lorsque $t \rightarrow +\infty$, car f_n est une densité. Elle est strictement croissante : si $0 \leq s < t$, les intégrales sur $[-t, -s]$ et $[s, t]$ sont strictement positives. Comme $1 - \alpha \in]0, 1[$, le théorème de la bijection donne un unique $t_{n,\alpha} > 0$ tel que $g_n(t_{n,\alpha}) = 1 - \alpha$. Pour $t < 0$, l'événement $[|T_n| \leq t]$ est vide, donc aucune autre solution réelle n'existe.

$$\exists! t_{n,\alpha} > 0, \quad \mathbb{P}(|T_n| \leq t_{n,\alpha}) = 1 - \alpha.$$

5. Moments de T_n et écart avec Y .

Dans cette question, $n \geq 3$. L'indépendance de Y et des X_i entraîne celle de Y et de toute fonction de S_n .

- a) La normale Y admet une espérance absolue finie. La question 3.b) et l'indépendance donnent

$$\mathbb{E}(|T_n|) = \sqrt{n} \mathbb{E}(|Y|) \mathbb{E}(S_n^{-1/2}) < +\infty.$$

Ainsi l'espérance existe et se factorise :

$$\mathbb{E}(T_n) = \sqrt{n} \mathbb{E}(Y) \mathbb{E}(S_n^{-1/2}) = 0.$$

- b) Comme Y^2 et $1/S_n$ sont indépendantes et d'espérance finie,

$$\mathbb{E}(T_n^2) = n \mathbb{E}(Y^2) \mathbb{E}(1/S_n) = \frac{n}{n-2}.$$

Avec a), on obtient

$$\mathbb{V}(T_n) = \frac{n}{n-2}.$$

- c) Le produit $T_n Y = \sqrt{n} Y^2 / \sqrt{S_n}$ est positif et admet une espérance finie par indépendance et 3.b). En écrivant $S_n = 2W_n$ et $\mathbb{E}(Y^2) = 1$,

$$\mathbb{E}(T_n Y) = \sqrt{\frac{n}{2}} \mathbb{E}(W_n^{-1/2}).$$

Le développement du carré donne donc

$$\begin{aligned} \mathbb{E}((T_n - Y)^2) &= \mathbb{E}(T_n^2) + \mathbb{E}(Y^2) - 2\mathbb{E}(T_n Y) \\ &= \frac{n}{n-2} + 1 - \sqrt{2n} \mathbb{E}(W_n^{-1/2}). \end{aligned}$$

Ainsi

$$\mathbb{E}((T_n - Y)^2) = \frac{2n-2}{n-2} - \sqrt{2n} \mathbb{E}\left(\frac{1}{\sqrt{W_n}}\right).$$

6. Étude de la suite u_n .

Pour $n \geq 2$, la formule de transfert et la densité de W_n donnent

$$\mathbb{E}(W_n^{-1/2}) = \frac{1}{\Gamma(n/2)} \int_0^{+\infty} x^{(n-1)/2-1} e^{-x} dx.$$

Cette intégrale converge car $(n-1)/2 > 0$; ceci justifie notamment l'existence de u_2 , non couverte par la question 3.b). Pour la suite, on dispose donc de la formule

$$u_n = \frac{\Gamma((n-1)/2)}{\Gamma(n/2)} > 0 \quad (n \geq 2).$$

- a) Puisque $W_{n+1} = W_n + X_{n+1}^2/2 > W_n > 0$, on a $W_{n+1}^{-1/2} < W_n^{-1/2}$. En prenant les espérances de ces variables d'espérance finie, on obtient $u_{n+1} \leq u_n$. En utilisant la formule précédente, $\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1$ et la question 2.b),

$$(u_n)_{n \geq 2} \text{ est décroissante,} \quad u_2 = \sqrt{\pi}.$$

- b) Les deux quotients de Gamma se simplifient :

$$u_{n+1} u_n = \frac{\Gamma(n/2)}{\Gamma((n+1)/2)} \frac{\Gamma((n-1)/2)}{\Gamma(n/2)} = \frac{\Gamma((n-1)/2)}{\Gamma((n+1)/2)}.$$

La relation $\Gamma((n+1)/2) = ((n-1)/2) \Gamma((n-1)/2)$ donne

$$\forall n \geq 2, \quad u_{n+1} u_n = \frac{2}{n-1}, \quad u_{n+1} = \frac{2}{(n-1)u_n}.$$

- c) Le tableau a $n-1$ entrées ; l'entrée d'indice $j-2$ contient u_j .

def suite_u(n):

 U = np.zeros(n - 1)

 U[0] = np.sqrt(np.pi)

 for j in range(2, n):

 U[j - 1] = 2 / ((j - 1) * U[j - 2])

 return U

$$\text{suite_u}(n) \text{ renvoie } (u_2, u_3, \dots, u_n), \quad n \geq 2.$$

- d) Les points représentant nu_n^2 semblent tendre vers 2. Puisque $u_n > 0$, cela suggère l'équivalence suivante :

$$\boxed{nu_n^2 \longrightarrow 2, \quad u_n \sim \sqrt{2/n}.}$$

Ici l'équivalence signifie précisément que $u_n/\sqrt{2/n} \rightarrow 1$. Elle est démontrée ci-dessous.

- e) Par décroissance, $u_{n+1} \leq u_n \leq u_{n-1}$ pour $n \geq 3$. En utilisant b) aux indices n et $n-1$, on obtient

$$\frac{2}{n-1} = u_n u_{n+1} \leq u_n^2 \leq u_{n-1} u_n = \frac{2}{n-2}.$$

Ainsi

$$\frac{n}{n-1} \leq \frac{nu_n^2}{2} \leq \frac{n}{n-2}.$$

Les deux bornes tendent vers 1. Le théorème d'encadrement et la continuité de la racine carrée donnent

$$\boxed{\frac{u_n}{\sqrt{2/n}} \longrightarrow 1, \quad u_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{n}}.}$$

7. Convergence en probabilité de T_n vers Y .

Par les questions 5.c) et 6.e), pour $n \geq 3$,

$$\mathbb{E}((T_n - Y)^2) = \frac{2n-2}{n-2} - \sqrt{2n} u_n \longrightarrow 2 - 2 = 0.$$

D'après 5.a), $\mathbb{E}(T_n - Y) = 0$, donc $\mathbb{V}(T_n - Y) = \mathbb{E}((T_n - Y)^2)$. Pour tout $\varepsilon > 0$, l'inégalité de Bienaymé–Tchebychev donne :

$$\mathbb{P}(|T_n - Y| > \varepsilon) \leq \frac{\mathbb{E}((T_n - Y)^2)}{\varepsilon^2} \longrightarrow 0.$$

C'est exactement la définition de la convergence en probabilité :

$$\boxed{T_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} Y.}$$

Aucun moment de T_2 n'est nécessaire à cette conclusion, qui concerne les indices tendant vers l'infini.

Partie II – Transformation de deux lois normales

Partie II – Indépendance après transformation

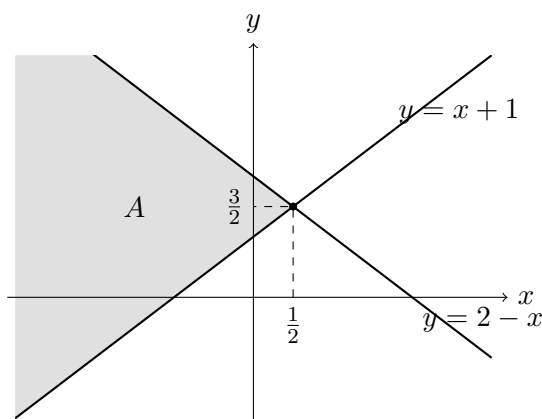
Pour $a, b \in \mathbb{R}$, on pose $A = \{(x, y) : x + y \leq a, x - y \leq b\}$, $c = (a + b)/2$ et $d = (a - b)/2$. Les variables X, Y sont indépendantes de loi $\mathcal{N}(0, 1)$. On utilise les deux formules de probabilités et d'intégrales explicitement admises dans cette partie de l'énoncé.

Corrigé – Partie II.

8. Représentation de A pour $a = 2$, $b = -1$.

Les inégalités s'écrivent $y \leq 2 - x$ et $y \geq x + 1$. Les deux droites se coupent en $(1/2, 3/2)$. La région est comprise entre ces deux droites, à gauche de leur point d'intersection, avec ses frontières incluses :

$$A = \{(x, y) : x \leq 1/2, \quad x + 1 \leq y \leq 2 - x\}.$$



9. Section de A lorsque $y \geq d$.

- a) Les deux conditions définissant A équivalent à $x \leq a - y$ et $x \leq b + y$. Or $y \geq (a - b)/2$ donne $a - y \leq b + y$. La plus restrictive est donc la première :

$$y \geq d \implies ((x, y) \in A \iff x \leq a - y).$$

- b) Pour $y \geq d$, la fonction indicatrice vaut 1 pour $x \leq a - y$ et 0 au-delà. Par définition de Φ ,

$$\int_{-\infty}^{+\infty} \mathbf{1}_A(x, y) \varphi(x) dx = \int_{-\infty}^{a-y} \varphi(x) dx = \Phi(a - y).$$

10. Section de A lorsque $y \leq d$.

Cette fois $b + y \leq a - y$, donc $(x, y) \in A$ équivaut à $x \leq b + y$. On en déduit

$$y \leq d \implies \int_{-\infty}^{+\infty} \mathbf{1}_A(x, y) \varphi(x) dx = \Phi(b + y).$$

Les deux expressions coïncident en $y = d$, où $a - d = b + d = c$.

11. Fermeture de A et première expression de sa probabilité.

- a) Les applications $(x, y) \mapsto x + y$ et $(x, y) \mapsto x - y$ sont linéaires, donc continues. La partie A est l'intersection de leurs images réciproques des fermés $] - \infty, a]$ et $] - \infty, b]$.

$$A \text{ est une partie fermée de } \mathbb{R}^2.$$

- b) La formule de probabilité admise dans l'énoncé est applicable puisque A est fermée. Les questions 9 et 10 donnent

$$\mathbb{P}((X, Y) \in A) = \int_{-\infty}^d \Phi(b + y) \varphi(y) dy + \int_d^{+\infty} \Phi(a - y) \varphi(y) dy.$$

Ces intégrales convergent, car $0 \leq \Phi \leq 1$ et φ est une densité. Dans la première intégrale simple, poser $y = d - z$; dans la seconde, poser $y = d + z$. Ces applications affines sont de classe $\mathcal{C}^1$, bijectives des demi-intervalles considérés, de dérivées non nulles. Les changements de variable conservent donc la nature des intégrales impropres. Comme $b + d = a - d = c$, on obtient

$$\mathbb{P}((X, Y) \in A) = \int_0^{+\infty} (\varphi(d + z) + \varphi(d - z)) \Phi(c - z) dz.$$

12. Une expression utilisant l'identité admise.

Pour $z \geq 0$ fixé, dans l'intégrale simple définissant $\Phi(c - z)$, on effectue la translation $s = t - z$. C'est une bijection affine de classe $\mathcal{C}^1$, de dérivée 1, de $] - \infty, c]$ sur $] - \infty, c - z]$; les intégrales impropres ont même nature. Ainsi

$$\Phi(c - z) = \int_{-\infty}^{c-z} \varphi(s) ds = \int_{-\infty}^c \varphi(t - z) dt.$$

En remplaçant cette quantité dans la formule de 11.b),

$$\mathbb{P}((X, Y) \in A) = \int_0^{+\infty} \left(\int_{-\infty}^c (\varphi(d + z) + \varphi(d - z)) \varphi(t - z) dt \right) dz.$$

Pour la suite, on utilise directement l'égalité admise à la question 12 de l'énoncé :

$$\mathbb{P}((X, Y) \in A) = \int_{-\infty}^c \left(\int_0^{+\infty} (\varphi(d + z) + \varphi(d - z)) \varphi(t - z) dz \right) dt.$$

13. Identité entre produits de densités normales.

L'identité algébrique

$$\left(\frac{u + v}{\sqrt{2}} \right)^2 + \left(\frac{u - v}{\sqrt{2}} \right)^2 = u^2 + v^2$$

montre que les exponentielles intervenant dans les deux produits de densités sont identiques. Le facteur multiplicatif est $1/(2\pi)$ dans les deux cas. Donc

$$\varphi(u) \varphi(v) = \varphi \left(\frac{u + v}{\sqrt{2}} \right) \varphi \left(\frac{u - v}{\sqrt{2}} \right).$$

14. Calcul des intégrales intérieures.

Fixons $t \in \mathbb{R}$. La question 13 donne d'abord

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

Dans l'intégrale simple en z , le changement affine $r = (d-t+2z)/\sqrt{2}$, de dérivée $\sqrt{2}$, est une bijection $\mathcal{C}^1$ de $[0, +\infty[$ sur $[(d-t)/\sqrt{2}, +\infty[$. La convergence est celle d'une intégrale de densité normale. Il donne

$$\begin{aligned} \int_0^{+\infty} \varphi(d+z)\varphi(t-z) dz &= \frac{1}{\sqrt{2}} \varphi\left(\frac{t+d}{\sqrt{2}}\right) \left(1 - \Phi\left(\frac{d-t}{\sqrt{2}}\right)\right) \\ &= \frac{1}{\sqrt{2}} \varphi\left(\frac{t+d}{\sqrt{2}}\right) \Phi\left(\frac{t-d}{\sqrt{2}}\right). \end{aligned}$$

La dernière égalité utilise $1 - \Phi(s) = \Phi(-s)$, conséquence de la symétrie de la densité normale. De même, la question 13 et la parité de φ donnent

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

Cette fois $r = (t+d-2z)/\sqrt{2}$ est une bijection affine $\mathcal{C}^1$, de dérivée $-\sqrt{2}$, de $[0, +\infty[$ sur $] -\infty, (t+d)/\sqrt{2}]$. Le changement de variable, dans cette seule intégrale simple convergente, donne

$$\int_0^{+\infty} \varphi(d-z)\varphi(t-z) dz = \frac{1}{\sqrt{2}} \varphi\left(\frac{t-d}{\sqrt{2}}\right) \Phi\left(\frac{t+d}{\sqrt{2}}\right).$$

On additionne les deux expressions, puis on les remplace dans l'identité admise en 12 :

$$\begin{aligned} \mathbb{P}((X, Y) \in A) &= \frac{1}{\sqrt{2}} \int_{-\infty}^c \left(\varphi\left(\frac{t+d}{\sqrt{2}}\right) \Phi\left(\frac{t-d}{\sqrt{2}}\right) \right. \\ &\quad \left. + \varphi\left(\frac{t-d}{\sqrt{2}}\right) \Phi\left(\frac{t+d}{\sqrt{2}}\right) \right) dt. \end{aligned}$$

15. Calcul final de la probabilité.

La fonction réelle

$$H(t) = \Phi\left(\frac{t+d}{\sqrt{2}}\right) \Phi\left(\frac{t-d}{\sqrt{2}}\right)$$

est de classe $\mathcal{C}^1$, puisque $\Phi' = \varphi$. La dérivation de ce produit de fonctions d'une seule variable donne

$$H'(t) = \frac{1}{\sqrt{2}} \left(\varphi\left(\frac{t+d}{\sqrt{2}}\right) \Phi\left(\frac{t-d}{\sqrt{2}}\right) + \varphi\left(\frac{t-d}{\sqrt{2}}\right) \Phi\left(\frac{t+d}{\sqrt{2}}\right) \right).$$

Sur tout segment $[-R, c]$ avec R assez grand, $\int_{-R}^c H'(t) dt = H(c) - H(-R)$. Comme $\Phi(s) \rightarrow 0$ en $-\infty$ et $0 \leq \Phi \leq 1$, $H(-R) \rightarrow 0$. La formule de 14 devient donc

$$\mathbb{P}((X, Y) \in A) = \Phi\left(\frac{c+d}{\sqrt{2}}\right) \Phi\left(\frac{c-d}{\sqrt{2}}\right) = \Phi\left(\frac{a}{\sqrt{2}}\right) \Phi\left(\frac{b}{\sqrt{2}}\right).$$

16. Loi et indépendance de W et Z .

Pour des réels w, z , choisir $a = \sqrt{2}w$ et $b = \sqrt{2}z$ donne

$$\mathbb{P}(W \leq w, Z \leq z) = \Phi(w)\Phi(z), \quad W = \frac{X+Y}{\sqrt{2}}, \quad Z = \frac{X-Y}{\sqrt{2}}.$$

En faisant tendre z vers $+\infty$, les événements considérés augmentent vers $[W \leq w]$, et on obtient $\mathbb{P}(W \leq w) = \Phi(w)$. De même, $\mathbb{P}(Z \leq z) = \Phi(z)$. La fonction de répartition du couple est donc le produit de ses fonctions de répartition marginales, ce qui caractérise l'indépendance.

$W \sim \mathcal{N}(0, 1), \quad Z \sim \mathcal{N}(0, 1), \quad W \text{ et } Z \text{ sont indépendantes.}$

Partie III – Théorème de Cochran et estimation gaussienne

Partie III – Décomposition orthogonale et intervalle de confiance

La base $(a_1, \dots, a_n)$ est orthonormée, avec $a_1 = (1, \dots, 1)/\sqrt{n}$. Le théorème de Cochran admis dans l'énoncé affirme que les coordonnées $Y_k = \langle X, a_k \rangle$ d'un vecteur de normales centrées réduites indépendantes sont elles-mêmes des normales centrées réduites mutuellement indépendantes.

Corrigé – Partie III.

17. Décomposition du vecteur centré.

- a) Dans une base orthonormée, $x = \sum_{k=1}^n \langle x, a_k \rangle a_k$. Or $\langle x, a_1 \rangle = \sqrt{n} \bar{x}$ et $\langle x, a_1 \rangle a_1 = (\bar{x}, \dots, \bar{x})$. En soustrayant ce terme,

$$\sum_{k=2}^n \langle x, a_k \rangle a_k = (x_1 - \bar{x}, \dots, x_n - \bar{x}).$$

- b) Par orthonormalité, le carré de la norme du membre de gauche est $\sum_{k=2}^n \langle x, a_k \rangle^2$. Le carré de la norme du membre de droite est $\sum_{i=1}^n (x_i - \bar{x})^2$. Ainsi

$$\sum_{k=2}^n \langle x, a_k \rangle^2 = \sum_{i=1}^n (x_i - \bar{x})^2.$$

18. Transformation de deux normales et covariance.

- a) Pour $n = 2$, les vecteurs $a_1 = (1, 1)/\sqrt{2}$ et $a_2 = (1, -1)/\sqrt{2}$ forment une base orthonormée. Le théorème de Cochran donne directement

$$\frac{X_1 + X_2}{\sqrt{2}} \text{ et } \frac{X_1 - X_2}{\sqrt{2}} \text{ sont indépendantes, de loi } \mathcal{N}(0, 1).$$

C'est le résultat de la partie II.

- b) Les variables R_1 et R_2 admettent des seconds moments, car ce sont des combinaisons linéaires de normales indépendantes. Elles sont centrées. Par bilinéarité de la covariance, avec $\text{Cov}(X_1, X_2) = 0$ et $\mathbb{V}(X_1) = \mathbb{V}(X_2) = 1$,

$$\text{Cov}(R_1, R_2) = \beta_1 + \beta_2.$$

Si R_1, R_2 sont indépendantes, leur covariance est nulle. Réciproquement, si cette covariance est nulle, $\beta_2 = -\beta_1$. L'hypothèse $(\beta_1, \beta_2) \neq (0, 0)$ implique alors $\beta_1 \neq 0$ et

$$R_1 = \sqrt{2} \frac{X_1 + X_2}{\sqrt{2}}, \quad R_2 = \sqrt{2} \beta_1 \frac{X_1 - X_2}{\sqrt{2}}.$$

Ces deux variables sont des fonctions séparées des deux variables indépendantes de a), donc elles sont indépendantes.

$$R_1 \text{ et } R_2 \text{ indépendantes} \iff \text{Cov}(R_1, R_2) = 0 \iff \beta_1 + \beta_2 = 0.$$

19. Moyenne empirique et somme des carrés centrés.

- a) La stabilité par somme de normales indépendantes donne $\sum_i X_i \sim \mathcal{N}(0, n)$. En divisant par n , on obtient

$$\bar{X} \sim \mathcal{N}\left(0, \frac{1}{n}\right).$$

- b) La définition de a_1 et la question 17.b) donnent respectivement

$$\bar{X} = \frac{Y_1}{\sqrt{n}}, \quad U = \sum_{k=2}^n Y_k^2.$$

Ces identités valent presque sûrement si des modifications sur un événement de probabilité nulle ont été effectuées comme le permet l'énoncé.

- c) D'après Cochran, Y_1 est indépendante du vecteur $(Y_2, \dots, Y_n)$; par fonctions séparées, $\bar{X}$ et U sont donc indépendantes. Les $n - 1$ variables $Y_2, \dots, Y_n$ sont des normales centrées réduites indépendantes. Par la définition du khi-deux comme somme de leurs carrés,

$$\bar{X} \text{ et } U \text{ sont indépendantes, } \quad U \sim \chi^2(n-1).$$

La question 2 donne en particulier $\mathbb{E}(U) = n - 1$.

20. Estimation de σ^2 et statistique de Student.

Les Z_i sont indépendantes de loi $\mathcal{N}(\mu, \sigma^2)$, avec $\sigma > 0$. Les variables $X_i = (Z_i - \mu)/\sigma$ sont alors indépendantes de loi $\mathcal{N}(0, 1)$.

- a) On a $Z_i = \mu + \sigma X_i$ et $\bar{Z} = \mu + \sigma \bar{X}$. Ainsi $Z_i - \bar{Z} = \sigma(X_i - \bar{X})$; en élevant au carré et en sommant,

$$V = \sigma^2 U.$$

- b) La question 19 donne l'existence de l'espérance de U , égale à $n - 1$. Donc

$$\mathbb{E}\left(\frac{V}{n-1}\right) = \frac{\sigma^2}{n-1} \mathbb{E}(U) = \sigma^2.$$

La statistique $V/(n-1)$ ne dépend pas des paramètres inconnus et son espérance est le paramètre à estimer. C'est donc un estimateur sans biais :

$$\widehat{\sigma^2} = \frac{V}{n-1} \text{ est un estimateur sans biais de } \sigma^2.$$

- c) Avec $\sigma > 0$ et a), la statistique proposée s'écrit

$$\frac{\bar{Z} - \mu}{\sqrt{V/(n(n-1))}} = \frac{\sigma \bar{X}}{\sigma \sqrt{U/(n(n-1))}} = \frac{\sqrt{n} \bar{X}}{\sqrt{U/(n-1)}} = \frac{Y_1}{\sqrt{(\sum_{k=2}^n Y_k^2)/(n-1)}}.$$

Le numérateur Y_1 est une normale centrée réduite indépendante des $n - 1$ normales $Y_2, \dots, Y_n$ figurant au dénominateur. C'est exactement la définition de la loi de Student de paramètre $n - 1$ rappelée en question 4. Ainsi

$$\frac{\bar{Z} - \mu}{\sqrt{V/(n(n-1))}} \sim \mathcal{T}(n-1).$$

Le résultat vaut également pour $n = 2$; il ne nécessite aucun moment de la loi de Student correspondante.

21. Intervalle de confiance de la moyenne μ .

Par définition de $t_{n-1,\alpha}$ et par la question 20.c),

$$\mathbb{P}\left(\left|\frac{\bar{Z} - \mu}{\sqrt{V/(n(n-1))}}\right| \leq t_{n-1,\alpha}\right) = 1 - \alpha.$$

Le dénominateur est strictement positif suivant l'hypothèse de l'énoncé. Pour chaque réalisation, l'inégalité entre valeurs absolues équivaut à

$$\bar{Z} - t_{n-1,\alpha}\sqrt{\frac{V}{n(n-1)}} \leq \mu \leq \bar{Z} + t_{n-1,\alpha}\sqrt{\frac{V}{n(n-1)}}.$$

On obtient l'intervalle aléatoire

$$I_\alpha = \left[\bar{Z} - t_{n-1,\alpha}\sqrt{\frac{V}{n(n-1)}} ; \bar{Z} + t_{n-1,\alpha}\sqrt{\frac{V}{n(n-1)}} \right].$$

Ses bornes sont calculables à partir des observations, de n et du quantile tabulé. Pour toutes les valeurs $\mu \in \mathbb{R}$ et $\sigma > 0$,

$$\mathbb{P}(\mu \in I_\alpha) = 1 - \alpha.$$

C'est un intervalle de confiance de niveau exact $1 - \alpha$, valable pour chaque $n \geq 2$ dans le modèle gaussien considéré.