DM12

# Correction du devoir maison

## Partie I. Fonction logistique et lois logistiques.

1. (a) La fonction $\Lambda$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ en tant que quotient de fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}$ dont le dénominateur ne s'annule pas. De plus, on a :

$$\forall x \in \mathbb{R}, \ \Lambda'(x) = \frac{e^{-x}}{(1+e^{-x})^2} > 0.$$

Donc $\Lambda$ est strictement croissante sur $\mathbb{R}$. Enfin, on a :

$$\lim_{x\to-\infty} \Lambda(x) = 0 \quad \text{et} \quad \lim_{x\to+\infty} \Lambda(x) = 1$$

$\Lambda$ est donc continue et strictement croissante. $\boxed{\Lambda \text{ réalise une bijection de } \mathbb{R} \text{ dans } ]\lim_{-\infty}\Lambda, \lim_{+\infty}\Lambda[=]0,1[.}$

**Remarque.** On notera que $\Lambda$ est une fonction de répartition d'une variable à densité car elle est de classe $\mathscr{C}^1$, croissante, et que $\lim\limits_{x\to-\infty} \Lambda(x) = 0$ et $\lim\limits_{x\to+\infty} \Lambda(x) = 1$.

Déterminons sa bijection réciproque. Soit pour cela $x \in \mathbb{R}$ et $y \in ]0,1[$. On résout :

$$\begin{aligned}
\Lambda(x) = y \Leftrightarrow \frac{1+e^{-x}}{} = y \Leftrightarrow 1+e^{-x} &= \frac{1}{y} \\
\Leftrightarrow e^{-x} = \frac{1}{y} - 1 &= \frac{1-y}{y} \\
\Leftrightarrow x = -\ln\left(\frac{1-y}{y}\right) &= \ln\left(\frac{y}{1-y}\right)
\end{aligned}$$

$$\boxed{\text{La bijection réciproque de } \Lambda \text{ est donc } L : x \in ]0,1[ \mapsto \ln\left(\frac{x}{1-x}\right).}$$

(b) D'après le calcul précédent :

$$\boxed{\forall x \in \mathbb{R}, \quad \Lambda'(x) = \frac{e^{-x}}{(1+e^{-x})^2}.}$$

(c) On pose la fonction $g$ définie sur $\mathbb{R}$ par $g : x \mapsto \Lambda(x) - x$. $g$ est dérivable en tant que somme de telles fonctions et on a :

$$\forall x \in \mathbb{R}, \quad g'(x) = \frac{e^{-x}}{(1+e^{-x})^2} - 1 = -\frac{1+e^{-x}+e^{-2x}}{(1+e^{-x})^2} < 0.$$

Donc $g$ est donc strictement décroissante. De plus, on a :

$$\lim_{-\infty} g = +\infty \quad \text{et} \quad \lim_{+\infty} g = -\infty.$$

Comme $g$ est continue et strictement décroissante sur $\mathbb{R}$, $g$ réalise une bijection de $\mathbb{R}$ dans $\mathbb{R}$. En particulier, $\boxed{\text{il existe un unique } x_0 \in \mathbb{R}, \text{ tel que } g(x_0) = 0, \text{ ce qui se réécrit } \Lambda(x_0) = x_0.}$

(d) Avec les notations précédentes, l'inégalité demandée est en fait :

$$\forall x \in \mathbb{R}, \quad |g(x) - g(x_0)| \leq |x - x_0|.$$

Or $g$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et sa dérivée vérifie :

$$\forall x \in \mathbb{R}, \ |g'(x)| = \left| -\frac{1+e^{-x}+e^{-2x}}{(1+e^{-x})^2} \right| = \frac{1+e^{-x}+e^{-2x}}{(1+e^{-x})^2} = \frac{(1+e^{-x})^2 - e^{-x}}{(1+e^{-x})^2} = 1 \underbrace{- \frac{e^{-x}}{(1+e^{-x})^2}}_{\leq 0} \leq 1.$$

Par l'inégalité des accroissements finis, on a bien :

$$\forall x \in \mathbb{R}, \quad |g(x) - g(x_0)| \leq |x - x_0| \quad \Rightarrow \quad \boxed{\forall x \in \mathbb{R}, \quad |\Lambda(x) - x| \leq |x - x_0|.}$$

2. (a) On utilise l'algorithme de dichotomie avec la fonction $g$ définie précédemment pour estimer la valeur de $x_0$ tel que $g(x_0) = 0$. Comme $g$ est décroissante et que $g(0) = 1/2 > 0$ et $g(1) = -1/(e+1) < 0$, on sait, par le théorème de la bijection, que $g$ s'annule sur $]0, 1[$. On peut donc restreindre la recherche sur $]0, 1[$ pour trouver une valeur approchée de $x_0$ : cela justifie les choix de $a$ et $b$.

   La condition `Lambda(c)>c` revient donc à $g(c) > 0$ donc comme $g$ est décroissante, la condition $g(c) > 0$ signifie que la fonction s'annule sur $]c, b[$, sinon, la fonction s'annule sur $]a, c[$. On propose donc :

```
7 Lambda(c)>c then a=c ; else b=c ; end ;
```

   (b) À la fin de la boucle `while`, on a $0 < b-a \leq \varepsilon$, et on renvoie $\dfrac{a+b}{2}$ en tant qu'approximation de $x_0$. Or on sait que $x_0 \in [a, b] = \left[a, \frac{a+b}{2}\right] \cup \left[\frac{a+b}{2}, b\right]$, union d'intervalles tous deux de longueur $\dfrac{b-a}{2}$, de sorte que :

$$\left|x_0 - \frac{a+b}{2}\right| \leq \frac{b-a}{2} \leq \frac{\varepsilon}{2}.$$

   Pour être assurée que l'erreur d'approximation commise ne dépasse pas $10^{-4}$, la valeur maximale que l'on peut affecter à la variable `eps` en ligne 4 est $\boxed{2 \times 10^{-4}.}$

   (c) La variable `x0` du programme `Scilab` n'est pas $x_0$, mais une approximation de ce réel à $\frac{\varepsilon}{2}$ près. La valeur renvoyée par l'instruction (10) n'est donc pas 0 a priori. On peut cependant utiliser l'inégalité de la question 1.(d) pour affirmer que :

$$|\Lambda(\mathtt{x0}) - \mathtt{x0}| \leq |\mathtt{x0} - x_0| \leq \frac{\varepsilon}{2}.$$

   Si $\varepsilon = 2 \times 10^{-4}$, $\boxed{\Lambda(\mathtt{x0}) - \mathtt{x0} \text{ est donc un réel compris entre } -10^{-4} \text{ et } 10^{-4}.}$

3. (a) On a vu en question 1.(a) que $\Lambda$ est une fonction de répartition d'une variable à densité. De plus, comme $\Lambda$ est de classe $\mathscr{C}^1$ sur tout $\mathbb{R}$, $\boxed{\text{sa dérivée } \lambda \text{ est bien une densité de probabilité.}}$

   (b) On rappelle que :

$$\forall x \in \mathbb{R}, \quad \lambda(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$

   Vérifions que $\lambda$ est paire. Elle est définie sur l'intervalle $\mathbb{R}$ symétrique par rapport à 0, et on a pour tout $x \in \mathbb{R}$ :

$$\lambda(-x) = \frac{e^x}{(1+e^x)^2} = \frac{e^x}{(e^x(e^{-x}+1))^2} = \frac{e^{-x}}{(e^{-x}+1)^2} = \lambda(x).$$

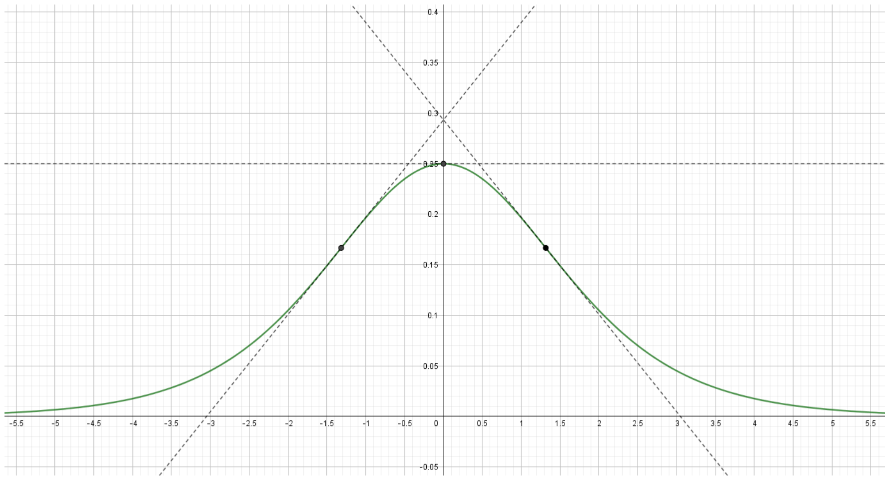
   $\boxed{\text{La fonction } \lambda \text{ est donc paire.}}$

   $\lambda$ est de classe $\mathscr{C}^2$ en tant que quotient de telles fonctions (dont le dénominateur ne s'annule pas) et on a pour tout $x \in \mathbb{R}$ :

$$\lambda'(x) = -\frac{e^{-x}(1-e^{-x})}{(1+e^{-x})^3} = -\frac{e^x(e^x-1)}{(e^x+1)^3} \quad \text{et} \quad \lambda''(x) = \frac{e^{-x}(1-4e^{-x}+e^{-2x})}{(1+e^{-x})^4} = \frac{e^x(e^{2x}-4e^x+1)}{(e^x+1)^4}.$$

   $\lambda'$ ne s'annule qu'en $x = 0$. De plus, $\lambda'$ est positive sur $\mathbb{R}_-$ et négative sur $\mathbb{R}_+$, de sorte que $\lambda$ est croissante sur $\mathbb{R}_-$ et décroissante sur $\mathbb{R}^+$ et sa courbe représentative présente une tangente horizontale en 0.

   Représentons la courbe représentative de $\lambda$.



Déterminons les points d'inflexion. Le signe de $\lambda''$ dépend uniquement du facteur :

$$e^{2x} - 4e^x + 1 = (e^x - 2)^2 - 3 = (e^x - 2 + \sqrt{3})(e^x - 2 - \sqrt{3}).$$

On effectue le tableau de signes suivant :

| $x$ | $-\infty$ | | $\ln(2-\sqrt{3})$ | | $\ln(2+\sqrt{3})$ | | $+\infty$ |
|---|---|---|---|---|---|---|---|
| $e^x - 2 + \sqrt{3}$ | | $-$ | $0$ | $+$ | $+$ | $+$ | |
| $e^x - 2 - \sqrt{3}$ | | $-$ | $-$ | $-$ | $0$ | $+$ | |
| $\lambda''(x)$ | | $+$ | $0$ | $-$ | $0$ | $+$ | |

$\boxed{\text{La courbe représentative de } \lambda \text{ admet deux points d'inflexions en } \ln(2-\sqrt{3}) \text{ et } \ln(2+\sqrt{3}).}$

De plus, $\lambda$ est convexe sur $]-\infty, \ln(2-\sqrt{3})[$ et $]\ln(2+\sqrt{3}), +\infty[$ et concave sur $]\ln(2-\sqrt{3}), \ln(2+\sqrt{3})[$. Notons que comme $\lambda$ est paire, les points d'inflexion en $\ln(2-\sqrt{3}$ et $\ln(2+\sqrt{3})$ sont symétriques par rapport à l'axe des ordonnées. Et la courbe représentative de $\lambda$ traverse les tangentes en ces deux points.

4. (a) Soit $Z$ une variable aléatoire suivant la loi logistique standard. Montrons que $Z$ admet des moments à tout ordre. Soit $k \in \mathbb{N}^*$. $Z$ admet un moment d'ordre $k$ si et seulement si

$$\int_{-\infty}^{+\infty} x^k \times \frac{e^{-x}}{(1+e^{-x})^2} \,\mathrm{d}x$$

converge absolument. La fonction $x \mapsto \left| x^k \times \dfrac{e^{-x}}{(1+e^{-x})^2} \right|$ est continue et positive sur $]-\infty, +\infty[$. Comme elle est paire sur $\mathbb{R}$, on sait que :

$$\int_{-\infty}^{+\infty} x^k \frac{e^{-x}}{(1+e^{-x})^2} \,\mathrm{d}x \text{ converge absolument } \Leftrightarrow \int_0^{+\infty} \underbrace{x^k \frac{e^{-x}}{(1+e^{-x})^2}}_{\geq 0} \,\mathrm{d}x \text{ converge (absolument).}$$

Prouvons la convergence de cette dernière intégrale (généralisée en $+\infty$) par théorème de comparaison :

- Par croissance comparée au voisinage de $+\infty$ :

$$x^2 \times \left( x^k \times \frac{e^{-x}}{(1+e^{-x})^2} \right) \underset{x\to+\infty}{\sim} x^{k+2} e^{-x} \underset{x\to+\infty}{\longrightarrow} 0.$$

Ainsi, on a $x^k \times \dfrac{e^{-x}}{(1+e^{-x})^2} = o\left(\dfrac{1}{x^2}\right)$.

- $\dfrac{1}{x^2} \geq 1$ pour tout $x \geq 1$.

- $\displaystyle\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x$ converge en tant qu'intégrale de Riemann en $+\infty$ d'exposant $2>1$.

Par comparaison d'intégrales de fonctions positives, l'intégrale $\displaystyle\int_0^{+\infty} x^k \times \frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x$ converge, et donc $\displaystyle\int_{-\infty}^{+\infty} x^k \frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x$ converge absolument par parité. On en déduit que $\boxed{Z \text{ admet un moment d'ordre } k, \text{ pour tout } k\in\mathbb{N}^*.}$

Si $Y \hookrightarrow \mathscr{L}(r,s)$, alors $Z=\dfrac{Y-r}{s}$ suit une loi logistique standard et donc admet un moment d'ordre $k$ pour tout $k\in\mathbb{N}^*$. Et :

$$Y^k = (sZ+r)^k = \sum_{j=0}^{k}\binom{k}{j}s^jZ^jr^{k-j}.$$

admet une espérance comme combinaison linéaire de variables aléatoires admettant une espérance. Donc $\boxed{Y \text{ admet un moment de n'importe quel ordre.}}$

Calculons l'espérance de $Z$. La fonction $x\mapsto x\lambda(x)$ est impaire et son intégrale $\displaystyle\int_{-\infty}^{+\infty} x\lambda(x)\,\mathrm{d}x$ converge. Donc $\boxed{E(Z) \text{ existe et vaut } 0.}$ Si $Y$ suit la loi logistique $\mathscr{L}(r,s)$, $\frac{Y-r}{s}$ suit une loi logistique standard, et on a :

$$E(Y)=E\left(s\frac{Y-r}{s}+r\right)=sE\left(\frac{Y-r}{s}\right)+r\ \boxed{=r.}$$

(b) On utilise la méthode d'inversion pour simuler la loi logistique standard (qui a été présentée dans le **TP5. Simulation de variables aléatoires continues.**). Rappelons comment cela fonctionne. Soit $U\hookrightarrow\mathscr{U}(]0,1[)$. Montrons que $L(U)$ suit la loi logistique standard. En effet, si on note $F$ la fonction de répartition de $L(U)$, alors pour tout $x\in\mathbb{R}$, on a :

$$F(x)=P(L(U)\leq x) \underset{L \text{ strict. croiss.}}{=} P(U\leq \underbrace{\Lambda(x)}_{\in]0,1[})=\Lambda(x).$$

Ainsi $\Lambda(U)$ suit bien une loi logistique standard. Et pour simuler une loi logistique $\mathscr{L}(r,s)$, on utilise le fait que si $Z$ suit une loi logistique standard, on a $sZ+r\hookrightarrow\mathscr{L}(r,s)$.

On propose donc le programme suivant : on initialise une matrice `S` au format adapté `(n,p)`, puis on la remplit en utilisant deux boucles imbriquées parcourant lignes et colonnes de réalisations de la loi logistique $\mathscr{L}(r,s)$ par méthode d'inversion.

```
function S=grandlogis(n,p,r,s)
S = zeros(n,p)
for i=1:n
   for j=1:p
         U = rand()       // on simule une loi uniforme sur ]0,1[
         Z = log(U/(1-U)) // L(U)
         S(i,j) = s*Z + r
   end
end
endfunction
```

(c) Soit $(Y_1,\dots,Y_n)$ un $n$-échantillon de loi mère la loi logistique $\mathscr{L}(r,s)$. On a montré dans le TD19 - Exercice 4, que la variance empirique $S_n^2=\dfrac{1}{n-1}\displaystyle\sum_{k=1}^{n}(Y_k-\overline{Y_n})^2$ est un estimateur sans biais et convergent de la variance de $\mathscr{L}(r,s)$ (qui admet bien un moment d'ordre 4). Rappelons enfin qu'en `Scilab`, la commande `stdev(x)` calcule l'écart-type d'un vecteur `x`.

On peut donc procéder comme suit (en supposant que les variables `r` et `s` ont été définies par l'utilisateur préalablement).

```
n = 10000 // n suffisamment grand
Y = grandlogis(1,n,r,s)
var = stdev(Y)^2
disp(var)
```

5. (a) Posons $Z_1 = \ln(U_1)$ et $Z_2 = -\ln(U_2)$. On écrit :

$$Z = Z_1 + Z_2.$$

On commence par déterminer la loi puis une densité de $Z_1$ et $Z_2$. Soit pour cela $F$ la fonction de répartition de $Z_1$. On a pour tout $x \in \mathbb{R}$ :

$$F(x) = P(\ln(U_1) \leq x) \underbrace{=}_{\text{exp croiss.}} P(U_1 \leq e^x) = 1 - \exp(-e^x).$$

On en déduit que $Z_1$ est une variable à densité (car $F$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$) et admet comme densité :

$$f : x \mapsto e^x \exp(-e^x).$$

$Z_2$ est de même loi que $-Z_1$. Par transformation affine d'une variable à densité, $Z_2$ est aussi à densité, de densité :

$$g : x \mapsto \frac{1}{|-1|} f\left(\frac{x-0}{-1}\right) = f(-x) = e^{-x}\exp(-e^{-x}).$$

Montrons que $Z$ est à densité en justifiant le produit de convolution :

- On a $f$ continue sur $]-\infty,+\infty[$ et

$$\lim_{-\infty} f = \lim_{+\infty} f = 0$$

  donc $f$ est bornée.
- $Z_1 = \ln(U_1)$ et $Z_2 = -\ln(U_2)$ sont indépendantes par lemme de coalition (car $U_1$ et $U_2$ sont indépendantes).

On en déduit que $Z$ est une variable à densité, de densité (continue sur $\mathbb{R}$) :

$$h : x \mapsto f * g(x) = \int_{-\infty}^{+\infty} f(t)g(x-t)\,\mathrm{d}t.$$

Calculons maintenant $f * g$. Pour tout $x \in \mathbb{R}$, on a :

$$\begin{aligned} f * g(x) &= \int_{-\infty}^{+\infty} f(t)g(x-t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} e^t \exp(-e^t)e^{-(x-t)}\exp(-e^{-(x-t)})\,\mathrm{d}t \\ &= e^{-x}\int_{-\infty}^{+\infty} e^{2t}\exp(-e^t(1+e^{-x}))\,\mathrm{d}t. \end{aligned}$$

On effectue le changement de variable $u = e^t$ (licite car $t \mapsto e^t$ est de classe $\mathscr{C}^1$ et strictement croissante sur $\mathbb{R}$). On a $\mathrm{d}u = e^t\,\mathrm{d}t$, et $u : 0 \to +\infty$ lorsque $t : -\infty \to +\infty$. On obtient :

$$h(x) = e^{-x}\int_0^{+\infty} u\exp(-u(1+e^{-x}))du = \frac{e^{-x}}{1+e^{-x}} \times \underbrace{\int_0^{+\infty} u(1+e^{-x})\exp(-u(1+e^{-x}))du}_{=(*)}.$$

On reconnait dans $(*)$ l'espérance d'une variable aléatoire qui suit une loi de paramètre $\lambda = 1+e^{-x}$, qui vaut $\dfrac{1}{\lambda} = \dfrac{1}{1+e^{-x}}$. On en déduit que :

$$h(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \lambda(x)$$

Ainsi $Z$ admet $\lambda$ pour densité, et $\boxed{Z \text{ suit bien une loi logistique standard.}}$

(b) On en déduit le programme suivant permettant de simuler une variable aléatoire suivant la loi logistique standard.

```
U1 = grand(1,1,'exp',1)
U2 = grand(1,1,'exp',1) // on simule U1 et U2
Z = log(U1/U2)
disp(Z)
```

## Partie II. Variance de la loi logistique standard.

6. (a) On a $\boxed{P_0 = 1 \text{ et } P_1 = 3(X-1)-1 = 3X-4.}$

(b) Soit $n \in \mathbb{N}^*$. Pour tout $k \in [\![1,n]\!]$, on a $\deg((X-1)^{n-k}) < n$ et $\deg((X-1)^n) = n$ donc $\boxed{\deg(P_n) = n.}$ De plus, son coefficient dominant est :

$$(-1)^0\binom{2n+1}{1} \boxed{= 2n+1.}$$

Pour le coefficient du monôme en $X^{n-1}$, deux termes sont à considérer : le coefficient d'ordre $X^{n-1}$ dans $(X-1)^n$ et le coefficient d'ordre $n-1$ dans $(X-1)^{n-1}$. On a d'une part :

$$(X-1)^n = \sum_{j=0}^{n}\binom{n}{j}(-1)^j X^{n-j}.$$

La contribution de $(-1)^0\dbinom{2n+1}{1}(X-1)^n$ pour le monôme $X^{n-1}$ est :

$$-(2n+1)\binom{n}{1} = -n(2n+1).$$

On en déduit que le coefficient devant $X^{n-1}$ pour $P_n$ vaut :

$$(-1)\times\binom{2n+1}{3} - n(2n+1) = -\left(\frac{(2n+1)2n(2n-1)+6n(2n+1)}{6}\right) \boxed{= -\frac{2}{3}n(n+1)(2n+1).}$$

(c) On sait que $P_n$ est de degré $n$ et on connait son coefficient dominant. En posant $z_1, \dots, z_n$ ses racines (non nécessairement deux à deux distinctes), on peut écrire :

$$P_n = (2n+1)\prod_{k=1}^{n}(X-z_k)$$

Si on développe cette expression, on s'aperçoit que le coefficient devant le monôme $X^{n-1}$ vaut :

$$-(2n+1)\sum_{k=1}^{n} z_k.$$

En effet en développant, les termes de degré $n-1$ sont ceux pour lesquels on a choisi $(n-1)$ fois $X$ et une fois l'un des $-z_i$, et il y en a $n$ que l'on somme. En identifiant alors les coefficients devant le monôme $X^{n-1}$ obtenus ci-dessus et à la question précédente, on obtient :

$$-(2n+1)\sum_{k=1}^{n} z_k \;=\; -\frac{2}{3}n(n+1)(2n+1) \;\Rightarrow\; \boxed{\sum_{k=1}^{n} z_k \;=\; \frac{2}{3}n(n+1).}$$

7. (a) Tout d'abord, on a par la formule de Moivre :

$$(\cos x + i\sin x)^{2n+1} = (e^{ix})^{2n+1} = e^{i(2n+1)x} = \cos((2n+1)x) + i\sin((2n+1)x).$$

On a donc la première égalité :

$$\boxed{\operatorname{Im}\left((\cos x + i\sin x)^{2n+1}\right) = \sin\left((2n+1)x\right).}$$

Ensuite, par le binôme de Newton, on écrit :

$$(\cos x + i\sin x)^{2n+1} = \sum_{k=0}^{2n+1}\binom{2n+1}{k} i^k \sin^k(x)\cos^{2n+1-k}(x).$$

Or, on sait que (par une récurrence élémentaire si nécessaire) :

$$i^k \;=\; \begin{cases} 1 & \text{si } \exists j\in\mathbb{Z},\ k=4j \\ i & \text{si } \exists j\in\mathbb{Z},\ k=4j+1 \\ -1 & \text{si } \exists j\in\mathbb{Z},\ k=4j+2 \\ -i & \text{si } \exists j\in\mathbb{Z},\ k=4j+3 \end{cases}$$

Si on prend la partie imaginaire de la somme, seuls les indices impairs restent. On obtient donc :

$$\begin{aligned}\operatorname{Im}(\cos x + i\sin x)^{2n+1} &= \sum_{j=0}^{n}\binom{2n+1}{2j+1}(-1)^j\sin^{2j+1}(x)\cos^{2n+1-2j-1}(x)\\ &\boxed{= \sum_{j=0}^{n}\binom{2n+1}{2j+1}(-1)^j\sin^{2j+1}(x)\cos^{2(n-j)}(x).}\end{aligned}$$

(b) Soit $x\in]0,\pi[$ (ce qui assure que $\sin x\neq 0$), on a :

$$\begin{aligned}\sin^{2n+1}(x)P_n\left(\frac{1}{\sin^2(x)}\right) &= \sin^{2n+1}(x)\sum_{k=0}^{n}(-1)^k\binom{2n+1}{2k+1}\left(\frac{1}{\sin^2 x}-1\right)^{n-k}\\ &= \sin^{2n+1}(x)\sum_{k=0}^{n}(-1)^k\binom{2n+1}{2k+1}\left(\frac{1-\sin^2 x}{\sin^2 x}\right)^{n-k}\\ &= \sum_{k=0}^{n}(-1)^k\binom{2n+1}{2k+1}\cos^{2(n-k)}(x)\sin^{2n+1-2(n-k)}(x)\\ &= \sum_{k=0}^{n}(-1)^k\binom{2n+1}{2k+1}\cos^{2(n-k)}(x)\sin^{2k+1}(x)\\ &\boxed{= \sin((2n+1)x).}\end{aligned}$$

On peut donc conclure.

(c) Pour tout $k \in [\![1,n]\!]$, posons $x_k = \dfrac{k\pi}{2n+1} \in ]0,\pi[$. On a :

$$P_n\left(\frac{1}{\sin^2(x_k)}\right) = \frac{\sin((2n+1)x_k)}{\sin^{2n+1}(x_k)} = \frac{\sin(k\pi)}{\sin^{2n+1}(x_k)} = 0.$$

Ainsi, $z_k = \dfrac{1}{\sin^2(x_k)}$ est racine de $P_n$ pour tout $k \in [\![1,n]\!]$. Et les réels $z_1, \dots, z_n$ sont deux à deux distincts car les $x_1, \dots, x_n$ le sont et que la fonction $x \mapsto \dfrac{1}{\sin^2(x)}$ est injective (elle est strictement décroissante sur $]0,\pi[$ car dérivable de dérivée $x \mapsto -2\dfrac{\cos(x)}{\sin^3(x)} < 0$). Comme $P_n$ est de degré $n$, on a ainsi trouvé les $n$ racines du polynôme $P_n$. À l'aide de la question 6.(c), on a obtient donc :

$$\sum_{k=1}^{n} z_k = \frac{2}{3}n(n+1) \quad \Rightarrow \quad \boxed{\sum_{k=1}^{n} \frac{1}{\sin^2(k\pi/(2n+1))} = \frac{2}{3}n(2n+1).}$$

8. (a) On peut étudier les fonctions $x \mapsto x - \sin x$ et $x \mapsto \tan x - x$ sur $[0,\pi/2[$. On peut aussi remarquer que la fonction sin est concave sur $[0,\pi/2[$ et la fonction tan est convexe (en effet, elles sont $\mathscr{C}^2$ sur cet intervalle, et la dérivée seconde de sin (resp. tan) est la fonction $x \mapsto -\sin x$ qui est négative sur $[0,\pi/2[$ (resp. $x \mapsto 2\dfrac{\sin x}{\cos^3 x}$ qui est positive sur $[0,\pi/2[$)). La tangente en 0 à la fois pour la courbe de la fonction sinus et de la fonction tangente est la droite d'équation :

$$y = x.$$

Par propriété de la convexité, on a donc

$$\forall x \in ]0,\pi/2[, \quad \sin x \leq x \leq \tan x.$$

Ensuite, comme $\sin x > 0$ sur $]0,\pi/2[$, on peut écrire :

$$\sin x \leq x \leq \tan x \Rightarrow \sin^2 x \leq x^2 \leq \tan^2 x \Longrightarrow \frac{1}{\sin^2 x} \geq \frac{1}{x^2} \geq \frac{\cos^2 x}{\sin^2 x} = \frac{1-\sin^2 x}{\sin^2 x} = \frac{1}{\sin^2 x} - 1.$$

Ce qui est l'inégalité demandée.

(b) On remarque que pour tout $k \in [\![1,n]\!]$, on a :

$$0 < \frac{k\pi}{2n+1} \leq \underbrace{\frac{n}{2n+1}}_{<1/2}\pi < \frac{\pi}{2}.$$

En utilisant l'une des inégalités de la question 7.(c), on obtient :

$$\sum_{k=1}^{n} \frac{1}{\sin^2(k\pi/(2n+1))} \leq \sum_{k=1}^{n}\left(\frac{1}{(k\pi/(2n+1))^2} + 1\right) = \frac{(2n+1)^2}{\pi^2}\sum_{k=1}^{n}\frac{1}{k^2} + n,$$

ce qui donne encore :

$$\frac{(2n+1)^2}{\pi^2}\sum_{k=1}^{n}\frac{1}{k^2} \geq \sum_{k=1}^{n}\frac{1}{\sin^2(k\pi/(2n+1))} - n = \frac{2}{3}n^2 - \frac{1}{3}n = \frac{1}{3}n(2n-1).$$

D'autre part, par la deuxième inégalité de la question 7.(c) :

$$\frac{2}{3}n(n+1) = \sum_{k=1}^{n}\frac{1}{\sin^2(k\pi/(2n+1))} \geq \sum_{k=1}^{n}\frac{1}{(k\pi/(2n+1))^2} = \frac{(2n+1)^2}{\pi^2}\sum_{k=1}^{n}\frac{1}{k^2}.$$

On obtient donc :

$$\boxed{\frac{1}{3}n(2n-1) \leq \frac{(2n+1)^2}{\pi^2}\sum_{k=1}^{n}\frac{1}{k^2} \leq \frac{2}{3}n(n+1).}$$

(c) On a pour tout $n \in \mathbb{N}^*$ :

$$\frac{\pi^2}{3}\frac{n(2n-1)}{(2n+1)^2} \leq \sum_{k=1}^{n} \frac{1}{k^2} \leq \frac{2\pi^2}{3}\frac{n(n+1)}{(2n+1)^2}.$$

Or, on a :

$$\frac{n(2n-1)}{(2n+1)^2} \underset{n\to+\infty}{\sim} \frac{2n^2}{4n^2} \underset{n\to+\infty}{\sim} \frac{1}{2} \quad \text{et} \quad \frac{n(n+1)}{(2n+1)^2} \underset{n\to+\infty}{\sim} \frac{n^2}{4n^2} \underset{n\to+\infty}{\sim} \frac{1}{4}.$$

Par théorème des gendarmes, $\lim\limits_{n\to+\infty} \sum\limits_{k=1}^{n} \frac{1}{k^2}$ existe (et donc la série $\sum \frac{1}{k^2}$ converge), et on a :

$$\boxed{\sum_{k=1}^{+\infty} \frac{1}{k^2} = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \frac{\pi^2}{6}.}$$

9. (a) On rappelle que $E(Z) = 0$ et que $Z$ admet des moments de tout ordre, donc en particulier une variance. Par la formule de Koenig-Huygens, on a :

$$V(Z) = E(Z^2) = \int_{-\infty}^{+\infty} x^2 \frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x = 2\int_{0}^{+\infty} \underbrace{x^2 \frac{e^{-x}}{(1+e^{-x})^2}}_{\text{fct. paire}}\,\mathrm{d}x,$$

et cette dernière intégrale converge.

Soit $A > 0$. On effectue une intégration par partie sur l'intégrale $\displaystyle\int_0^A \frac{xe^{-x}}{1+e^{-x}}$.

$$\begin{array}{c|ccc c}
+ & \boxed{\dfrac{e^{-x}}{1+e^{-x}}} & & x & \\
 & & \searrow & & \boxed{\mathscr{C}^1} \\
- & -\dfrac{e^{-x}}{(1+e^{-x})^2} & \overset{\int}{\leftarrow} & \boxed{\dfrac{x^2}{2}} &
\end{array}$$

On obtient :

$$\begin{aligned}
\int_0^A x\frac{e^{-x}}{1+e^{-x}}\,\mathrm{d}x &= \left[\frac{1}{2}x^2\frac{e^{-x}}{1+e^{-x}}\right]_0^A + \frac{1}{2}\int_0^A x^2\frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x \\
&= \frac{1}{2}\frac{A^2e^{-A}}{1+e^{-A}} + \frac{1}{2}\int_0^A x^2\frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x
\end{aligned}$$

Or, par croissance comparée

$$\frac{A^2e^{-A}}{1+e^{-A}} \underset{A\to+\infty}{\sim} A^2e^{-A} \underset{A\to+\infty}{\to} 0.$$

En passant à la limite lorsque $A \to +\infty$, on obtient que l'intégrale $\displaystyle\int_0^{+\infty} x\frac{e^{-x}}{1+e^{-x}}\,\mathrm{d}x$ converge (ce qu'on savait déjà puisque $E(Z)$ existe), et on a :

$$\int_0^{+\infty} x\frac{e^{-x}}{1+e^{-x}}\,\mathrm{d}x = \frac{1}{2}\int_0^{+\infty} x^2\frac{e^{-x}}{(1+e^{-x})^2}\,\mathrm{d}x \boxed{= \frac{1}{4}V(Z).}$$

(b) Soit $A > 0$ et soit $n \in \mathbb{N}^*$. On a par linéarité de l'intégrale :

$$\begin{aligned}
\sum_{k=0}^{n}(-1)^k\int_0^A xe^{-(k+1)x}\,\mathrm{d}x &= \int_0^A xe^{-x}\left(\sum_{k=0}^{n}(-e^{-x})^k\right)\mathrm{d}x \underset{e^{-x}\neq 1}{=} \int_0^A xe^{-x}\frac{1-(-e^{-x})^{n+1}}{1+e^{-x}}\,\mathrm{d}x \\
&= \int_0^A \frac{xe^{-x}}{1+e^{-x}}\,\mathrm{d}x + (-1)^{n+1}\int_0^A \frac{xe^{-(n+2)x}}{1+e^{-x}}\,\mathrm{d}x.
\end{aligned}$$

Toutes les intégrales en jeu convergent en $+\infty$. En effet, on a :

- $\forall k \in [\![0,n]\!]$, $xe^{-(k+1)x} = o\left(\frac{1}{x^2}\right)$ et $\frac{xe^{-(n+2)x}}{1+e^{-x}} \sim xe^{-(n+2)x} = o\left(\frac{1}{x^2}\right)$ par croissance comparée.
- $\frac{1}{x^2} \geq 0$ pour tout $x \geq 1$.
- $\int_1^{+\infty} \frac{\mathrm{d}x}{x^2}$ converge en tant qu'intégrale de Riemann en $+\infty$ d'exposant $2 > 1$.

on conclut par le critère de comparaison d'intégrales de fonctions positives que tout converge en $+\infty$. On peut donc passer à la limite lorsque $A \to +\infty$, ce qui donne :

$$\boxed{\sum_{k=0}^{n}(-1)^k \int_0^{+\infty} xe^{-(k+1)x}\,\mathrm{d}x = \int_0^{+\infty} \frac{xe^{-x}}{1+e^{-x}}\,\mathrm{d}x + \underbrace{(-1)^{n+1}\int_0^{+\infty} \frac{xe^{-(n+2)x}}{1+e^{-x}}\,\mathrm{d}x}_{=I_n}.}$$

(c) On a pour tout $x \in \mathbb{R}$ :

$$0 \leq \frac{1}{1+e^{-x}} \leq 1.$$

Par croissance de l'intégrale (tout converge) :

$$0 \leq |I_n| \leq \int_0^{+\infty} xe^{-(n+2)x}\,\mathrm{d}x$$

Or, si on considère $X \hookrightarrow \mathscr{E}(n+2)$, alors on a :

$$\int_0^{+\infty} xe^{-(n+2)x}\,\mathrm{d}x = \frac{1}{n+2}\int_0^{+\infty}(n+2)xe^{-(n+2)x}\,\mathrm{d}x = \frac{E(X)}{n+2} = \frac{1}{(n+2)^2}.$$

On a donc $\lim\limits_{n\to+\infty} \int_0^{+\infty} xe^{-(n+2)x}dx = 0$, et par théorème des gendarmes $\lim\limits_{n\to+\infty} |I_n|$ existe et vaut 0. Ainsi, on a $\boxed{\lim\limits_{n\to+\infty} I_n = 0.}$

En passant à la limite quand $n \to +\infty$ dans l'égalité du 9.(c), on obtient :

$$\boxed{\int_0^{+\infty} \frac{xe^{-x}}{1+e^{-x}}\,\mathrm{d}x = \sum_{k=0}^{+\infty}(-1)^k \int_0^{+\infty} xe^{-(k+1)x}\,\mathrm{d}x = \sum_{k=0}^{+\infty}(-1)^k \frac{1}{(k+1)^2}}$$

où l'intégrale $\int_0^{+\infty} xe^{-(k+1)x}dx$ a déjà été calculée au début de cette question.

(d) On a montré que :

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

On souhaite déterminer la valeur de la somme suivante :

$$\sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)^2} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n^2}.$$

Soit $n \in \mathbb{N}^*$, et notons $T_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$. On sépare les termes pairs et impairs dans la somme $S_{2n} = \sum\limits_{k=1}^{2n} \frac{(-1)^{k-1}}{k^2}$.

$$\begin{aligned} S_{2n} &= \sum_{i=1}^{n} \frac{(-1)^{2i-1}}{(2i)^2} + \sum_{j=0}^{n-1} \frac{(-1)^{2j+1-1}}{(2j+1)^2} = -\sum_{i=1}^{n} \frac{1}{(2i)^2} + \sum_{j=0}^{n-1} \frac{1}{(2j+1)^2} \\ &= -2\sum_{i=1}^{n} \frac{1}{(2i)^2} + \underbrace{\sum_{i=1}^{n} \frac{1}{(2i)^2} + \sum_{j=0}^{n-1} \frac{1}{(2j+1)^2}}_{=T_{2n}} = -\frac{2}{4}\sum_{i=1}^{n} \frac{1}{i^2} + T_{2n} = T_{2n} - \frac{1}{2}T_n \end{aligned}$$

Or, on a montré à la question 8.(c) que $\lim_{n\to+\infty} T_n = \frac{\pi^2}{6}$. En tant que suite extraite d'une suite convergente, on a $\lim_{n\to+\infty} T_{2n} = \frac{\pi^2}{6}$. Par somme de suites convergentes, $\lim_{n\to+\infty} S_{2n}$ existe et on a :

$$\lim_{n\to+\infty} S_{2n} = \frac{\pi^2}{6} - \frac{\pi^2}{12} = \frac{\pi^2}{12}.$$

Or la suite de terme général $S_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k^2}$ converge d'après la question 9.(c), et sa limite vaut $\frac{V(Z)}{4}$ d'après la question 9.(a). En tant que suite extraite d'une suite convergente, $(S_{2n})$ converge donc aussi vers $\frac{V(Z)}{4}$. Et par unicité de la limite, on en déduit que :

$$\frac{V(Z)}{4} = \frac{\pi^2}{12} \quad \Rightarrow \quad \boxed{V(Z) = \frac{\pi^2}{3}.}$$

10. (a) Les fonctions $g_1 : x \mapsto |\ln(x)|e^{-x}$ et $g_2 : x \mapsto (\ln x)^2 e^{-x}$ sont continues et positives sur $]0, +\infty[$. Les intégrales $I' = \int_0^{+\infty} g_1(x)\,\mathrm{d}x$ et $J = \int_0^{+\infty} g_2(x)\,\mathrm{d}x$ sont donc généralisées en 0 et en $+\infty$.

En 0, on a :

- $\sqrt{x}|\ln(x)|e^{-x}x \underset{\sim}{\to} 0\sqrt{x}|\ln(x)| \underset{x\to0}{\longrightarrow} 0$ par croissances comparées, de sorte que $g_1(x) \underset{x\to0}{=} o\left(\frac{1}{\sqrt{x}}\right)$. De même, on a $g_2(x) \underset{x\to0}{=} o\left(\frac{1}{\sqrt{x}}\right)$.
- $\frac{1}{\sqrt{x}} \geq 0$ pour tout $x \in ]0, 1]$.
- $\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x$ converge en tant qu'intégrale de Riemann en 0 d'exposant $1/2 < 1$.

Par théorème de comparaison, les intégrales $\int_0^1 g_1(x)\,\mathrm{d}x$ et $\int_0^1 g_2(x)\,\mathrm{d}x$ convergent.

En $+\infty$, on a :

- $x^2|\ln(x)|e^{-x} = \frac{\ln(x)}{x} \times x^3 e^{-x} \underset{x\to+\infty}{\longrightarrow} 0$ par croissances comparées. On a donc $g_1(x) \underset{x\to+\infty}{=} o\left(\frac{1}{x^2}\right)$. De même, on a $g_2(x) \underset{x\to+\infty}{=} o\left(\frac{1}{x^2}\right)$.
- $\frac{1}{x^2} \geq 0$ pour tout $x \geq 1$.
- $\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x$ converge en tant qu'intégrale de Riemann en $+\infty$ d'exposant $2 > 1$.

Par théorème de comparaison, les intégrales $\int_1^{+\infty} g_1(x)\,\mathrm{d}x$ et $\int_1^{+\infty} g_2(x)\,\mathrm{d}x$ convergent.

Par Chasles, les intégrales $I'$ et $J$ convergent. Ainsi $\boxed{I \text{ et } J \text{ convergent absolument, donc convergent.}}$

(b) Soit $U \hookrightarrow \mathscr{E}(1)$. Les intégrales $I$ et $J$ convergeant absolument, on en déduit par théorème de transfert que $E(\ln(U)$ et $E(\ln(U)^2)$ existent et valent :

$$E(\ln(U)) = \int_0^{+\infty} \ln(x)e^{-x}\,\mathrm{d}x = I \quad \text{et} \quad E(\ln(U)^2) = \int_0^{+\infty} (\ln x)^2 e^{-x} dx = J.$$

Par la formule de Koenig Huygens, on a : $J - I^2 = E(\ln(U)^2) - E(\ln(U))^2 = V(\ln(U))$.
Or, par la question 5.(a), on sait que si $U_1$ et $U_2$ sont deux variables indépendantes de loi exponentielle de paramètre 1, alors :

$$Z = \ln(U_1) - \ln(U_2)$$

suit une loi logistique standard. Ainsi, par indépendance (résultant du lemme de coalition), on a :

$$V(Z) = V(\ln(U_1)) + V(-\ln(U_2)) = V(\ln(U_1)) + V(\ln(U_2)) = 2(J - I^2)$$

Comme de plus $V(Z) = \dfrac{\pi^2}{3}$ par la question 9.(d), on a $\boxed{J - I^2 = \dfrac{\pi^2}{6}.}$

## Partie III. Estimation à partir de données binaires.

11. Comme $f$ est continue, $F$ est donc de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et admet pour dérivée $f$. Comme $f$ est strictement positive, $F$ est strictement croissante. De plus par propriété d'une fonction de répartition, on a :

$$\lim_{-\infty} F = 0 \text{ et } \lim_{+\infty} F = 1.$$

Comme $F$ est continue et strictement croissante, $\boxed{\text{elle réalise une bijection de } \mathbb{R} \text{ dans } ]0,1[.}$

12. Les variables aléatoires $(Y_1, \ldots, Y_n)$ sont mutuellement indépendantes, de même loi de Bernoulli de paramètre $F(\theta)$, qui admet une espérance égale à $F(\theta)$ et une variance égale à $F(\theta)(1-F(\theta))$. Par le théorème limite central, on sait que :

$$\frac{\overline{Y_n} - E(\overline{Y_n})}{\sqrt{V(\overline{Y_n})}} \xrightarrow{\mathscr{L}} Z \quad \text{où} \quad Z \hookrightarrow \mathscr{N}(0,1).$$

Or par linéarité de l'espérance, on a $E(\overline{Y_n}) = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} E(Y_j) = F(\theta)$, et par indépendance :

$$V(\overline{Y_n}) = \frac{1}{n^2}\sum_{k=1}^{n} V(Y_k) = \frac{F(\theta)(1-F(\theta))}{n}.$$

On a donc :

$$\sqrt{n}\frac{\overline{Y_n} - F(\theta)}{\sqrt{F(\theta)(1-F(\theta))}} \xrightarrow{\mathscr{L}} Z \quad \Longrightarrow \quad \sqrt{n}(\overline{Y_n} - F(\theta)) \xrightarrow{\mathscr{L}} \sqrt{F(\theta)(1-F(\theta))}Z.$$

Donc $\boxed{\sqrt{n}(\overline{Y_n} - F(\theta)) \text{ converge en loi vers la variable } \sqrt{F(\theta)(1-F(\theta))}Z \hookrightarrow \mathscr{N}(0, F(\theta)(1-F(\theta))).}$

13. (a) Les variables $Y_1, \ldots, Y_n$ sont à valeurs dans $\{0,1\}$ donc $\overline{Y_n}$ est à valeurs dans $\{0, 1/n, 2/n, \ldots, (n-1)/n, 1\}$. Ainsi, on a :

$$P_\theta(E_n) = 1 - P_\theta(\overline{Y_n} = 0) - P_\theta(\overline{Y_n} = 1) = 1 - P_\theta(Y_1 = 0, \ldots, Y_n = 0) - P_\theta(Y_1 = 1, \ldots, Y_n = 1).$$

Or par indépendance des $(Y_i)$, et puisque les $Y_i$ suivent tous une loi de Bernoulli de paramètre $F(\theta)$, on a :

$$P_\theta(E_n) = 1 - P_\theta(Y_1 = 0)^n - P_\theta(Y_1 = 1)^n \boxed{= 1 - (1 - F(\theta))^n - F(\theta)^n.}$$

Comme $F$ est à valeurs dans $]0,1[$, on a $F(\theta) \in ]0,1[$ et $(1 - F(\theta)) \in ]0,1[$, de sorte que :

$$\boxed{\lim_{n \to +\infty} P_\theta(E_n) = 1.}$$

(b) i. Commençons par remarquer que :

$$\{\overline{Y_n} \leq F(x)\} \cap E_n = \{\omega \in E_n, \overline{Y_n}(\omega) \leq F(x)\}.$$

Si $\omega \in E_n$, alors $T_n(\omega) = F^{-1}(\overline{Y_n}(\omega))$ et donc $T_n(\omega) \leq x \Leftrightarrow F^{-1}(\overline{Y_n}(\omega)) \leq x \Leftrightarrow \overline{Y_n}(\omega) \leq F(x)$ par stricte croissance de $F$. On obtient donc :

$$\boxed{\{\omega \in E_n,\, T_n(\omega) \leq x\} = \{\omega \in E_n,\, \overline{Y_n}(\omega) \leq F(x)\}.}$$

Montrons maintenant que pour tout $x \in \mathbb{R}$, $[T_n \leq x] \in \mathscr{A}$. Tout d'abord comme $\overline{Y_n}$ est une variable aléatoire (en tant que somme de variables aléatoires), les ensembles que $E_n = [0 < Y_n < 1]$ et $[\overline{Y_n} \leq F(x)]$ sont éléments de la tribu $\mathscr{A}$. De plus, on a :

$$[T_n \leq x] = ([T_n \leq x] \cap E_n) \cup \left([T_n \leq x] \cap \overline{E_n}\right).$$

Or si $\omega \in \overline{E_n}$, alors $T_n(\omega) = 0$ et on a :

$$[T_n \leq x] \cap \overline{E_n} = \begin{cases} \emptyset & \text{si } x < 0, \\ \overline{E_n} & \text{si } x \geq 0. \end{cases}$$

Dans tous les cas, $[T_n \leq x] \cap \overline{E_n}$ appartient à $\mathscr{A}$. D'autre part, $[T_n \leq x] \cap E_n = [\overline{Y_n} \leq F(x)] \cap E_n$ appartient bien à la tribu $\mathscr{A}$ (stabilité par intersection finie ou dénombrable), et $\boxed{[T_n \leq x] \in \mathscr{A}}$ aussi (stabilité par union finie ou dénombrable).

ii. D'une part, on a :

$$[\overline{Y_n} \leq F(x)] \cap E_n = [T_n \leq x] \cap E_n \subset [T_n \leq x].$$

Par croissance de $P_\theta$, on a :

$$\boxed{P_\theta\left([\overline{Y_n} \leq F(x)] \cap E_n\right) \leq P_\theta(T_n \leq x).}$$

D'autre part, par la formule des probabilités totales avec le SCE $(E_n, \overline{E_n})$, on a :

$$P_\theta(T_n \leq x) = P_\theta([T_n \leq x] \cap E_n) + P_\theta(\underbrace{[T_n \leq x] \cap \overline{E_n}}_{\subset \overline{E}_n}) \leq P_\theta([T_n \leq x] \cap E_n) + P_\theta(\overline{E}_n)$$
$$\boxed{\leq P_\theta([T_n \leq x] \cap E_n) + 1 - P_\theta(E_n).}$$

(c) On va utiliser pour cela l'encadrement de la question précédente. Puisque $\lim\limits_{n \to +\infty} 1 - P_\theta(E_n) = 0$, $P_\theta(T_n \leq x)$ admet la même limite (si elle existe) que $P_\theta([\overline{Y_n} \leq F(x)] \cap E_n)$ par théorème des gendarmes. On va donc s'intéresser à la convergence de $P_\theta([\overline{Y_n} \leq F(x)] \cap E_n)$.

Rappelons pour commencer que puisque les $(Y_i)_i$ sont i.i.d. et admettent une variance, on sait par la loi faible des grands nombres que :

$$\overline{Y_n} \overset{P_\theta}{\to} F(\theta).$$

Autrement dit, pour tout $\varepsilon > 0$, on a :

$$\lim_{n \to +\infty} P_\theta(|\overline{Y_n} - F(\theta)| \geq \varepsilon) = 0.$$

Soit $x \neq \theta$ fixé. On a deux cas à considérer :

- $x < \theta$. Alors $F(x) < F(\theta)$ par croissance stricte de $F$, et on a :

$$P_\theta(\overline{Y_n} \leq F(x)) = P_\theta\left(\overline{Y_n} - F(\theta) \leq -(F(\theta) - F(x))\right) = P_\theta(F(\theta) - \overline{Y_n} \geq F(\theta) - F(x))$$
$$\leq P_\theta(|\overline{Y_n} - F(\theta)| \geq \underbrace{F(\theta) - F(x)}_{>0}) \underset{n \to +\infty}{\to} 0 \text{ par la LFGN.}$$

Comme $P_\theta(\overline{Y_n} \leq F(x)) \geq 0$, on en déduit par encadrement que $\lim\limits_{n \to +\infty} P_\theta(\overline{Y_n} \leq F(x)) = 0$. Et puisque :

$$0 \leq P_\theta([\overline{Y_n} \leq F(x)] \cap E_n) \leq P_\theta(\overline{Y_n} \leq x),$$

on obtient $\lim\limits_{n \to +\infty} P_\theta([\overline{Y_n} \leq F(x)] \cap E_n) = 0$.

- $x > \theta$. Dans ce cas, on a $F(x) - F(\theta) > 0$ et :

$$P_\theta(\overline{Y_n} \leq F(x)) = P_\theta(\overline{Y_n} - F(\theta) \leq F(x) - F(\theta)) = 1 - P_\theta(\overline{Y_n} - F(\theta) > F(x) - F(\theta))$$
$$\geq 1 - P_\theta(|\overline{Y_n} - F(\theta)| \geq F(x) - F(\theta)) \underset{n\to+\infty}{\longrightarrow} 1.$$

Par théorème d'encadrement, on en déduit que $\lim\limits_{n\to+\infty} P_\theta(\overline{Y_n} \leq F(x)) = 1$. Or on a :

$$P_\theta([\overline{Y_n} \leq F(x)] \cap E_n) = P_\theta(\overline{Y_n} \leq F(x)) + P_\theta(E_n) - P_\theta([\overline{Y_n} \leq F(x)] \cup E_n)$$

Puisque $E_n \subset [\overline{Y_n} \leq F(x)] \cup E_n$ et $\lim P_\theta(E_n) = 1$, on en déduit que $\lim P_\theta([\overline{Y_n} \leq F(x)] \cup E_n) = 1$ et donc :

$$\lim_{n\to+\infty} P_\theta([\overline{Y_n} \leq F(x)] \cap E_n) = 1 + 1 - 1 = 1.$$

Avec la remarque faite au début de cette question, on obtient donc :

$$\boxed{\lim_{n\to+\infty} P_\theta(T_n \leq x) = \lim_{n\to+\infty} P_\theta([\overline{Y_n} \leq F(x)] \cap E_n) = \begin{cases} 0 & \text{si } x < \theta, \\ 1 & \text{si } x > \theta. \end{cases}}$$

(d) Tout d'abord, $T_n$ est bien une variable aléatoire définie à partir de l'échantillon $(Y_1, \ldots, Y_n)$ et qui ne dépend pas du paramètre inconnu $\theta$ : c'est donc bien un estimateur de $\theta$.

Soit $\varepsilon > 0$. On a :

$$P_\theta(|T_n - \theta| \geq \varepsilon) = 1 - P_\theta(|T_n - \theta| < \varepsilon) = 1 - P_\theta(\theta - \varepsilon < T_n < \theta + \varepsilon)$$
$$= 1 - P_\theta(T_n < \theta + \varepsilon) + P_\theta(T_n \leq \theta - \varepsilon). \qquad (*)$$

Comme $\varepsilon > 0$, on a $\theta - \varepsilon < \theta$, et donc par la question précédente :

$$\lim_{n\to+\infty} P_\theta(T_n \leq \theta - \varepsilon) = 0.$$

D'autre part, on a $[T_n \leq \theta + \varepsilon/2] \subset [T_n < \theta + \varepsilon]$, d'où par croissance de la probabilité :

$$P_\theta(T_n < \theta + \varepsilon) \geq P_\theta(T_n \leq \theta + \varepsilon/2).$$

Or on a $P_\theta(T_n \leq \theta + \varepsilon/2) \underset{n\to+\infty}{\longrightarrow} 1$ par la question précédente (puisque $\theta + \varepsilon/2 > \theta$). Par théorème des gendarmes, $\lim\limits_{n\to+\infty} P_\theta(T_n < \theta + \varepsilon)$ existe et vaut 1.
En reprenant $(*)$, on obtient donc :

$$\lim_{n\to+\infty} P_\theta(|T_n - \theta| \geq \varepsilon) = 0.$$

Ainsi $\boxed{(T_n)_{n\in\mathbb{N}^*} \text{ est une suite d'estimateurs convergente de } \theta.}$

**Pour aller plus loin.**

Le résultat obtenu à la question 13.(c) montre que $(T_n)$ converge en loi vers la variable constante égale à $\theta$. On a montré à la question 13.(d) que $(T_n)$ converge aussi en probabilité vers $\theta$. On a en fait le résultat suivant (que l'on vient ici en partie de démontrer et qui est hors programme) :

> Une suite $(Y_n)$ converge en probabilité vers une constante $c$ si et seulement si $(Y_n)$ converge en loi vers $c$.

14. (a) i. La fonction $f$ est continue et strictement positive (donc ne s'annule pas). La fonction $x \mapsto \dfrac{1}{f(x)}$ est donc continue sur $\mathbb{R}$. Par définition de la continuité en $\theta$, on sait que pour tout $\varepsilon > 0$ (en particulier pour celui qu'on a fixé), on a :

$$\boxed{\exists \alpha > 0, \quad \forall x \in [\theta - \alpha, \theta + \alpha], \quad \left| \frac{1}{f(x)} - \frac{1}{f(\theta)} \right| \leq \varepsilon.}$$

ii. Soit $\omega \in [|T_n - \theta| \leq \alpha] \cap E_n$. Si $\overline{Y_n}(\omega) = F(\theta)$, alors $U_n(\omega) = \dfrac{1}{f(\theta)}$ et on a bien $\omega \in B_n(\varepsilon)$ dans ce cas.

Supposons à présent que $\overline{Y_n}(\omega) \neq F(\theta)$. Puisque $\omega \in E_n$, on a $T_n(\omega) = F^{-1}(\overline{Y_n}(\omega))$. On a donc :

$$U_n(\omega) = \frac{T_n(\omega) - \theta}{\overline{Y_n}(\omega) - F(\theta)} = \frac{T_n(\omega) - \theta}{F(T_n(\omega)) - F(\theta)} = \left( \frac{F(T_n(\omega)) - F(\theta)}{T_n(\omega) - \theta} \right)^{-1}.$$

On voit apparaître un taux d'accroissement de la fonction $F$. Pour tout $x \in [\theta - \alpha, \theta + \alpha]$ avec $x \neq \theta$, on a par le théorème des accroissements finis (appliqué entre $\theta$ et $x$) qu'il existe $c_x \in [\theta - \alpha, \theta + \alpha]$ tel que :

$$\frac{F(x) - F(\theta)}{x - \theta} = f(c_x).$$

En particulier comme $c_x \in [\theta - \alpha, \theta + \alpha]$, on a par la question 14.(a).i :

$$\left| \frac{x - \theta}{F(x) - F(\theta)} - \frac{1}{f(\theta)} \right| = \left| \frac{1}{f(c_x)} - \frac{1}{f(\theta)} \right| \leq \varepsilon.$$

Avec $x = T_n(\omega)$, on a bien $T_n(\omega) \in [\theta - \alpha, \theta + \alpha]$ (car $\omega \in [|T_n - \theta| \leq \alpha]$) et $\overline{Y_n}(\omega) = F(T_n(\omega)) \neq F(\theta) \Rightarrow T_n(\omega) \neq \theta$. D'où :

$$\left| U_n(\omega) - \frac{1}{f(\theta)} \right| = \left| \frac{T_n(\omega) - \theta}{\overline{Y_n}(\omega) - F(\theta)} - \frac{1}{f(\theta)} \right| \leq \varepsilon,$$

et donc $\omega \in B_n(\varepsilon)$. On a donc montré l'inclusion :

$$\boxed{[|T_n - \theta| \leq \alpha] \cap E_n \subset B_n(\varepsilon).}$$

(b) Soit $\varepsilon > 0$. À l'aide de la question précédente, on a (par croissance de la probabilité) :

$$P\left( \left| U_n - \frac{1}{f(\theta)} \right| \right) > \varepsilon) = 1 - P\left( \left| U_n - \frac{1}{f(\theta)} \right| \right) \leq \varepsilon) \leq 1 - P\left( [|T_n - \theta| \leq \alpha] \cap E_n \right).$$

Or on a :

$$P\left( [|T_n - \theta| \leq \alpha] \cap E_n \right) = P(|T_n - \theta| \leq \alpha) + P(E_n) - P\big([|T_n - \theta| \leq \alpha] \cup E_n\big).$$

On a :

$$\lim_{n \to +\infty} P(|T_n - \theta| \leq \alpha) = \lim_{n \to +\infty} (1 - P(|T_n - \theta| > \alpha)) = 1 \quad \text{car} \quad T_n \xrightarrow{P} \theta.$$

et

$$\lim_{n \to +\infty} P(E_n) = 1 \quad \text{et} \quad \lim_{n \to +\infty} P([|T_n - \theta| \leq \alpha] \cup E_n) = 1 \quad \text{car} \quad P([|T_n - \theta| \leq \alpha] \cup E_n) \geq P(E_n).$$

On obtient donc :

$$\lim_{n \to +\infty} P\big([|T_n - \theta| \leq \alpha] \cap E_n\big) = 1 + 1 - 1 = 1.$$

Par théorème des gendarmes, on en déduit que :

$$\lim_{n\to+\infty} P\left(\left|U_n - \frac{1}{f(\theta)}\right| > \varepsilon\right) = 0.$$

Ainsi $\boxed{(U_n) \text{ converge en probabilité vers } \frac{1}{f(\theta)}.}$

**Remarque.** Rappelons que pour montrer une convergence en probabilité, on peut regarder la limite de $P\left(\left|U_n - \frac{1}{f(\theta)}\right| > \varepsilon\right)$ ou de $P\left(\left|U_n - \frac{1}{f(\theta)}\right| \geq \varepsilon\right)$, avec une inégalité stricte ou large de manière indifférente.

(c) On écrit :

$$\sqrt{n}(T_n - \theta) = U_n \times \sqrt{n}(\overline{Y_n} - F(\theta)).$$

On a montré que $\sqrt{n}(\overline{Y_n} - F(\theta)) \xrightarrow{\mathscr{L}} Z$ où $Z \hookrightarrow \mathscr{N}(0, F(\theta)(1-F(\theta)))$ et $U_n \xrightarrow{P} \frac{1}{f(\theta)}$. Par le théorème de Slutsky, on obtient donc :

$$\sqrt{n}(T_n - \theta) \xrightarrow{\mathscr{L}} \frac{1}{f(\theta)} Z.$$

Ainsi $\boxed{\sqrt{n}(T_n - \theta) \text{ converge en loi vers } \frac{1}{f(\theta)} Z \hookrightarrow \mathscr{N}\left(0, \frac{F(\theta)(1-F(\theta))}{f(\theta)^2}\right).}$

# Partie IV. Régression logistique.

15. (a) Commençons par noter que puisque $M \in \mathscr{M}_{p,k}(\mathbb{R})$, on a $M^tM \in \mathscr{M}_p(\mathbb{R})$.

Pour montrer que $M^tM$ est inversible, on va montrer que l'équation $M^tMX = 0_p$ d'inconnue $X \in \mathscr{M}_{p,1}(\mathbb{R})$ admet pour unique solution $X = 0_{p,1}$. Soit donc $X \in \mathscr{M}_{p,1}(\mathbb{R})$ tel que $M^tMX = 0_{p,1}$. On a :

$$\begin{aligned} M^tMX = 0_{p,1} \quad &\Rightarrow \quad {}^tXM^tMX = 0_{1,1} \quad \Rightarrow \quad {}^t({}^tMX)^tMX = 0_{1,1} \\ &\Rightarrow \quad \|{}^tMX\|^2 = 0 \quad \Rightarrow \quad {}^tMX = 0_{k,1}. \end{aligned}$$

Or l'application $X \in \mathscr{M}_{p,1}(\mathbb{R}) \mapsto^t MX \in \mathscr{M}_{k,1}(\mathbb{R})$ est injective. En effet, la matrice ${}^tM$ est de rang $p$ (on rappelle que ${}^tM$ et $M$ ont même rang) et par le théorème du rang, on a :

$$\dim \operatorname{Ker}({}^tM) = \dim \mathscr{M}_{p,1}(\mathbb{R}) - \operatorname{rg}({}^tM) = 0.$$

On en déduit que $X = 0_{p,1}$. Ainsi $\boxed{M^tM \text{ est bien inversible.}}$

(b) Soit $H \in \mathscr{M}_{k,1}(\mathbb{R})$. On cherche à minimiser la quantité suivante, dépendant de $U \in \mathscr{M}_{p,1}(\mathbb{R})$ :

$${}^t({}^tMU - H)({}^tMU - H) = \|{}^tMU - H\|^2$$

Posons pour cela $F = \operatorname{Im}({}^tM) = \{V \in \mathscr{M}_{k,1}(\mathbb{R}),\ \exists U \in \mathscr{M}_{p,1}(\mathbb{R}),\ {}^tMU = V\}$. Par le théorème de projection orthogonale, on sait que la fonction :

$$\Phi : \begin{array}{ccc} F & \to & \mathbb{R} \\ V & \mapsto & \|V - H\|^2 \end{array}$$

admet un minimum global atteint en un unique point $V = p_F(H)$, projeté orthogonal de $H$ sur $F$. Et par injectivité de $X \mapsto {}^tMX$, on en déduit qu'il existe un unique vecteur $U \in \mathscr{M}_{p,1}(\mathbb{R})$ tel que ${}^tMU = p_F(H)$.

**Remarque.** On vient ici de retrouver le résultat obtenu à la fin du **Chapitre 16. Projection orthogonale.** sur les pseudo-solutions d'un système linéaire. On va ici aller plus loin :

nous allons donner l'expression de cette pseudo-solution $U$ en fonction de $M$ et $H$. Cette expression a également été obtenu dans le **TP 7. Statistiques descriptives bivariées.** lorsqu'on a déterminé l'équation de la droite des moindres carrés.

Par propriété de la projection orthogonale, on sait que pour tout $V \in F$, on a

$$H - p_F(H) = H - {}^tMU \perp V.$$

En particulier pour tout $W \in \mathscr{M}_{p,1}(\mathbb{R})$, on a ${}^tMW \in F$, de sorte que :

$$\begin{aligned} H - {}^tMU \perp {}^tMW &\Rightarrow {}^t({}^tMW)(H - {}^tMU) = 0 \Rightarrow {}^tWM(H - {}^tMU) = 0 \\ &\Rightarrow {}^tW \cdot (MH - M^tMU) = 0. \end{aligned}$$

En particulier pour $W = (MH - M^tMU)$, on obtient :

$$\|MH - M^tMU\|^2 = 0 \;\Rightarrow\; MH = M^tMU \underbrace{\Rightarrow}_{M^tM \text{ inv.}} \boxed{U = (M^tM)^{-1}MH.}$$

(c) Supposons que l'on connaisse les lois des variables $Y_{i,n}$. Cela signifie que l'on connait la valeur de $b(x^{(1)}), \dots, b(x^{(n)})$ et donc, par bijectivité de la fonction $\Lambda$, les valeurs de $\langle \alpha, x^{(1)} \rangle, \dots, \langle \alpha, x^{(k)} \rangle$. Autrement dit, il existe $\beta_1, \dots, \beta_k$ des réels connus tels que :

$$\begin{cases} \langle \alpha, x^{(1)} \rangle &= \beta_1 \\ \langle \alpha, x^{(2)} \rangle &= \beta_2 \\ \quad\vdots & \\ \langle \alpha, x^{(k)} \rangle &= \beta_k \end{cases} \quad \Leftrightarrow \quad {}^tMA = B \;\text{ où }\; B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}.$$

Dans le cas où $M$ n'est pas de rang $p$ (et donc où $X \mapsto {}^tMX$ n'est pas injective), l'équation matricielle ${}^tMA = B$ d'inconnu $A$ admet potentiellement plusieurs solutions. $\boxed{\text{On ne peut pas déterminer } \alpha.}$

16. (a) On rappelle que $L$ est la bijection réciproque de la fonction de répartition $\Lambda$. Ainsi, d'après la partie III, question 13, on sait que :

$$\forall i \in [\![1, k]\!], \quad T_{i,n} \xrightarrow{P} \langle \alpha, x^{(i)} \rangle.$$

En effet, ici les $Y_{i,k}$ sont des variables de Bernoulli de paramètre $b(x^{(i)}) = \Lambda(\langle \alpha, x^{(i)} \rangle)$, et on a bien le résultat voulu en prenant $\theta = \langle \alpha, x^{(i)} \rangle$.

On en déduit que (en composant par la fonction continue $x \mapsto c_i x$) :

$$\forall i \in [\![1, k]\!], \quad c_iT_{i,n} \xrightarrow{P} c_i\langle \alpha, x^{(i)} \rangle$$

Maintenant il nous faut utiliser le résultat hors programme suivant (énoncé et démontré dans l'exercice 17.7) :

> **Lemme.**
> Soit $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires et $X$ et $Y$ deux variables aléatoires, toutes définies sur un même espace probabilisé. On suppose que $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$.
>
> Alors on a $X_n + Y_n \xrightarrow{P} X + Y$.

Démontrons ce résultat. Soit pour cela $\varepsilon > 0$ fixé. Commençons par établir l'inclusion suivante :

$$[|X_n + Y_n - (X+Y)| \geq \varepsilon] \subset [|X_n - X| \geq \frac{\varepsilon}{2}] \cup [|Y_n - Y| \geq \frac{\varepsilon}{2}].$$

Soit $\omega \in \Omega$ tel que :

$$|X_n(\omega) + Y_n(\omega) - (X(\omega) + Y(\omega))| \geq \varepsilon. \qquad (*)$$

Par l'inégalité triangulaire, on a :

$$\varepsilon \leq |X_n(\omega) - X(\omega)| + |Y_n(\omega) - Y(\omega)|.$$

Supposons que l'on ait $|X_n(\omega) - X(\omega)| < \dfrac{\varepsilon}{2}$ et $|Y_n(\omega) - Y(\omega)| < \dfrac{\varepsilon}{2}$. On aurait alors par somme :

$$|X_n(\omega) - X(\omega)| + |Y_n(\omega) - Y(\omega)| < \varepsilon$$

ce qui est contradictoire avec l'inégalité $(*)$. Ainsi on a $|X_n(\omega) - X(\omega)| \geq \dfrac{\varepsilon}{2}$ ou $|Y_n(\omega) - Y(\omega)| \geq \dfrac{\varepsilon}{2}$. D'où l'inclusion d'évènements :

$$[|X_n + Y_n - (X+Y)| \geq \varepsilon] \subset \left[|X_n - X| \geq \frac{\varepsilon}{2}\right] \cup \left[|Y_n - Y| \geq \frac{\varepsilon}{2}\right].$$

Maintenant par croissance d'une probabilité, on a :

$$\begin{aligned} P\left(|X_n + Y_n - (X+Y)| \geq \varepsilon\right) &\leq P\left(\left[|X_n - X| \geq \frac{\varepsilon}{2}\right] \cup \left[|Y_n - Y| \geq \frac{\varepsilon}{2}\right]\right) \\ &\leq P\left(|X_n - X| \geq \frac{\varepsilon}{2}\right) + P\left(|Y_n - Y| \geq \frac{\varepsilon}{2}\right). \end{aligned}$$

Puisque $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$, on a :

$$\lim_{n\to+\infty} P\left(|X_n - X| \geq \frac{\varepsilon}{2}\right) = 0 = \lim_{n\to+\infty} P\left(|Y_n - Y| \geq \frac{\varepsilon}{2}\right).$$

Par théorème des gendarmes (une probabilité étant positive), on obtient que $\lim\limits_{n\to+\infty} P(|X_n + Y_n - (X+Y)| \geq \varepsilon)$ existe et vaut 0. Ceci étant vrai pour tout $\varepsilon > 0$, on a donc montré que $X_n + Y_n \xrightarrow{P} X + Y$.

Par une récurrence immédiate, on peut généraliser ce résultat comme suit.

> **Lemme.**
> Soit $k \in \mathbb{N}^*$ et $(X_n^{(1)})_{n\in\mathbb{N}}, \dots, (X_n^{(k)})_{n\in\mathbb{N}}$ des suites de variables aléatoires et $X^{(1)}, \dots, X^{(k)}$ des variables aléatoires toutes définies sur un même espace probabilisé. On suppose que $X_n^{(i)} \xrightarrow{P} X^{(i)}$ pour tout $i \in [\![1, k]\!]$.
>
> Alors on a $X_n^{(1)} + \cdots + X_n^{(k)} \xrightarrow{P} X^{(1)} + \cdots + X^{(k)}$.

Ce dernier lemme permet de conclure que :

$$\boxed{\sum_{i=1}^{k} c_i T_{i,n} \xrightarrow{P} \sum_{i=1}^{k} c_i \langle \alpha, x^{(i)} \rangle.}$$

(b) Notons :

$$(M^tM)^{-1}M = C = (c_{i,j})_{\substack{1\leq i\leq p\\ 1\leq j\leq k}}$$

Soit $i \in [\![1,p]\!]$. Alors on a par définition de $A_{i,n}$ que :

$$A_{i,n} = [CH_n]_{i,1} = \sum_{j=1}^{k} c_{i,j}T_{j,n}$$

D'après la question précédente, on a :

$$A_{i,n} \xrightarrow{P} A_i = \sum_{j=1}^{k} c_{i,j}\langle \alpha, x^{(j)}\rangle.$$

Notons $\tilde{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_p \end{pmatrix}$. Avec $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}$, on a ${}^tMA = \begin{pmatrix} \displaystyle\sum_{j=1}^{p} \alpha_j x_j^{(1)} \\ \vdots \\ \displaystyle\sum_{j=1}^{p} \alpha_j x_j^{(k)} \end{pmatrix} = \begin{pmatrix} \langle \alpha, x^{(1)}\rangle \\ \vdots \\ \langle \alpha, x^{(k)}\rangle \end{pmatrix}$. On en déduit que :

$$\tilde{A} = C({}^tMA) = (M^tM)^{-1}M({}^tMA) = (M^tM)^{-1}(M^tM)A = A.$$

Ainsi $\tilde{A} = A$ et on peut conclure que :

$$\boxed{\forall i \in [\![1,p]\!], \quad A_{i,n} \xrightarrow{P} \alpha_i}$$