

Correction du devoir maison

Partie I. Fonction logistique et lors logistiques.

1. (a) La fonction Λ est de classe \mathcal{C}^1 sur \mathbb{R} en tant que quotient de fonctions de classe \mathcal{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 Λ est strictement croissante sur \mathbb{R} . Enfin, on a :

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

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

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

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

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

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

- (b) D'après le calcul précédent :

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

- (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 :

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

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

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

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

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

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

Or g est de classe \mathcal{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 :

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

2. (a) On utilise l'algorithme de dichotomie avec la fonction g définie précédemment pour estimer la valeur de x_0 tel que $g(x_0) = 0$.

Puisque $g(0) = 1/2 > 0$ et $g(1) = -1/(e + 1) < 0$, x_0 appartient à $]0, 1[$. 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 $\text{Lambda}(c) > c$ équivaut à $g(c) > 0$. Dans ce cas, g étant décroissante, x_0 appartient à l'intervalle $]c, b[$. Sinon, x_0 appartient à $]a, c[$. On propose donc :

```

9  | if Lambda(c) > c :
10 |     a = c
11 | else :
12 |     b = c

```

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

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

Pour être assuré 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 6 est $\boxed{2 \times 10^{-4}}$.

- (c) La variable `x0` du programme Python n'est pas x_0 , mais une approximation de ce réel à $\frac{\varepsilon}{2}$ près. La valeur renvoyée par l'instruction 14 n'est donc pas 0 a priori. On peut cependant utiliser l'inégalité de la question 1.(d) pour affirmer que :

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

Si l'on prend $\varepsilon = 2 \times 10^{-4}$, alors $\boxed{\Lambda(x_0) - x_0 \text{ est un réel compris entre } -10^{-4} \text{ et } 10^{-4}}$.

3. (a) On a vu en question 1.(a) que Λ est une fonction de répartition d'une variable à densité. De plus, comme Λ est de classe \mathcal{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 λ est paire. Elle est définie sur l'intervalle \mathbb{R} symétrique par rapport à 0, et on a pour tout $x \in \mathbb{R}$:

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

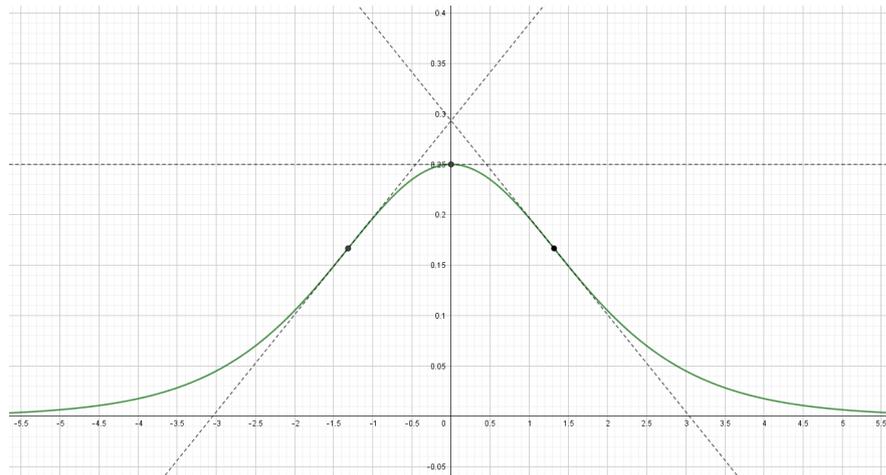
$\boxed{\text{La fonction } \lambda \text{ est donc paire.}}$

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

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

λ' ne s'annule qu'en $x = 0$. De plus, λ' est positive sur \mathbb{R}_- et négative sur \mathbb{R}_+ , de sorte que λ 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 λ .



Déterminons les points d'inflexion. Le signe de λ'' 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	-	+

La courbe représentative de λ admet deux points d'inflexion en $\ln(2 - \sqrt{3})$ et $\ln(2 + \sqrt{3})$. De plus, λ 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 λ 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 λ 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

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

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

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

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

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

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

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

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

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

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

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

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

admet une espérance comme combinaison linéaire de variables aléatoires admettant une espérance. Donc Y 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 $\int_{-\infty}^{+\infty} x\lambda(x) dx$ converge. Donc $E(Z)$ existe et vaut 0. Si Y suit la loi logistique $\mathcal{L}(r, s)$, $\frac{Y-r}{s}$ suit une loi logistique standard, et on a :

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

- (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 $U \hookrightarrow \mathcal{U}(]0, 1[)$. 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 $\mathcal{L}(r, s)$, on utilise le fait que si Z suit une loi logistique standard, on a $sZ + r \hookrightarrow \mathcal{L}(r, s)$.

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 $\mathcal{L}(r, s)$ par méthode d'inversion.

```

1 | def logis(n,p,r,s):
2 |     S = np.zeros(n,p)
3 |     for i in range(n) :
4 |         for j in range(p):
5 |             U = rd.random() # on simule une loi
6 |             Z = np.log(U/(1-U)) # on calcule L(U)
7 |             S[i,j] = s*Z+r
8 |     return S

```

- (c) Soit (Y_1, \dots, Y_n) un n -échantillon de loi mère la loi logistique $\mathcal{L}(r, s)$. On a montré dans le TD19 - Exercice 4, que la variance empirique $S_n^2 = \frac{1}{n} \sum_{k=1}^n (Y_k - \bar{Y}_n)^2$ est un estimateur asymptotiquement sans biais et convergent de la variance de $\mathcal{L}(r, s)$ (qui admet bien un moment d'ordre 4). Rappelons enfin qu'en Python, la commande `np.var(x)` calcule la variance 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).

```

1 | n = 10000 # n suffisamment grand
2 | Y = logis(1,n,r,s)
3 | print(np.var(Y))
    
```

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 \mathcal{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}) :

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

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

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

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

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

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

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

Ainsi Z admet λ 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.

```

1 | U1 = rd.exponential(1)
2 | U2 = rd.exponential(1)
3 | Z = np.log(U1/U2)
4 | print(Z)

```

Partie II. Variance de la loi logistique standard

6. (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 :

$$V(Z) = E(Z^2) = \int_{-\infty}^{+\infty} x^2 \frac{e^{-x}}{(1+e^{-x})^2} dx = 2 \int_0^{+\infty} \underbrace{x^2 \frac{e^{-x}}{(1+e^{-x})^2}}_{\text{fct. paire}} dx,$$

et cette dernière intégrale converge.

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

$$+ \left| \begin{array}{l} \boxed{\frac{e^{-x}}{1+e^{-x}}} \\ x \end{array} \right. \quad \swarrow \quad \boxed{\mathcal{E}^1}$$

$$- \left| \begin{array}{l} e^{-x} \\ (1+e^{-x})^2 \end{array} \right. \quad \nwarrow \quad \boxed{\frac{x^2}{2}}$$

On obtient :

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

Or, par croissance comparée

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

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

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

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

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

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

- $\forall k \in \llbracket 0, n \rrbracket$, $xe^{-(k+1)x} = o\left(\frac{1}{x^2}\right)$ 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 $x \geq 1$.
- $\int_1^{+\infty} \frac{dx}{x^2}$ 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 $A \rightarrow +\infty$, ce qui donne :

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

(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) :

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

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

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

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

En passant à la limite quand $n \rightarrow +\infty$ dans l'égalité du 5.(b), on obtient :

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

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

(d) On admet 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 $T_n = \sum_{k=1}^n \frac{1}{k^2}$. On sépare les termes pairs et impairs dans la somme

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

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

Or, on sait (résultat admis) que $\lim_{n \rightarrow +\infty} T_n = \frac{\pi^2}{6}$. En tant que suite extraite d'une suite convergente, on a $\lim_{n \rightarrow +\infty} T_{2n} = \frac{\pi^2}{6}$. Par somme de suites convergentes, $\lim_{n \rightarrow +\infty} S_{2n}$ existe et on a :

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

Or la suite de terme général $S_n = \sum_{k=1}^n \frac{(-1)^{k-1}}{k^2}$ converge d'après la question 5.(c), et sa limite vaut $\frac{V(Z)}{4}$ d'après la question 5.(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 :

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

7. (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 $I' = \int_0^{+\infty} g_1(x) dx$ et $J = \int_0^{+\infty} g_2(x) dx$ sont donc généralisées en 0 et en $+\infty$.

En 0, on a :

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

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

En $+\infty$, on a :

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

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

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

(b) Soit $U \hookrightarrow \mathcal{E}(1)$. 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 :

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

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 $V(Z) = \frac{\pi^2}{3}$ par la question 5.(d), on a $J - I^2 = \frac{\pi^2}{6}$.

Partie III. Estimation à partir de données binaires

7. Comme f est continue, F est donc de classe \mathcal{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, elle réalise une bijection de \mathbb{R} dans $]0, 1[$.

8. Les variables aléatoires (Y_1, \dots, 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 :

$$\frac{\bar{Y}_n - E(\bar{Y}_n)}{\sqrt{V(\bar{Y}_n)}} \xrightarrow{\mathcal{L}} Z \quad \text{où} \quad Z \hookrightarrow \mathcal{N}(0, 1).$$

Or par linéarité de l'espérance, on a $E(\bar{Y}_n) = \frac{1}{n} \sum_{j=1}^n E(Y_j) = F(\theta)$, et par indépendance :

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

On a donc :

$$\sqrt{n} \frac{\bar{Y}_n - F(\theta)}{\sqrt{F(\theta)(1 - F(\theta))}} \xrightarrow{\mathcal{L}} Z \quad \implies \quad \sqrt{n}(\bar{Y}_n - F(\theta)) \xrightarrow{\mathcal{L}} \sqrt{F(\theta)(1 - F(\theta))} Z.$$

Donc $\sqrt{n}(\bar{Y}_n - F(\theta))$ converge en loi vers la variable $\sqrt{F(\theta)(1 - F(\theta))} Z \hookrightarrow \mathcal{N}(0, F(\theta)(1 - F(\theta)))$.

9. (a) Les variables Y_1, \dots, Y_n sont à valeurs dans $\{0, 1\}$ donc \bar{Y}_n est à valeurs dans $\{0, 1/n, 2/n, \dots, (n-1)/n, 1\}$. Ainsi, on a :

$$P_\theta(E_n) = 1 - P_\theta(\bar{Y}_n = 0) - P_\theta(\bar{Y}_n = 1) = 1 - P_\theta(Y_1 = 0, \dots, Y_n = 0) - P_\theta(Y_1 = 1, \dots, Y_n = 1).$$

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

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

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 :

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

(b) i. Commençons par remarquer que :

$$\{\bar{Y}_n \leq F(x)\} \cap E_n = \{\omega \in E_n, \bar{Y}_n(\omega) \leq F(x)\}.$$

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

$$\{\omega \in E_n, T_n(\omega) \leq x\} = \{\omega \in E_n, \bar{Y}_n(\omega) \leq F(x)\}.$$

Montrons maintenant que pour tout $x \in \mathbb{R}$, $[T_n \leq x] \in \mathcal{A}$. 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 \mathcal{A} . De plus, on a :

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

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 \mathcal{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 \mathcal{A} (stabilité par intersection finie ou dénombrable), et $\boxed{[T_n \leq x] \in \mathcal{A}}$ 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_θ , on a :

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

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

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

(c) On va utiliser pour cela l'encadrement de la question précédente. Puisque $\lim_{n \rightarrow +\infty} 1 - P_\theta(E_n) = 0$, $P_\theta(T_n \leq x)$ 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 :

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

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

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

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 :

$$\begin{aligned} P_\theta(\overline{Y_n} \leq F(x)) &= P_\theta(\overline{Y_n} - F(\theta) \leq -(F(\theta) - F(x))) = P_\theta(F(\theta) - \overline{Y_n} \geq F(\theta) - F(x)) \\ &\leq P_\theta(|\overline{Y_n} - F(\theta)| \geq \underbrace{F(\theta) - F(x)}_{>0}) \xrightarrow{n \rightarrow +\infty} 0 \text{ par la LFGN.} \end{aligned}$$

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

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

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

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

$$\begin{aligned} P_\theta(\overline{Y}_n \leq F(x)) &= P_\theta(\overline{Y}_n - F(\theta) \leq F(x) - F(\theta)) = 1 - P_\theta(\overline{Y}_n - F(\theta) > F(x) - F(\theta)) \\ &\geq 1 - P_\theta(|\overline{Y}_n - F(\theta)| \geq F(x) - F(\theta)) \xrightarrow{n \rightarrow +\infty} 1. \end{aligned}$$

Par théorème d'encadrement, on en déduit que $\lim_{n \rightarrow +\infty} P_\theta(\overline{Y}_n \leq F(x)) = 1$. 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 :

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

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

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

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

Soit $\varepsilon > 0$. On a :

$$\begin{aligned} P_\theta(|T_n - \theta| \geq \varepsilon) &= 1 - P_\theta(|T_n - \theta| < \varepsilon) = 1 - P_\theta(\theta - \varepsilon < T_n < \theta + \varepsilon) \\ &= 1 - P_\theta(T_n < \theta + \varepsilon) + P_\theta(T_n \leq \theta - \varepsilon). \quad (*) \end{aligned}$$

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

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

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 $P_\theta(T_n \leq \theta + \varepsilon/2) \xrightarrow{n \rightarrow +\infty} 1$ par la question précédente (puisque $\theta + \varepsilon/2 > \theta$). Par théorème des gendarmes, $\lim_{n \rightarrow +\infty} P_\theta(T_n < \theta + \varepsilon)$ existe et vaut 1.

En reprenant (*), on obtient donc :

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

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



Pour aller plus loin.

Le résultat obtenu à la question 10.(c) montre que (T_n) converge en loi vers la variable constante égale à θ . On a montré à la question 10.(d) que (T_n) converge aussi en probabilité vers θ . 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 .

10. (a) i. La fonction f est continue et strictement positive (donc ne s'annule pas). La fonction $x \mapsto \frac{1}{f(x)}$ est donc continue sur \mathbb{R} . Par définition de la continuité en θ , 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 $U_n(\omega) = \frac{1}{f(\theta)}$ 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 θ 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) :

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

Or on a :

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

On a :

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

et

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

On obtient donc :

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

Par théorème des gendarmes, on en déduit que :

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

Ainsi (U_n) 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 :

$$\sqrt{n}(T_n - \theta) = U_n \times \sqrt{n}(\bar{Y}_n - F(\theta)).$$

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

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

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

Partie IV. Régression logistique

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

Pour montrer que $M^t M$ est inversible, on va montrer que l'équation $M^t M X = 0_p$ d'inconnue $X \in \mathcal{M}_{p,1}(\mathbb{R})$ admet pour unique solution $X = 0_{p,1}$. Soit donc $X \in \mathcal{M}_{p,1}(\mathbb{R})$ tel que $M^t M X = 0_{p,1}$. On a :

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

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

$$\dim \text{Ker}({}^t M) = \dim \mathcal{M}_{p,1}(\mathbb{R}) - \text{rg}({}^t M) = 0.$$

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

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

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

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

$$\Phi : \begin{array}{l} F \rightarrow \mathbb{R} \\ V \mapsto \|V - H\|^2 \end{array}$$

admet un minimum global atteint en un unique point $V = p_F(H)$, projeté orthogonal de H sur F . Et par injectivité de $X \mapsto {}^t M X$, on en déduit qu'il existe un unique vecteur $U \in \mathcal{M}_{p,1}(\mathbb{R})$ tel que ${}^t M U = p_F(H)$.

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 6. 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 $W \in \mathcal{M}_{p,1}(\mathbb{R})$, 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 :

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

- (c) Supposons que l'on connaisse les lois des variables $Y_{i,n}$. Cela signifie que l'on connaît la valeur de $b(x^{(1)}), \dots, b(x^{(n)})$ et donc, par bijectivité de la fonction Λ , les valeurs de $\langle \alpha, x^{(1)} \rangle, \dots, \langle \alpha, x^{(k)} \rangle$. Autrement dit, il existe β_1, \dots, β_k des réels connus tels que :

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

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. On ne peut pas déterminer α .

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

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

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 $x \mapsto c_i x$) :

$$\forall i \in \llbracket 1, k \rrbracket, \quad c_i T_{i,n} \xrightarrow{P} c_i \langle \alpha, x^{(i)} \rangle$$

Rappelons à présent le résultat suivant.

Rappel.

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 $X_n + Y_n \xrightarrow{P} X + Y$.

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

Propriété.

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 $i \in \llbracket 1, k \rrbracket$.

Alors on a $X_n^{(1)} + \dots + X_n^{(k)} \xrightarrow{P} X^{(1)} + \dots + X^{(k)}$.

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 :

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

Soit $i \in \llbracket 1, p \rrbracket$. 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 ${}^t M A = \begin{pmatrix} \sum_{j=1}^p \alpha_j x_j^{(1)} \\ \vdots \\ \sum_{j=1}^p \alpha_j x_j^{(k)} \end{pmatrix} = \begin{pmatrix} \langle \alpha, x^{(1)} \rangle \\ \vdots \\ \langle \alpha, x^{(k)} \rangle \end{pmatrix}$. On en

déduit que :

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

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

$$\boxed{\forall i \in \llbracket 1, p \rrbracket, A_{i,n} \xrightarrow{P} \alpha_i}$$