



PRÉPASIMPLE

prepasimple.fr

HEC — 2019

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Corrigé

Partie I : Fonction logistique et lois logistiques

- 1) a) On montre que Λ est continue et strictement croissante, on détermine ses limites, puis on résout l'équation $\Lambda(x) = y$.

La fonction Λ est dérivable sur $\mathbb{R}$ (quotient de fonctions dérivables, le dénominateur $1 + e^{-x}$ ne s'annulant pas) et, d'après le calcul de b), $\Lambda'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} > 0$. Elle est donc continue et strictement croissante sur $\mathbb{R}$. Comme $e^{-x} \rightarrow +\infty$ lorsque $x \rightarrow -\infty$ et $e^{-x} \rightarrow 0$ lorsque $x \rightarrow +\infty$, on a $\lim_{-\infty} \Lambda = 0$ et $\lim_{+\infty} \Lambda = 1$. Par le théorème de la bijection, Λ réalise une bijection de $\mathbb{R}$ sur $]0, 1[$.

Soit $y \in]0, 1[$ et $x \in \mathbb{R}$. Comme $y \neq 0$ et $1 - y > 0$:

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

Conclusion

Λ est une bijection de $\mathbb{R}$ sur $]0, 1[$, de réciproque $L : y \mapsto \ln\left(\frac{y}{1-y}\right)$.

- b) Avec $v(x) = 1 + e^{-x}$, $v'(x) = -e^{-x}$ et $\Lambda = \frac{1}{v}$, on a $\Lambda' = -\frac{v'}{v^2}$. On remarque aussi que $1 - \Lambda(x) = \frac{e^{-x}}{1 + e^{-x}}$, d'où une seconde écriture utile.

Conclusion

pour tout $x \in \mathbb{R}$, $\Lambda'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} = \Lambda(x)(1 - \Lambda(x))$. En particulier, comme $t(1-t) = \frac{1}{4} - \left(t - \frac{1}{2}\right)^2 \leq \frac{1}{4}$ pour tout réel t , on a $0 < \Lambda'(x) \leq \frac{1}{4}$.

- c) Posons $g(x) = \Lambda(x) - x$ pour $x \in \mathbb{R}$. La fonction g est dérivable et $g'(x) = \Lambda'(x) - 1 \leq \frac{1}{4} - 1 < 0$: g est continue et strictement décroissante. Lorsque $x \rightarrow -\infty$, $\Lambda(x) \rightarrow 0$ et $-x \rightarrow +\infty$, donc $g(x) \rightarrow +\infty$; lorsque $x \rightarrow +\infty$, $\Lambda(x) \rightarrow 1$ et $-x \rightarrow -\infty$, donc $g(x) \rightarrow -\infty$. Par le théorème de la bijection, g est une bijection de $\mathbb{R}$ sur $\mathbb{R}$: il existe un unique réel x_0 tel que $g(x_0) = 0$.

De plus, $g(0) = \frac{1}{2} > 0$ et $g(1) = \frac{1}{1+e^{-1}} - 1 = -\frac{e^{-1}}{1+e^{-1}} < 0$, donc $x_0 \in]0, 1[$.

Conclusion

il existe un unique réel x_0 tel que $\Lambda(x_0) = x_0$, et $0 < x_0 < 1$.

- d) La fonction Λ est de classe C^1 sur $\mathbb{R}$ et $|\Lambda'| \leq \frac{1}{4}$ d'après b). Par l'inégalité des accroissements finis, pour tout $x \in \mathbb{R}$:

$$|\Lambda(x) - x_0| = |\Lambda(x) - \Lambda(x_0)| \leq \frac{1}{4}|x - x_0| \leq |x - x_0|$$

Conclusion

pour tout $x \in \mathbb{R}$, $|\Lambda(x) - x_0| \leq |x - x_0|$.

- 2) a) Code adapté en Python. Le script Scilab de l'énoncé d'origine est remplacé par un code Python ; les numéros de lignes cités dans les questions sont ceux indiqués en commentaire.

Avec la fonction g de 1) c), strictement décroissante et nulle en x_0 :

$\Lambda(c) > c \iff g(c) > 0 \iff c < x_0$. La méthode de dichotomie conserve l'invariant $a < x_0 \leq b$: il est vrai au départ car $g(0) > 0 > g(1)$ (c'est pourquoi on choisit $a = 0$ et $b = 1$: ces valeurs encadrent x_0 d'après 1) c)). À chaque passage, si $\Lambda(c) > c$, alors $x_0 > c$ et on remplace a par c ; sinon $x_0 \leq c$ et on remplace b par c . La longueur $b - a$ est divisée par 2.

Python

```
import numpy as np                # (1)
def Lambda(x):                   # (2)
    return 1 / (1 + np.exp(-x)) # (3)

a = 0                            # (4)
b = 1                            # (5)
eps = 2e-4                      # (6)
while b - a > eps:               # (7)
    c = (a + b) / 2              # (8)
    if Lambda(c) > c:            # (9)
        a = c                   # (10)
    else:                        # (11)
        b = c                   # (12)
x0 = (a + b) / 2                 # (13)
```

Conclusion

la ligne (10) est $a = c$ et la ligne (12) est $b = c$; les valeurs $a = 0$ et $b = 1$ conviennent car $\Lambda(0) > 0$ et $\Lambda(1) < 1$, de sorte que $x_0 \in]0, 1[$.

- b)** En fin de boucle, $b - a \leq \text{eps}$ et $x_0 \in]a, b]$. La valeur renvoyée est le milieu $\frac{a+b}{2}$, donc l'erreur vérifie $\left| \frac{a+b}{2} - x_0 \right| \leq \frac{b-a}{2} \leq \frac{\text{eps}}{2}$. Pour garantir une erreur au plus égale à 10^{-4} , il suffit donc que $\frac{\text{eps}}{2} \leq 10^{-4}$.

Remarque : comme $b - a$ prend successivement les valeurs 2^{-m} , une analyse plus fine montre que toute valeur de eps strictement inférieure à $2^{-12} \approx 2,44 \times 10^{-4}$ convient aussi (la boucle s'arrête alors à $b - a = 2^{-13}$); la borne générale attendue est celle donnée ci-dessus.

Conclusion

on peut affecter à eps la valeur 2×10^{-4} , qui assure une erreur d'au plus 10^{-4} .

- c) Notons $\tilde{x}$ la valeur approchée obtenue. D'après l'inégalité triangulaire et 1) d) (avec la constante $\frac{1}{4}$ obtenue dans sa preuve) :

$$|\Lambda(\tilde{x}) - \tilde{x}| \leq |\Lambda(\tilde{x}) - x_0| + |x_0 - \tilde{x}| \leq \frac{5}{4}|\tilde{x} - x_0| \leq \frac{5}{4} \times 10^{-4}$$

Ce nombre n'est en général pas nul : c'est une valeur approchée de $g(x_0) = 0$, de signe opposé à celui de $\tilde{x} - x_0$ puisque g est strictement décroissante. Le code a été exécuté avec $\text{eps} = 2 \times 10^{-4}$: il donne $\tilde{x} \approx 0,658997$ (la solution vaut $x_0 \approx 0,659046$) et l'instruction (14) affiche environ $3,8 \times 10^{-5}$, valeur positive car $\tilde{x} < x_0$.

Conclusion

l'instruction (14) affiche un nombre proche de 0, de valeur absolue au plus $\frac{5}{4} \times 10^{-4}$ (donc au plus 2×10^{-4}), mais généralement non nul.

- 3) a) D'après 1) b), $\lambda(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$ est continue et positive sur $\mathbb{R}$. Pour $A, B > 0$,

$$\int_{-A}^B \lambda(x) dx = \Lambda(B) - \Lambda(-A), \text{ qui tend vers } 1 - 0 = 1 \text{ lorsque } A \text{ et } B \text{ tendent vers } +\infty. \text{ L'intégrale } \int_{-\infty}^{+\infty} \lambda \text{ converge donc et vaut } 1.$$

Conclusion

λ est une densité de probabilité, et Λ est la fonction de répartition associée.

- b) Parité. Pour $x \in \mathbb{R}$, en multipliant numérateur et dénominateur par e^{-2x} :

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

Variations. Avec $\lambda = \Lambda(1 - \Lambda)$ et $\Lambda' = \lambda$:

$$\lambda' = \lambda(1 - \Lambda) - \Lambda\lambda = \lambda(1 - 2\Lambda)$$

Comme $\lambda > 0$ et Λ est strictement croissante avec $\Lambda(0) = \frac{1}{2}$, λ' est strictement positive sur $] -\infty, 0[$ et strictement négative sur $]0, +\infty[$: λ croît puis décroît, avec un maximum $\lambda(0) = \frac{1}{4}$. De plus $0 < \lambda(x) \leq e^{-|x|}$ (voir 4) a)), donc $\lambda(x) \rightarrow 0$ en $\pm\infty$: l'axe des abscisses est asymptote en $-\infty$ et en $+\infty$.

Points d'inflexion. En dérivant $\lambda' = \lambda(1 - 2\Lambda)$:

$$\lambda'' = \lambda'(1 - 2\Lambda) - 2\lambda^2 = \lambda(1 - 2\Lambda)^2 - 2\lambda\Lambda(1 - \Lambda) = \lambda(6\Lambda^2 - 6\Lambda + 1)$$

Le trinôme $6t^2 - 6t + 1$ a pour racines $t_{\pm} = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$, toutes deux dans $]0, 1[$; il est positif à l'extérieur de $[t_-, t_+]$ et négatif à l'intérieur. Comme $\lambda > 0$ et Λ est une bijection strictement croissante de $\mathbb{R}$ sur $]0, 1[$, λ'' s'annule en changeant de signe exactement aux points $L(t_{\pm})$. Or :

$$L(t_+) = \ln\left(\frac{t_+}{1-t_+}\right) = \ln\left(\frac{3+\sqrt{3}}{3-\sqrt{3}}\right) = \ln\left(\frac{(3+\sqrt{3})^2}{6}\right) = \ln(2+\sqrt{3})$$

et $L(t_-) = -\ln(2+\sqrt{3})$ (car $t_- = 1 - t_+$ et $L(1-t) = -L(t)$). En ces points, $\lambda = t_{\pm}(1-t_{\pm}) = \frac{1}{4} - \frac{3}{36} = \frac{1}{6}$.

La courbe est en cloche, symétrique par rapport à l'axe des ordonnées. Le code suivant (exécuté) la trace.

Python

```
import numpy as np

import matplotlib.pyplot as plt

x = np.linspace(-6, 6, 400)
```

```
y = np.exp(-x) / (1 + np.exp(-x)) ** 2
plt.plot(x, y)
plt.show()
```

Conclusion

λ est paire, croissante sur $] -\infty, 0]$ et décroissante sur $[0, +\infty[$, de maximum $\frac{1}{4}$ en 0, de limite 0 en $\pm\infty$; sa courbe admet deux points d'inflexion, $\left(\pm \ln(2 + \sqrt{3}), \frac{1}{6}\right)$, avec $\ln(2 + \sqrt{3}) \approx 1,32$.

- 4) a) Soit Z de loi logistique standard. Pour $x \geq 0$, $1 + e^{-x} \geq 1$, donc $\lambda(x) \leq e^{-x}$; par parité, $\lambda(x) \leq e^{-|x|}$ pour tout réel x . Soit $m \in \mathbb{N}$. On a $0 \leq |x|^m \lambda(x) \leq |x|^m e^{-|x|}$, et $\int_0^{+\infty} x^m e^{-x} dx = \Gamma(m+1) = m!$ converge; par parité, $\int_{-\infty}^{+\infty} |x|^m e^{-|x|} dx$ converge. Par comparaison de fonctions positives, $\int_{-\infty}^{+\infty} |x|^m \lambda(x) dx$ converge : Z admet un moment d'ordre m .

Comme $x \mapsto x\lambda(x)$ est impaire et que son intégrale converge absolument, $E(Z) = 0$.

Soit Y de loi $\mathcal{L}(r, s)$: $Z = \frac{Y - r}{s}$ suit la loi logistique standard et $Y = r + sZ$. Par la formule du binôme,

$$Y^m = \sum_{j=0}^m \binom{m}{j} r^{m-j} s^j Z^j$$

est une combinaison linéaire de variables admettant une espérance : Y admet un moment d'ordre m . Enfin $E(Y) = r + sE(Z) = r$.

Conclusion

une variable de loi $\mathcal{L}(r, s)$ admet des moments de tout ordre, et son espérance vaut r .

- b) Code adapté en Python. La fonction Scilab `grandlogis` devient une fonction Python renvoyant un tableau à n lignes et p colonnes.

Fonction de répartition. Si Y suit $\mathcal{L}(r, s)$, pour $y \in \mathbb{R}$, comme $s > 0$:

$$F_Y(y) = P\left(Z \leq \frac{y-r}{s}\right) = \Lambda\left(\frac{y-r}{s}\right).$$

Méthode d'inversion. Soit U une variable de loi uniforme sur $]0, 1[$ et $Y' = r + sL(U)$.

Pour $y \in \mathbb{R}$, comme L et Λ sont des bijections réciproques strictement croissantes :

$$P([Y' \leq y]) = P\left([L(U) \leq \frac{y-r}{s}]\right) = P\left([U \leq \Lambda\left(\frac{y-r}{s}\right)]\right) = \Lambda\left(\frac{y-r}{s}\right)$$

donc Y' suit la loi $\mathcal{L}(r, s)$. On applique $u \mapsto r + s \ln\left(\frac{u}{1-u}\right)$ coefficient par coefficient à un tableau de simulations uniformes indépendantes (la valeur 0, renvoyée avec une probabilité négligeable par `rd.random`, est sans incidence théorique).

Python

```
import numpy as np

import numpy.random as rd

def grandlogis(n, p, r, s):
    U = rd.random((n, p))
    S = r + s * np.log(U / (1 - U))
    return S
```

Les coefficients de S sont des fonctions de coefficients indépendants de U : ce sont des simulations indépendantes de la loi $\mathcal{L}(r, s)$. Code exécuté : `grandlogis(3, 4, 2, 0.5)` renvoie bien un tableau de taille $(3, 4)$.

Conclusion

la fonction ci-dessus convient.

- c) Soit $Z_1, \dots, Z_N$ des variables indépendantes de loi logistique standard. D'après a), $E(Z_1) = 0$, donc $V(Z_1) = E(Z_1^2)$, et Z_1^2 admet une variance (moment d'ordre 4). Les variables $Z_1^2, \dots, Z_N^2$ sont indépendantes, de même loi, d'espérance $V(Z_1)$: par la loi faible des grands nombres,

$$\frac{1}{N} \sum_{j=1}^N Z_j^2$$

converge en probabilité vers $V(Z_1)$.

Python

```
N = 100000
S = grandlogis(1, N, 0, 1)
v = np.mean(S ** 2)
print(v)
```

On peut aussi utiliser `np.var(S)`, qui calcule la variance empirique

$$\frac{1}{N} \sum_{j=1}^N (z_j - \bar{z})^2$$

(convention de division par N). Code exécuté : on obtient $v \approx 3,30$, à comparer avec la valeur exacte $\frac{\pi^2}{3} \approx 3,29$ établie en 9) d). Cette observation n'est pas une preuve.

Conclusion

on simule un grand nombre N de variables logistiques standard indépendantes avec `grandlogis(1, N, 0, 1)` et on calcule la moyenne de leurs carrés, qui approche la variance d'après la loi faible des grands nombres.

- 5) a) Question adaptée au programme actuel : l'indication est précisée (densités de $\ln(U_1)$ et $-\ln(U_2)$, convolution, changement de variable $u = e^t$) afin d'éviter tout calcul par densité conjointe.

Les variables U_1 et U_2 sont presque sûrement strictement positives (car $P([U_i \leq 0]) = 0$), donc Z est défini presque sûrement et $Z = W_1 + W_2$ avec $W_1 = \ln(U_1)$ et $W_2 = -\ln(U_2)$.

Loi de W_1 . Pour $w \in \mathbb{R}$, comme $\exp$ est strictement croissante :

$P([W_1 \leq w]) = P([U_1 \leq e^w]) = 1 - \exp(-e^w)$. Cette fonction est de classe C^1 sur $\mathbb{R}$: W_1 est à densité, de densité $g_1(w) = e^w \exp(-e^w)$.

Loi de W_2 . $P([W_2 \leq w]) = P([U_2 \geq e^{-w}]) = \exp(-e^{-w})$, de classe C^1 sur $\mathbb{R}$: W_2 admet pour densité $g_2(w) = e^{-w} \exp(-e^{-w})$.

Les variables W_1 et W_2 sont indépendantes (fonctions respectivement de U_1 et de U_2 , indépendantes). La densité g_1 est bornée : avec $u = e^w$, $g_1(w) = ue^{-u} \leq e^{-1}$, car $u \mapsto ue^{-u}$ atteint son maximum en $u = 1$. Par le théorème de convolution, $Z = W_1 + W_2$ admet pour densité :

$$h(z) = \int_{-\infty}^{+\infty} g_1(t) g_2(z-t) dt = \int_{-\infty}^{+\infty} e^t e^{-e^t} e^{t-z} e^{-e^{t-z}} dt = e^{-z} \int_{-\infty}^{+\infty} e^{2t} \exp(-e^t \beta) dt$$

où $\beta = 1 + e^{-z} > 0$. Sur un segment $[a, b]$, le changement de variable $u = e^t$ ($t \mapsto e^t$ de classe C^1 , $du = e^t dt$) donne :

$$\int_a^b e^{2t} e^{-\beta e^t} dt = \int_{e^a}^{e^b} u e^{-\beta u} du$$

Lorsque $a \rightarrow -\infty$ et $b \rightarrow +\infty$, $e^a \rightarrow 0$ et $e^b \rightarrow +\infty$. Or, par le changement de variable affine $v = \beta u$, $\int_0^{+\infty} u e^{-\beta u} du = \frac{1}{\beta^2} \int_0^{+\infty} v e^{-v} dv = \frac{\Gamma(2)}{\beta^2} = \frac{1}{\beta^2}$ (intégrale convergente).
Donc :

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

Conclusion

$Z = \ln \left(\frac{U_1}{U_2} \right)$ admet λ pour densité : elle suit la loi logistique standard.

- b) Code adapté en Python. L'appel à la fonction Scilab `grand` est remplacé par `rd.exponential(1)`, dont l'argument est l'inverse du paramètre, ici égal à 1.

D'après a), il suffit de simuler deux variables indépendantes de loi exponentielle de paramètre 1 (deux appels successifs donnent des simulations indépendantes) et de calculer le logarithme de leur quotient.

Python

```
import numpy as np
import numpy.random as rd

def simullogis():
    u1 = rd.exponential(1)
    u2 = rd.exponential(1)
    return np.log(u1 / u2)
```

Code exécuté (version vectorisée sur 10^5 simulations) : moyenne empirique $\approx 0,02$ et variance empirique $\approx 3,28$, cohérentes avec $E(Z) = 0$ et $V(Z) = \frac{\pi^2}{3}$.

Conclusion

`simullogis()` renvoie une simulation de la loi logistique standard.

Partie II : Variance de la loi logistique standard

- 6) a) Pour $n = 0$, la somme n'a qu'un terme : $P_0 = \binom{1}{1}(X - 1)^0 = 1$. Pour $n = 1$:
- $$P_1 = \binom{3}{1}(X - 1) - \binom{3}{3} = 3X - 3 - 1.$$

Conclusion

$$P_0 = 1 \text{ et } P_1 = 3X - 4.$$

- b) Soit $n \in \mathbb{N}^*$. Le terme d'indice k est de degré $n - k$. Seuls les termes $k = 0$ et $k = 1$ contribuent aux coefficients de X^n et X^{n-1} . Par la formule du binôme,
- $$(X - 1)^n = X^n - nX^{n-1} + (\text{termes de degré} \leq n - 2) \text{ et}$$
- $$(X - 1)^{n-1} = X^{n-1} + (\text{termes de degré} \leq n - 2). \text{ Donc :}$$

$$P_n = (2n + 1)X^n - \left[n(2n + 1) + \binom{2n + 1}{3} \right] X^{n-1} + (\text{termes de degré} \leq n - 2)$$

Or

$$\binom{2n + 1}{3} = \frac{(2n + 1)(2n)(2n - 1)}{6} = \frac{n(2n + 1)(2n - 1)}{3},$$

donc :

$$\begin{aligned} n(2n + 1) + \binom{2n + 1}{3} &= n(2n + 1) \left(1 + \frac{2n - 1}{3} \right) = \frac{n(2n + 1)(2n + 2)}{3} \\ &= \frac{2n(n + 1)(2n + 1)}{3} \end{aligned}$$

Conclusion

pour $n \geq 1$, P_n est de degré n , ses deux monômes de plus hauts degrés étant $(2n + 1)X^n$ et $-\frac{2n(n + 1)(2n + 1)}{3}X^{n-1}$. (Pour $n = 1$, on retrouve $3X - 4$.)

- c) D'après b), P_n est de degré n et de coefficient dominant $a = 2n + 1$. Comme $t_1, \dots, t_n$ sont des racines deux à deux distinctes de P_n , le polynôme

$$\prod_{k=1}^n (X - t_k)$$

divise P_n (cours sur les racines). Le quotient est de degré $n - n = 0$, donc constant, et en comparant les coefficients dominants, il vaut a . Ainsi

$$P_n = a \prod_{k=1}^n (X - t_k).$$

Montrons par récurrence sur $m \in \llbracket 1, n \rrbracket$ que

$$\prod_{k=1}^m (X - t_k) = X^m - \left(\sum_{k=1}^m t_k \right) X^{m-1} + R_m$$

avec $\deg R_m \leq m - 2$. C'est vrai pour $m = 1$ ($R_1 = 0$). Si c'est vrai au rang m , on multiplie par $X - t_{m+1}$: le coefficient de X^m devient

$$- \sum_{k=1}^m t_k - t_{m+1},$$

et les autres termes sont de degré au plus $m - 1$.

En comparant les coefficients de X^{n-1} dans

$$P_n = a \prod_{k=1}^n (X - t_k)$$

et dans b) :

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

Conclusion

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

et

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

- 7) a) Question adaptée au programme actuel : l'égalité, obtenue dans l'énoncé d'origine par un calcul sur les nombres complexes, est établie ici par une récurrence utilisant uniquement les formules d'addition.

Notons $c = \cos(x)$ et $s = \sin(x)$, et utilisons la convention $\binom{m}{-1} = \binom{m}{m+1} = 0$. Les formules à établir s'écrivent, en notant $k = 2j$ dans la première et $k = 2j + 1$ dans la seconde :

$$\begin{aligned} \cos(mx) &= \sum_{\substack{0 \leq k \leq m \\ k \text{ pair}}} (-1)^{k/2} \binom{m}{k} c^{m-k} s^k \quad \text{et} \quad \sin(mx) \\ &= \sum_{\substack{0 \leq k \leq m \\ k \text{ impair}}} (-1)^{(k-1)/2} \binom{m}{k} c^{m-k} s^k \end{aligned}$$

Initialisation ($m = 0$) : $\cos(0) = 1 = \binom{0}{0} c^0 s^0$ et $\sin(0) = 0$, somme vide.

Hérédité. Supposons les formules vraies au rang m . Par les formules d'addition, $\cos((m+1)x) = c \cos(mx) - s \sin(mx)$, donc :

$$\cos((m+1)x) = \sum_{k \text{ pair}, k \leq m} (-1)^{k/2} \binom{m}{k} c^{m+1-k} s^k + \sum_{k \text{ impair}, k \leq m} (-1)^{(k+1)/2} \binom{m}{k} c^{m-k} s^{k+1}$$

Dans la seconde somme, on pose $k' = k + 1$, pair, compris entre 2 et $m + 1$: le terme devient $(-1)^{k'/2} \binom{m}{k'-1} c^{m+1-k'} s^{k'}$. En regroupant, avec les conventions ci-dessus et la formule de Pascal :

$$\begin{aligned}\cos((m+1)x) &= \sum_{k \text{ pair}, 0 \leq k \leq m+1} (-1)^{k/2} \left[\binom{m}{k} + \binom{m}{k-1} \right] c^{m+1-k} s^k \\ &= \sum_{k \text{ pair}, 0 \leq k \leq m+1} (-1)^{k/2} \binom{m+1}{k} c^{m+1-k} s^k\end{aligned}$$

De même, $\sin((m+1)x) = c \sin(mx) + s \cos(mx)$ donne :

$$\sin((m+1)x) = \sum_{k \text{ impair}, k \leq m} (-1)^{(k-1)/2} \binom{m}{k} c^{m+1-k} s^k + \sum_{k \text{ pair}, k \leq m} (-1)^{k/2} \binom{m}{k} c^{m-k} s^{k+1}$$

et, en posant $k' = k + 1$ (impair) dans la seconde somme, $(-1)^{k/2} = (-1)^{(k'-1)/2}$, d'où :

$$\begin{aligned}\sin((m+1)x) &= \sum_{k \text{ impair}, 1 \leq k \leq m+1} (-1)^{(k-1)/2} \left[\binom{m}{k} + \binom{m}{k-1} \right] c^{m+1-k} s^k \\ &= \sum_{k \text{ impair}, 1 \leq k \leq m+1} (-1)^{(k-1)/2} \binom{m+1}{k} c^{m+1-k} s^k\end{aligned}$$

Les formules sont donc vraies pour tout $m \in \mathbb{N}$.

Avec $m = 2n + 1$, les indices impairs $k \leq 2n + 1$ s'écrivent $k = 2j + 1$ avec $0 \leq j \leq n$, et $m - k = 2(n - j)$. En renommant j en k :

Conclusion

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

b) Soit $x \in]0, \pi[$; alors $\sin(x) > 0$. En divisant l'égalité de a) par $\sin^{2n+1}(x)$:

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

Or $\frac{\cos^2(x)}{\sin^2(x)} = \frac{1 - \sin^2(x)}{\sin^2(x)} = \frac{1}{\sin^2(x)} - 1.$

Conclusion

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

c) Soit $n \in \mathbb{N}^*$. Pour $k \in \llbracket 1, n \rrbracket$, posons $x_k = \frac{k\pi}{2n+1}$ et $t_k = \frac{1}{\sin^2(x_k)}$. Comme

$1 \leq k \leq n < \frac{2n+1}{2}$, on a $0 < x_1 < \dots < x_n < \frac{\pi}{2}$. La fonction $\sin$ est strictement croissante et strictement positive sur $]0, \frac{\pi}{2}[$, donc les réels $\sin^2(x_k)$ sont deux à deux distincts et non nuls, et les t_k sont deux à deux distincts.

Par 7) b) appliquée à $x_k \in]0, \pi[$: $P_n(t_k) = \frac{\sin(k\pi)}{\sin^{2n+1}(x_k)} = 0$. Ainsi $t_1, \dots, t_n$ sont n racines réelles deux à deux distinctes de P_n , et 6) c) s'applique.

Conclusion

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

8) a) Méthode 1 — Par étude de fonctions.

Soit $\varphi(t) = t - \sin(t)$ et $\psi(t) = \tan(t) - t$ sur $\left[0, \frac{\pi}{2}\right]$. Ces fonctions sont dérivables, $\varphi'(t) = 1 - \cos(t) \geq 0$ et $\psi'(t) = 1 + \tan^2(t) - 1 = \tan^2(t) \geq 0$: elles sont croissantes et nulles en 0, donc positives. D'où $\sin(x) \leq x \leq \tan(x)$.

Méthode 2 — Par la convexité.

Sur $\left[0, \frac{\pi}{2}\right]$, $\sin'' = -\sin \leq 0$: $\sin$ est concave, donc sa courbe est en dessous de sa

tangente en 0, d'équation $y = t$; ainsi $\sin(x) \leq x$. De même $\tan'' = 2 \tan(1 + \tan^2) \geq 0$: $\tan$ est convexe sur $\left[0, \frac{\pi}{2}\right]$, donc au-dessus de sa tangente en 0, d'équation $y = t$; ainsi $x \leq \tan(x)$.

La méthode 1 est la plus directe ; la méthode 2 montre que ces inégalités traduisent la position d'une courbe par rapport à une tangente, réflexe utile pour de nombreuses inégalités classiques.

Les trois nombres $\sin(x)$, x , $\tan(x)$ sont strictement positifs. De $0 < \sin(x) \leq x$ on tire $\frac{1}{x^2} \leq \frac{1}{\sin^2(x)}$. De $0 < x \leq \tan(x)$ on tire :

$$\frac{1}{x^2} \geq \frac{1}{\tan^2(x)} = \frac{\cos^2(x)}{\sin^2(x)} = \frac{1 - \sin^2(x)}{\sin^2(x)} = \frac{1}{\sin^2(x)} - 1$$

Conclusion

$$\sin(x) \leq x \leq \tan(x) \text{ et } \frac{1}{\sin^2(x)} - 1 \leq \frac{1}{x^2} \leq \frac{1}{\sin^2(x)}.$$

b) Soit $n \in \mathbb{N}^*$. On applique a) à $x_k = \frac{k\pi}{2n+1} \in \left]0, \frac{\pi}{2}\right[$ (voir 7) c)) pour $k \in \llbracket 1, n \rrbracket$:

$$\frac{1}{x_k^2} = \frac{(2n+1)^2}{\pi^2} \cdot \frac{1}{k^2}. \text{ En sommant les } n \text{ inégalités :}$$

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

Avec 7) c), le membre de gauche vaut $\frac{2n(n+1)}{3} - n = \frac{2n^2 - n}{3} = \frac{n(2n-1)}{3}.$

Conclusion

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

c) En multipliant l'encadrement de b) par $\frac{\pi^2}{(2n+1)^2} > 0$:

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

Les deux membres extrêmes sont des quotients de polynômes en n de degré 2 : le premier est équivalent à $\frac{2\pi^2 n^2}{12n^2}$ et le second à $\frac{2\pi^2 n^2}{12n^2}$, donc ils tendent tous deux vers $\frac{\pi^2}{6}$. Par encadrement, la suite des sommes partielles converge vers $\frac{\pi^2}{6}$.

Conclusion

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

9) a) D'après 4) a), Z admet une variance et $E(Z) = 0$, donc

$$V(Z) = E(Z^2) = \int_{-\infty}^{+\infty} x^2 \lambda(x) dx \text{ (théorème de transfert, intégrale convergente). La}$$

$$\text{fonction } x \mapsto x^2 \lambda(x) \text{ est paire, donc } V(Z) = 2 \int_0^{+\infty} x^2 \lambda(x) dx.$$

Intégration par parties sur un segment $[0, A]$, avec $u(x) = x^2$ et $v(x) = \Lambda(x) - 1 = -\frac{e^{-x}}{1+e^{-x}}$, fonctions de classe C^1 avec $v' = \lambda$:

$$\int_0^A x^2 \lambda(x) dx = [x^2 (\Lambda(x) - 1)]_0^A + \int_0^A 2x (1 - \Lambda(x)) dx = -\frac{A^2 e^{-A}}{1+e^{-A}} + 2 \int_0^A \frac{x e^{-x}}{1+e^{-x}} dx$$

Lorsque $A \rightarrow +\infty$, $0 \leq \frac{A^2 e^{-A}}{1+e^{-A}} \leq A^2 e^{-A} \rightarrow 0$ par croissances comparées. Le membre de gauche converge, donc $\int_0^{+\infty} \frac{x e^{-x}}{1+e^{-x}} dx$ converge (ce que confirme la majoration $0 \leq \frac{x e^{-x}}{1+e^{-x}} \leq x e^{-x}$) et :

$$\int_0^{+\infty} x^2 \lambda(x) dx = 2 \int_0^{+\infty} \frac{x e^{-x}}{1 + e^{-x}} dx$$

Conclusion

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

b) Soit $x \geq 0$ et $q = e^{-x} \in]0, 1]$. Comme $-q \neq 1$, la somme géométrique donne

$$\sum_{k=0}^n (-q)^k = \frac{1 - (-q)^{n+1}}{1 + q},$$

d'où :

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

En multipliant par $x e^{-x} = xq$:

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

Chaque intégrale sur $[0, +\infty[$ des termes du membre de droite converge :

$0 \leq x e^{-(k+1)x} \leq x e^{-x}$ et $0 \leq \frac{x e^{-(n+2)x}}{1 + e^{-x}} \leq x e^{-x}$, avec $\int_0^{+\infty} x e^{-x} dx = \Gamma(2) = 1$. Par linéarité des intégrales convergentes :

Conclusion

$$\int_0^{+\infty} \frac{x e^{-x}}{1 + e^{-x}} dx = \sum_{k=0}^n (-1)^k \int_0^{+\infty} x e^{-(k+1)x} dx + I_n.$$

c) Pour $a > 0$, le changement de variable affine $u = ax$ donne

$$\int_0^{+\infty} x e^{-ax} dx = \frac{1}{a^2} \int_0^{+\infty} u e^{-u} du = \frac{1}{a^2}. \text{ Donc :}$$

$$|I_n| \leq \int_0^{+\infty} \frac{x e^{-(n+2)x}}{1 + e^{-x}} dx \leq \int_0^{+\infty} x e^{-(n+2)x} dx = \frac{1}{(n+2)^2} \rightarrow 0$$

Notons $K = \int_0^{+\infty} \frac{x e^{-x}}{1 + e^{-x}} dx$. D'après b) et ce calcul,

$$\sum_{k=0}^n \frac{(-1)^k}{(k+1)^2} = K - I_n,$$

qui tend vers K . La série converge donc et a pour somme K .

Conclusion

$I_n \rightarrow 0$ et

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

d) Notons

$$S_m = \sum_{k=0}^{m-1} \frac{(-1)^k}{(k+1)^2} = \sum_{j=1}^m \frac{(-1)^{j-1}}{j^2}.$$

Pour j pair, $\frac{(-1)^{j-1}}{j^2} = \frac{1}{j^2} - \frac{2}{j^2}$; pour j impair, il vaut $\frac{1}{j^2}$. Donc, pour $N \in \mathbb{N}^*$:

$$S_{2N} = \sum_{j=1}^{2N} \frac{1}{j^2} - 2 \sum_{i=1}^N \frac{1}{(2i)^2} = \sum_{j=1}^{2N} \frac{1}{j^2} - \frac{1}{2} \sum_{i=1}^N \frac{1}{i^2}$$

D'après 8) c), $S_{2N} \rightarrow \frac{\pi^2}{6} - \frac{\pi^2}{12} = \frac{\pi^2}{12}$. Comme la suite (S_m) converge d'après c), sa limite est celle de la sous-suite (S_{2N}) :

$$\sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)^2} = \frac{\pi^2}{12}.$$

Avec a) et c) :

$$V(Z) = 4 \times \frac{\pi^2}{12}$$

Conclusion

$$V(Z) = \frac{\pi^2}{3}.$$

10) a) Pour $m \in \{1, 2\}$, la fonction $x \mapsto (\ln(x))^m e^{-x}$ est continue sur $]0, +\infty[$; on étudie la convergence absolue en 0 et en $+\infty$.

En 0 : $0 \leq |\ln(x)|^m e^{-x} \leq |\ln(x)|^m$ et, par croissances comparées, $x^{1/2} |\ln(x)|^m \rightarrow 0$, donc $|\ln(x)|^m = o\left(\frac{1}{x^{1/2}}\right)$; l'intégrale $\int_0^1 \frac{dx}{x^{1/2}}$ converge ($\frac{1}{2} < 1$).

En $+\infty$: $x^2 |\ln(x)|^m e^{-x} \rightarrow 0$ par croissances comparées, donc $|\ln(x)|^m e^{-x} = o\left(\frac{1}{x^2}\right)$, et $\int_1^{+\infty} \frac{dx}{x^2}$ converge.

Par comparaison de fonctions positives, les deux intégrales convergent absolument.

Conclusion

$$\int_0^{+\infty} \ln(x) e^{-x} dx \text{ et } \int_0^{+\infty} (\ln(x))^2 e^{-x} dx \text{ convergent.}$$

- b) U_1 admet pour densité $x \mapsto e^{-x}$ sur $]0, +\infty[$ (et 0 ailleurs). Par le théorème de transfert, les intégrales de a) étant absolument convergentes, $\ln(U_1)$ admet une espérance et un moment d'ordre 2 : $E(\ln(U_1)) = I$ et $E((\ln(U_1))^2) = J$. Par Koenig-Huygens, $V(\ln(U_1)) = J - I^2$, et de même $V(\ln(U_2)) = J - I^2$.

Les variables $\ln(U_1)$ et $-\ln(U_2)$ sont indépendantes, donc :

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

Or d'après 5) a), $Z = \ln\left(\frac{U_1}{U_2}\right)$ suit la loi logistique standard, donc $V(Z) = \frac{\pi^2}{3}$ par 9) d).

Conclusion

$$J - I^2 = \frac{\pi^2}{6}. \text{ (Une simulation de } 10^6 \text{ valeurs de } \ln(U_1) \text{ donne une variance empirique } \approx 1,640, \text{ pour } \frac{\pi^2}{6} \approx 1,645.)$$

Partie III : Estimation à partir de données binaires

- 11) Comme f est une densité continue sur $\mathbb{R}$, $F(x) = \int_{-\infty}^x f(t) dt$ est de classe C^1 sur $\mathbb{R}$ avec $F' = f > 0$: F est continue et strictement croissante. Ses limites sont 0 en $-\infty$ et 1 en $+\infty$ (fonction de répartition). Par le théorème de la bijection, F réalise une bijection de $\mathbb{R}$ sur $]0, 1[$.

Conclusion

F est une bijection de $\mathbb{R}$ sur $]0, 1[$; sa réciproque F^{-1} est continue et strictement croissante sur $]0, 1[$.

- 12) Les variables Y_j sont indépendantes, de même loi de Bernoulli de paramètre $F(\theta)$, d'espérance $F(\theta)$ et de variance $\sigma^2 = F(\theta)(1 - F(\theta))$, strictement positive car $F(\theta) \in]0, 1[$ d'après 11). Par le théorème limite central :

$\frac{\sqrt{n}(\bar{Y}_n - F(\theta))}{\sigma}$ converge en loi vers une variable N de loi $\mathcal{N}(0, 1)$.

Par composition avec la fonction continue $t \mapsto \sigma t$, $\sqrt{n}(\bar{Y}_n - F(\theta))$ converge en loi vers σN , qui suit la loi $\mathcal{N}(0, \sigma^2)$.

Conclusion

la suite converge en loi vers une variable de loi normale centrée de variance $F(\theta)(1 - F(\theta))$.

- 13) a) La variable $n\bar{Y}_n$ suit la loi binomiale $\mathcal{B}(n, F(\theta))$ (somme de n variables de Bernoulli indépendantes de paramètre $F(\theta)$). L'événement contraire de E_n est la réunion disjointe de $[\bar{Y}_n = 0] = [n\bar{Y}_n = 0]$ et $[\bar{Y}_n = 1] = [n\bar{Y}_n = n]$. Donc :

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

Comme $F(\theta)$ et $1 - F(\theta)$ appartiennent à $]0, 1[$, leurs puissances n -ièmes tendent vers 0.

Conclusion

$P_\theta(E_n) = 1 - (1 - F(\theta))^n - F(\theta)^n$, qui tend vers 1.

- b) Question adaptée au programme actuel : la démonstration du fait que $[T_n \leq x]$ appartient à la tribu $\mathcal{A}$ (question de mesurabilité) est remplacée par l'admission que T_n est une variable aléatoire.

- (i) Soit $\omega \in E_n$. Alors $\bar{Y}_n(\omega) \in]0, 1[$ et $T_n(\omega) = F^{-1}(\bar{Y}_n(\omega))$. Comme F est strictement croissante (11) :

$$T_n(\omega) \leq x \iff F(T_n(\omega)) \leq F(x) \iff \bar{Y}_n(\omega) \leq F(x)$$

Conclusion

$$\{\omega \in E_n ; T_n(\omega) \leq x\} = [\bar{Y}_n \leq F(x)] \cap E_n.$$

- (ii) $[T_n \leq x]$ est la réunion disjointe de $[T_n \leq x] \cap E_n$ et de $[T_n \leq x] \cap \bar{E}_n$. Le premier événement est $[\bar{Y}_n \leq F(x)] \cap E_n$ d'après (i) ; le second est inclus dans $\bar{E}_n$, donc sa probabilité est comprise entre 0 et $1 - P_\theta(E_n)$.

Conclusion

$$P_\theta([\bar{Y}_n \leq F(x)] \cap E_n) \leq P_\theta([T_n \leq x]) \leq P_\theta([\bar{Y}_n \leq F(x)] \cap E_n) + 1 - P_\theta(E_n).$$

- c) On a $E(\bar{Y}_n) = F(\theta)$ et $V(\bar{Y}_n) = \frac{\sigma^2}{n}$. Par l'inégalité de Bienaymé-Tchebychev, pour tout $\delta > 0$: $P_\theta(|\bar{Y}_n - F(\theta)| \geq \delta) \leq \frac{\sigma^2}{n\delta^2} \rightarrow 0$.

Cas $x < \theta$. Alors $\delta = F(\theta) - F(x) > 0$ (stricte croissance de F), et $[\bar{Y}_n \leq F(x)] \subset [|\bar{Y}_n - F(\theta)| \geq \delta]$. Le majorant de b) (ii) vérifie donc :

$$0 \leq P_\theta([T_n \leq x]) \leq \frac{\sigma^2}{n\delta^2} + 1 - P_\theta(E_n) \rightarrow 0$$

d'après a). Par encadrement, $P_\theta([T_n \leq x]) \rightarrow 0$.

Cas $x > \theta$. Alors $\delta = F(x) - F(\theta) > 0$ et $[\bar{Y}_n > F(x)] \subset [|\bar{Y}_n - F(\theta)| \geq \delta]$, donc

$P_\theta([\bar{Y}_n \leq F(x)]) \geq 1 - \frac{\sigma^2}{n\delta^2}$. Pour des événements B et E , $P(B \cap E) = P(B) - P(B \cap \bar{E}) \geq P(B) - P(\bar{E})$, donc avec b) (ii) :

$$1 \geq P_\theta([T_n \leq x]) \geq 1 - \frac{\sigma^2}{n\delta^2} - (1 - P_\theta(E_n)) \rightarrow 1$$

Conclusion

$\lim_{n \rightarrow +\infty} P_\theta([T_n \leq x])$ vaut 0 si $x < \theta$ et 1 si $x > \theta$.

- d) Pour tout n , T_n est une fonction de $(Y_1, \dots, Y_n)$ qui ne dépend pas de θ (F est connue) : c'est un estimateur de θ . Soit $\varepsilon > 0$. On a
 $[|T_n - \theta| > \varepsilon] \subset [T_n \leq \theta - \varepsilon] \cup [T_n > \theta + \varepsilon]$, donc :

$$P_\theta(|T_n - \theta| > \varepsilon) \leq P_\theta([T_n \leq \theta - \varepsilon]) + 1 - P_\theta([T_n \leq \theta + \varepsilon])$$

D'après c), appliqué en $x = \theta - \varepsilon < \theta$ et $x = \theta + \varepsilon > \theta$, le majorant tend vers $0 + 1 - 1 = 0$.

Conclusion

(T_n) converge en probabilité vers θ : c'est une suite convergente d'estimateurs de θ .

- 14) a) (i) La fonction f est continue et strictement positive sur $\mathbb{R}$, donc $\frac{1}{f}$ est continue sur $\mathbb{R}$, en particulier en θ . Par définition de la continuité en θ , il existe $\alpha > 0$ tel que pour tout x vérifiant $|x - \theta| \leq \alpha$, $\left| \frac{1}{f(x)} - \frac{1}{f(\theta)} \right| \leq \varepsilon$.

Conclusion

un tel $\alpha > 0$ existe.

- (ii) Soit $\omega \in [|T_n - \theta| \leq \alpha] \cap E_n$.

Si $\bar{Y}_n(\omega) = F(\theta)$, alors $U_n(\omega) = \frac{1}{f(\theta)}$ et $\omega \in B_n(\varepsilon)$.

Sinon, comme $\omega \in E_n$, $t = T_n(\omega) = F^{-1}(\bar{Y}_n(\omega))$ vérifie $F(t) = \bar{Y}_n(\omega) \neq F(\theta)$, donc $t \neq \theta$. La fonction F est dérivable de dérivée f ; par l'égalité des accroissements finis, il

existe c_ω strictement compris entre θ et t tel que $F(t) - F(\theta) = f(c_\omega)(t - \theta)$, avec $f(c_\omega) > 0$. Donc :

$$U_n(\omega) = \frac{t - \theta}{F(t) - F(\theta)} = \frac{1}{f(c_\omega)}$$

Comme $|c_\omega - \theta| \leq |t - \theta| \leq \alpha$, (i) donne $\left| U_n(\omega) - \frac{1}{f(\theta)} \right| \leq \varepsilon$.

Conclusion

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

b) Soit $\varepsilon > 0$ et α comme en a). En passant aux contraires dans a) (ii) :

$$\overline{B_n(\varepsilon)} \subset [|T_n - \theta| > \alpha] \cup \overline{E_n}. \text{ Donc :}$$

$$P_\theta \left(\left[\left| U_n - \frac{1}{f(\theta)} \right| > \varepsilon \right] \right) \leq P_\theta (\overline{B_n(\varepsilon)}) \leq P_\theta ([|T_n - \theta| > \alpha]) + 1 - P_\theta(E_n)$$

Le premier terme tend vers 0 par 13) d), le second par 13) a).

Conclusion

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

c) Question adaptée au programme actuel : la conclusion directe, qui reposerait sur le théorème de Slutsky (produit d'une suite convergeant en loi par une suite convergeant en probabilité vers une constante), est remplacée par une démonstration guidée par encadrement des fonctions de répartition.

(i) Soit $\omega \in \Omega$. Si $\bar{Y}_n(\omega) \neq F(\theta)$, par définition de U_n :

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

Si $\bar{Y}_n(\omega) = F(\theta)$, alors $W_n(\omega) = 0$; de plus $\bar{Y}_n(\omega) \in]0, 1[$, donc $\omega \in E_n$ et $T_n(\omega) = F^{-1}(F(\theta)) = \theta$, si bien que les deux membres sont nuls.

Conclusion

$$\sqrt{n} (T_n - \theta) = U_n W_n.$$

(ii) Sur $B_n(\varepsilon)$, on a $0 < c - \varepsilon \leq U_n \leq c + \varepsilon$.

Majoration. Soit $\omega \in [U_n W_n \leq x] \cap B_n(\varepsilon)$. Si $W_n(\omega) \geq 0$, alors

$(c - \varepsilon)W_n(\omega) \leq U_n(\omega)W_n(\omega) \leq x$, donc $W_n(\omega) \leq \frac{x}{c - \varepsilon} \leq M_\varepsilon$. Si $W_n(\omega) < 0$, alors

$(c + \varepsilon)W_n(\omega) \leq U_n(\omega)W_n(\omega) \leq x$, donc $W_n(\omega) \leq \frac{x}{c + \varepsilon} \leq M_\varepsilon$. Ainsi

$[U_n W_n \leq x] \cap B_n(\varepsilon) \subset [W_n \leq M_\varepsilon]$, et :

$$\begin{aligned} P_\theta([U_n W_n \leq x]) &\leq P_\theta([U_n W_n \leq x] \cap B_n(\varepsilon)) + P_\theta(\overline{B_n(\varepsilon)}) \leq P_\theta([W_n \leq M_\varepsilon]) \\ &\quad + P_\theta(\overline{B_n(\varepsilon)}) \end{aligned}$$

Minoration. Soit $\omega \in [W_n \leq m_\varepsilon] \cap B_n(\varepsilon)$. Si $W_n(\omega) \geq 0$, alors

$U_n(\omega)W_n(\omega) \leq (c + \varepsilon)W_n(\omega) \leq (c + \varepsilon)m_\varepsilon \leq x$ car $m_\varepsilon \leq \frac{x}{c + \varepsilon}$. Si $W_n(\omega) < 0$, alors

$U_n(\omega)W_n(\omega) \leq (c - \varepsilon)W_n(\omega) \leq (c - \varepsilon)m_\varepsilon \leq x$ car $m_\varepsilon \leq \frac{x}{c - \varepsilon}$. Ainsi

$[W_n \leq m_\varepsilon] \cap B_n(\varepsilon) \subset [U_n W_n \leq x]$, et :

$$P_\theta([U_n W_n \leq x]) \geq P_\theta([W_n \leq m_\varepsilon] \cap B_n(\varepsilon)) \geq P_\theta([W_n \leq m_\varepsilon]) - P_\theta(\overline{B_n(\varepsilon)})$$

Conclusion

l'encadrement demandé est établi.

(iii) Notons G la fonction de répartition de la loi $\mathcal{N}(0, \sigma^2)$: $G(t) = \Phi\left(\frac{t}{\sigma}\right)$, continue sur $\mathbb{R}$. D'après 12), $P_\theta([W_n \leq t]) \rightarrow G(t)$ pour tout réel t .

Soit $x \in \mathbb{R}$ et $\eta > 0$. Lorsque $\varepsilon \rightarrow 0^+$, $\frac{x}{c-\varepsilon}$ et $\frac{x}{c+\varepsilon}$ tendent vers $\frac{x}{c}$, donc m_ε et M_ε aussi. Par continuité de G en $\frac{x}{c}$, on peut fixer $\varepsilon \in]0, c[$ tel que $G(M_\varepsilon) \leq G\left(\frac{x}{c}\right) + \eta$ et $G(m_\varepsilon) \geq G\left(\frac{x}{c}\right) - \eta$. Pour ce ε fixé, $P_\theta([W_n \leq M_\varepsilon]) \rightarrow G(M_\varepsilon)$, $P_\theta([W_n \leq m_\varepsilon]) \rightarrow G(m_\varepsilon)$ et, d'après la preuve de b), $P_\theta(\overline{B_n(\varepsilon)}) \rightarrow 0$. Il existe donc n_0 tel que, pour $n \geq n_0$, ces trois quantités sont à moins de η de leurs limites. Avec (i) et (ii), pour $n \geq n_0$:

$$G\left(\frac{x}{c}\right) - 3\eta \leq P_\theta([\sqrt{n}(T_n - \theta) \leq x]) \leq G\left(\frac{x}{c}\right) + 3\eta$$

Ainsi $P_\theta([\sqrt{n}(T_n - \theta) \leq x]) \rightarrow G\left(\frac{x}{c}\right) = \Phi\left(\frac{x}{c\sigma}\right)$ pour tout réel x , qui est la fonction de répartition de la loi $\mathcal{N}(0, c^2\sigma^2)$.

Conclusion

$(\sqrt{n}(T_n - \theta))$ converge en loi vers une variable de loi normale centrée de variance $\frac{F(\theta)(1-F(\theta))}{f(\theta)^2}$.

Partie IV : Régression logistique

15) a) Pour $Y \in \mathcal{M}_{m,1}(\mathbb{R})$, notons $\|Y\|^2 = {}^t Y Y$, carré de la norme euclidienne. La matrice $M {}^t M$ est carrée d'ordre p . Soit $X \in \mathcal{M}_{p,1}(\mathbb{R})$ tel que $M {}^t M X = 0$. Alors :

$${}^t X M {}^t M X = {}^t ({}^t M X) ({}^t M X) = \|{}^t M X\|^2 = 0$$

donc ${}^t M X = 0$. Or $\text{rg}({}^t M) = \text{rg}(M) = p$ et ${}^t M \in \mathcal{M}_{k,p}(\mathbb{R})$ a p colonnes : ses colonnes

forment une famille libre, donc ${}^tMX = 0$ impose $X = 0$. Le noyau de $M {}^tM$ est réduit à $\{0\}$.

Conclusion

$M {}^tM$ est inversible. (Au passage : ${}^tMX \neq 0$ pour tout $X \neq 0$.)

b) Notons $G = M {}^tM$ (symétrique, inversible), $U_0 = G^{-1}MH$ et, pour $U \in \mathcal{M}_{p,1}(\mathbb{R})$, $\varphi(U) = \|{}^tMU - H\|^2$, qui est l'unique coefficient de la matrice de l'énoncé.

Méthode 1 — Par orthogonalité (Pythagore).

On a $M ({}^tMU_0 - H) = GU_0 - MH = 0$. Pour tout U , en posant $D = U - U_0$: ${}^tMU - H = {}^tMD + ({}^tMU_0 - H)$, et ces deux vecteurs sont orthogonaux car ${}^t({}^tMD) ({}^tMU_0 - H) = {}^tDM ({}^tMU_0 - H) = 0$. Par le théorème de Pythagore :

$$\varphi(U) = \|{}^tMD\|^2 + \varphi(U_0) \geq \varphi(U_0)$$

avec égalité si et seulement si ${}^tMD = 0$, c'est-à-dire $D = 0$ d'après a).

Méthode 2 — Par le gradient et la hessienne.

En développant (le réel ${}^tH {}^tMU$ est égal à son transposé ${}^tU MH$) :

$\varphi(U) = {}^tUGU - 2 {}^tU MH + {}^tHH$. C'est une fonction polynomiale des coordonnées $u_1, \dots, u_p$ de U , donc de classe C^2 sur l'ouvert convexe $\mathbb{R}^p$. Comme G est symétrique, son gradient est $\nabla\varphi(U) = 2(GU - MH)$ et sa hessienne est la matrice constante $2G$. Le gradient s'annule si et seulement si $GU = MH$, c'est-à-dire $U = U_0$: c'est l'unique point critique. Si $GX = \mu X$ avec $X \neq 0$, alors $\mu\|X\|^2 = {}^tXGX = \|{}^tMX\|^2 > 0$ d'après a), donc $\mu > 0$: la hessienne a toutes ses valeurs propres strictement positives en tout point de l'ouvert convexe $\mathbb{R}^p$. Par le critère d'extremum global, φ admet en U_0 un minimum

global. Il est unique, car tout point de minimum global sur l'ouvert $\mathbb{R}^p$ est un point critique, donc égal à U_0 .

La méthode 1 est géométrique : tMU_0 est la projection orthogonale de H sur l'image de tM , ce qui explique le résultat ; la méthode 2 est plus calculatoire mais s'applique à toute fonction de classe C^2 , sans structure euclidienne apparente.

Conclusion

le minimum de $\|{}^tMU - H\|^2$ est atteint en une unique matrice, $U = (M {}^tM)^{-1}MH$.

- c) La loi de $Y_{i,n}$ est la loi de Bernoulli de paramètre $\Lambda(\langle \alpha, x^{(i)} \rangle)$. Comme Λ est injective, les lois des $Y_{i,n}$ déterminent exactement les k réels $\langle \alpha, x^{(i)} \rangle$, qui sont les coefficients de la matrice colonne tMA , et rien de plus.

Si $\text{rg}(M) < p$, alors $\text{rg}({}^tM) < p$ et, par le théorème du rang appliqué à l'application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^k$ de matrice tM , son noyau n'est pas réduit à $\{0\}$. Pour $V \neq 0$ dans ce noyau, le vecteur α' de matrice $A' = A + V \neq A$ vérifie ${}^tMA' = {}^tMA$: les paramètres α et α' donnent aux $Y_{i,n}$ exactement les mêmes lois.

Conclusion

si $\text{rg}(M) < p$, plusieurs vecteurs α distincts sont compatibles avec les mêmes lois, qui ne permettent donc pas de définir α . (Lorsque $\text{rg}(M) = p$, au contraire, tMA détermine $A = (M {}^tM)^{-1}M ({}^tMA)$.)

- 16) a) Fixons $i \in \llbracket 1, k \rrbracket$ et posons $\theta_i = \langle \alpha, x^{(i)} \rangle$. On applique la partie III avec $F = \Lambda$: d'après 3) a), Λ est la fonction de répartition de la loi logistique standard, de densité λ continue et strictement positive sur $\mathbb{R}$. Les variables $Y_{i,n}$, $n \in \mathbb{N}^*$, sont indépendantes de loi de Bernoulli de paramètre $\Lambda(\theta_i)$, et $\Lambda^{-1} = L$ d'après 1) a) : la variable $T_{i,n}$ est exactement la variable T_n de la question 13) associée à $\theta = \theta_i$. D'après 13) d), $(T_{i,n})_n$ converge en probabilité vers θ_i .

Par composition avec la fonction continue $t \mapsto c_i t$, $(c_i T_{i,n})_n$ converge en probabilité vers $c_i \theta_i$. La somme de deux suites convergeant en probabilité converge en probabilité vers la

somme des limites ; par récurrence sur le nombre de termes,

$$\left(\sum_{i=1}^k c_i T_{i,n} \right)_n$$

converge en probabilité vers

$$\sum_{i=1}^k c_i \theta_i.$$

Enfin

$$\sum_i c_i T_{i,n}$$

ne dépend que des observations $Y_{i,j}$ ($1 \leq i \leq k, 1 \leq j \leq n$) et non de α : c'est un estimateur.

Conclusion

$$\left(\sum_{i=1}^k c_i T_{i,n} \right)_{n \in \mathbb{N}^*}$$

est une suite convergente d'estimateurs de

$$\sum_{i=1}^k c_i \langle \alpha, x^{(i)} \rangle.$$

b) Notons $Q = (M^t M)^{-1} M \in \mathcal{M}_{p,k}(\mathbb{R})$, de coefficients $q_{j,i}$. Pour $j \in \llbracket 1, p \rrbracket$:

$$A_{j,n} = \sum_{i=1}^k q_{j,i} T_{i,n}.$$

D'après a) avec $c_i = q_{j,i}$, $(A_{j,n})_n$ est une suite convergente d'estimateurs de

$$\sum_{i=1}^k q_{j,i} \theta_i.$$

Or

$$\theta_i = \sum_{l=1}^p x_l^{(i)} \alpha_l$$

est le i -ème coefficient de ${}^t M A$. Donc

$$\sum_{i=1}^k q_{j,i} \theta_i$$

est le j -ème coefficient de :

$$Q {}^t M A = (M {}^t M)^{-1} M {}^t M A = A$$

c'est-à-dire α_j .

Interprétation : d'après 15) b), $(A_{1,n}, \dots, A_{p,n})$ est l'estimation des moindres carrés de α obtenue à partir des estimations $T_{i,n}$ des $\langle \alpha, x^{(i)} \rangle$.

Conclusion

pour tout $j \in \llbracket 1, p \rrbracket$, $(A_{j,n})_{n \in \mathbb{N}^*}$ est une suite convergente d'estimateurs de α_j .

Relevé des adaptations

- 2) : Code adapté en Python. Le script Scilab de dichotomie devient un code Python avec les mêmes blancs ; les lignes sont numérotées en commentaire, d'où la renumérotation des lignes citées (lignes (7), (2), (3), (4) et (10) d'origine devenues (10) et (12), (4), (5), (6) et (14)). Justification : seul Python est autorisé.
- 4) b) et c) : Code adapté en Python. La fonction Scilab grandlogis devient une fonction Python grandlogis(n, p, r, s) renvoyant un tableau à n lignes et p colonnes ; l'énoncé rappelle `rd.random((n, p))` et le calcul coefficient par coefficient. Justification : seul Python est autorisé ; ces syntaxes doivent être expliquées.
- 5) a) : Question adaptée au programme actuel. L'indication « changement de variable exponentiel » est précisée : densités de $\ln(U_1)$ et $-\ln(U_2)$, produit de convolution, changement de variable $u = e^t$. Justification : le référentiel exclut le calcul par densité conjointe ou intégrale double, et le changement de variable non affine doit être indiqué.
- 5) b) : Code adapté en Python. La fonction Scilab grand est remplacée par `rd.exponential(1)`, dont la convention de paramètre est rappelée. Justification : seul Python est autorisé, et le paramètre scale de NumPy doit être précisé.
- Préambule de la partie II et 6) c) : Question adaptée au programme actuel. Les définitions de la partie imaginaire, de la factorisation sur les complexes et de la somme des racines complexes sont supprimées ; 6) c) suppose que P_n admet n racines réelles distinctes et demande la factorisation réelle puis la somme de ces racines. Justification : aucun nombre complexe n'est autorisé.
- 7) a) : Question adaptée au programme actuel. L'égalité obtenue dans l'énoncé d'origine par un calcul sur les nombres complexes est remplacée par une récurrence sur les formules de $\cos(mx)$ et $\sin(mx)$, utilisant les formules d'addition. Justification : aucun nombre complexe n'est autorisé.

- 7) c) : référence à 6) c) et 7) b) précisée, pour que la question montre que les $\frac{1}{\sin^2\left(\frac{k\pi}{2n+1}\right)}$ sont n racines réelles distinctes de P_n . Justification : cohérence avec la nouvelle version de 6) c).
- 13) b) (i) : Question adaptée au programme actuel. «Montrer que $[T_n \leq x]$ est un élément de la tribu $\mathcal{A}$ » est remplacé par «On admet que T_n est une variable aléatoire». Justification : pas de théorie des tribus ni de mesurabilité.
- 14) c) : Question adaptée au programme actuel. La question est découpée en (i), (ii), (iii) : identité $\sqrt{n}(T_n - \theta) = U_n W_n$, encadrement des fonctions de répartition, puis conclusion. Justification : la conclusion directe relèverait du théorème de Slutsky, exclu.