LSTAT 2040 - TP 5 : Solutions

Méthodes d'estimation et propriétés

Exercice 1

(a) Un calcul simple montre que $\mu_1 = \mathrm{E}[X_1] = \theta/(\theta+1)$. L'inversion de cette relation donne

$$\theta = \frac{\mu_1}{1-\mu_1} =: g(\mu_1),$$

où $g : x \in (0,1) \mapsto x/(1-x) \in \mathbb{R}$. On déduit que l'estimateur des moments de θ est donné par

$$\hat{\theta}_{\mathrm{MM}} := g(\overline{X}_n) = \frac{\overline{X}_n}{1-\overline{X}_n}.$$

(b) Notons que, par la loi des grands nombres,

$$\overline{X}_n \xrightarrow{p} \mu_1 = \frac{\theta}{\theta+1}.$$

Observons également que la fonction g définie au point (a) est continue. Il s'ensuit du CMT que

$$\hat{\theta}_{\mathrm{MM}} = g(\overline{X}_n) \xrightarrow{p} g(\mu_1) = \theta.$$

(c) Notons que, par le CLT,

$$\sqrt{n}(\overline{X}_n - \mu_1) \xrightarrow{d} \mathrm{N}(0, \mathrm{Var}[X_1]) = \mathrm{N}\left(0, \frac{\theta}{(\theta+2)(\theta+1)^2}\right).$$

Puisque la fonction g définie au point (a) est différentiable au point $\mu_1 \in (0,1)$, on a par la méthode delta,

$$\sqrt{n}(\hat{\theta}_{\mathrm{MM}} - \theta) = \sqrt{n}(g(\overline{X}_n) - g(\mu_1)) \xrightarrow{d} g'(\mu_1)\mathrm{N}\left(0, \frac{\theta}{(\theta+2)(\theta+1)^2}\right) = \mathrm{N}\left(0, \frac{\theta(\theta+1)^2}{\theta+2}\right).$$

Exercice 2

(a) Commençons par écrire la log-vraisemblance de l'échantillon. Celle-ci est donnée par

$$\ell_n(\theta) = \sum_{i=1}^{n} \log f(X_i;\theta),$$

où $f(\cdot;\theta)$ est la fonction de densité associée à la distribution $\mathrm{N}(\theta,\theta)$. Un calcul simple montre que

$$\ell_n(\theta) = -\frac{n}{2}\left(\log(2\pi) + \log(\theta)\right) - \frac{1}{2\theta}\sum_{i=1}^{n}(X_i-\theta)^2.$$

Afin de maximiser cette fonction par rapport à $\theta \in (0,\infty)$, on commence par chercher les solutions positives de l'équation de premier ordre donnée par

$$\partial_\theta \ell_n(\theta) = 0.$$

On vérifie que cette équation est équivalente à l'équation de second ordre en θ donnée par

$$\theta^2 + \theta - \overline{X^2}_n = 0.$$

L'unique solution positive est donnée par

$$\theta_0 = \frac{-1 + \sqrt{1 + 4\overline{X^2}_n}}{2}.$$

De plus, on vérifie sans trop de difficultés que

$$\partial_\theta^2 \ell_n(\theta_0) < 0.$$

Cela implique que θ_0 est un maximum local de log-vraisemblance, et, au vue du théorème énoncé dans la note associée à ce TP, ce maximum est également global. On déduit que

$$\hat{\theta}_{\text{MLE}} = \theta_0 = \frac{-1 + \sqrt{1 + 4\overline{X^2}_n}}{2}.$$

(b) Par le CLT, on a

$$\sqrt{n}(\overline{X^2}_n - \mathrm{E}[X_1^2]) \xrightarrow{d} \mathrm{N}(0, \mathrm{Var}[X_1^2]).$$

En notant que $\mathrm{E}[X_1^2] = \mathrm{Var}[X_1] + \mathrm{E}[X_1]^2 = \theta + \theta^2$ et $\mathrm{E}[X_1^4] = \theta^4 + 6\theta^3 + 3\theta^2$, on trouve

$$\sqrt{n}(\overline{X^2}_n - (\theta + \theta^2) \xrightarrow{d} \mathrm{N}(0, 4\theta^3 + 2\theta^2).$$

Ensuite, on remarque que le MLE obtenu au point précédent satisfait

$$\hat{\theta}_{\text{MLE}} = h(\overline{X^2}_n),$$

où $h : x \in (0, \infty) \mapsto h(x) := (-1 + \sqrt{1 + 4x})/2 \in (0, \infty)$. La fonction h est différentiable $\theta + \theta^2$ et il suit de la méthode delta que

$$\sqrt{n}(\hat{\theta}_{\text{MLE}} - \theta) = \sqrt{n}(h(\overline{X^2}_n) - h(\theta + \theta^2)) \xrightarrow{d} h'(\theta + \theta^2)\mathrm{N}(0, 4\theta^3 + 2\theta^2) = \mathrm{N}\left(0, \frac{2\theta^2}{2\theta + 1}\right).$$

(c) L'information de Fisher contenue dans l'échantillon est donnée par $I_n(\theta) = nI(\theta)$ avec

$$I(\theta) = -\mathrm{E}[\partial_\theta^2 \log f(X_1; \theta)].$$

Le calcul du membre de droite donne

$$I(\theta) = \frac{2\theta + 1}{2\theta^2}.$$

(d) La variance asymptotique du MLE $\hat{\theta}_{\text{MLE}}$ est égale à l'inverse de l'information de Fisher. Cela implique que le MLE $\hat{\theta}_{\text{MLE}}$ est asymptotiquement efficace pour θ.

Exercice 3

Notons que la vraisemblance d'un échantillon iid $X_1, \ldots, X_n$ où $X_1 \sim \mathrm{Unif}[\theta, \theta + 2]$ est donnée par

$$L_n(\theta) = \prod_{i=1}^n \frac{1}{2}\mathrm{I}(\theta \leq X_i \leq \theta + 2) = \frac{1}{2^n}\mathrm{I}(X_{(n)} - 2 \leq \theta \leq X_{(1)}).$$

Ici, on a $X_{(1)} = 2.3$ et $X_{(n)} - 2 = 1.9$. On sait donc qu'en dehors de l'intervalle $[1.9, 2.3]$, la vraisemblance de notre échantillon vaut zéro. Le seul des trois estimateurs proposés à rendre celle-ci strictement positive est alors donné par $\hat{\theta}_2 = 1.94$, qui est le seul candidat pour être le MLE. Notons que, puisque la vraisemblance est constante sur $[1.9, 2.3]$, n'importe quel $\theta \in [1.9, 2.3]$ correspondrait au MLE. Ce dernier n'est donc pas unique.

Exercice 4

(a) Puisque les $X_i, i = 1, \ldots, n$ suivent une distribution uniforme discrète sur $\{1, \ldots, \theta\}$, on a $\mu_1 = \mathrm{E}[X_1] = (1+\theta)/2$. L'inversion de cette relation donne

$$\theta = 2\mu_1 - 1 =: g(\mu_1),$$

où $g : x \in \mathbb{R} \mapsto 2x - 1 \in \mathbb{R}$. On déduit que l'estimateur des moments de θ est donné par

$$\hat{\theta}_{\mathrm{MM}} := g(\overline{X}_n) = 2\overline{X}_n - 1.$$

(b) Puisque la fonction de masse commune des $X_i, i = 1, \ldots, n$ est donnée par $f(\cdot;\theta)$ où

$$f(x;\theta) = \frac{1}{\theta}\mathrm{I}(x \in \{1, \ldots, \theta\}),$$

la vraisemblance de l'échantillon est donnée par

$$L_n(\theta) = \theta^{-n}\mathrm{I}(\theta \geq X_{(n)}).$$

La vraisemblance est donc nulle pour toute valeur de θ plus petite que le maximum des $X_i, i = 1, \ldots, n$ et est décroissante pour les valeurs supérieures à $X_{(n)}$. On déduit que $\hat{\theta}_{\mathrm{MLE}} = X_{(n)}$.

(c) On calcule, pour cet échantillon

$$\hat{\theta}_{\mathrm{MM}} = 9.5 \qquad \text{et} \qquad \hat{\theta}_{\mathrm{MLE}} = 12.$$

L'estimateur du maximum de vraisemblance est meilleur que l'estimateur des moments pour deux raisons. La première est que l'estimateur des moments est plus petit que le maximum des boules obtenues. Or, θ correspond au plus grand numéro observable sur les boules de l'urne. Par conséquent, on est certain que θ est au minimum $x_4 = 12$. La seconde raison est que l'estimateur des moments n'est pas un nombre naturel, et θ est entier, par définition. Le MLE quant à lui ne souffre d'aucun de ces deux défauts.

Exercice 5

Supposons avoir un échantillon iid $X_1, \ldots, X_n$ dont la distribution commune a pour densité $f(\cdot;\theta)$ où l'on sait que $f(x;\theta) = 0$ si $x \notin [-\theta, \theta]$. Par exemple, notons, sans perte de généralité

$$f(x;\theta) = g(x;\theta)\mathrm{I}(-\theta \leq x \leq \theta).$$

Il est aisé de voir que la vraisemblance de l'échantillon peut s'écrire

$$L_n(\theta) = \prod_{i=1}^{n} g(X_i;\theta)\mathrm{I}(\theta \geq \max\{-X_{(1)}, X_{(n)}\}).$$

Par conséquent, pour l'échantillon considéré dans l'exercice, on sait que la vraisemblance doit être nulle pour $\theta < \max\{-(-10), 1\} = 10$. Le seul graphique satisfaisant cette condition est le graphe B.

Exercice 6

(a) Observons que

$$\mu_1 = \mathrm{E}[X_1] = k\theta \qquad \text{et} \qquad \mu_2 = \mathrm{E}[X_1^2] = k(k+1)\theta^2.$$

Une inversion de ces relations donne, après quelques manipulations algébriques élémentaires,

$$(k, \theta) = g(\mu_1, \mu_2),$$

avec $g : (x, y) \in D := \{(u, v) \in (0, \infty)^2 : v > u^2\} \mapsto g(x, y) := (x^2/(y - x^2), (y - x^2)/x)) \in (0, \infty)^2$. Il s'ensuit que l'estimateur des moments de (k, θ) est donné par

$$(\hat{k}_{\mathrm{MM}}, \hat{\theta}_{\mathrm{MM}}) = g(\overline{X}_n, \overline{X^2}_n) = \left(\frac{\overline{X}_n^2}{\overline{X^2}_n - \overline{X}_n^2}, \frac{\overline{X^2}_n - \overline{X}_n^2}{\overline{X}_n} \right).$$

(b) Commençons par constater qu'il suit du CLT multivarié que

$$\sqrt{n}\left\{(\overline{X}_n, \overline{X^2}_n) - (k\theta, k(k+1)\theta^2)\right\} \xrightarrow{d} \mathrm{N}_2((0,0),\Sigma),$$

avec

$$\Sigma = \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_1^2] \\ \mathrm{Cov}[X_1, X_1^2] & \mathrm{Var}[X_1^2] \end{pmatrix}.$$

En utilisant le fait que

$$\mu_3 = \mathrm{E}[X_1^3] = k(k+1)(k+2)\theta^3 \qquad \text{et} \qquad \mu_4 = \mathrm{E}[X_1^4] = k(k+1)(k+2)(k+3)\theta^4,$$

on peut vérifier, après quelques vérifications algébriques élémentaires, que la variance asymptotique Σ est donnée par

$$\Sigma = \begin{pmatrix} k\theta^2 & 2k(k+1)\theta^3 \\ 2k(k+1)\theta^3 & 2k(k+1)(2k+3)\theta^4 \end{pmatrix}.$$

Ensuite, observons que la fonction g définie au point (a) est différentiable en $(k\theta, k(k+1)\theta^2) \in D$. Ainsi, il suit de la méthode delta bivariée que

$$\begin{aligned}\sqrt{n}\left\{(\hat{k}_{\mathrm{MM}}, \hat{\theta}_{\mathrm{MM}}) - (k,\theta)\right\} &= \sqrt{n}\left\{g(\overline{X}_n, \overline{X^2}_n) - g(k\theta, k(k+1)\theta^2)\right\} \\ &\xrightarrow{d} \mathrm{N}_2\left((0,0), \nabla g(k\theta, k(k+1)\theta^2)\Sigma\nabla g(k\theta, k(k+1)\theta^2)^T\right),\end{aligned}$$

où

$$\nabla g(x,y) = \begin{pmatrix} \frac{2xy}{(y-x^2)^2} & \frac{-1}{(y-x^2)} \\ \frac{-(x^2+y)}{x^2} & \frac{1}{x} \end{pmatrix}.$$

est la matrice jacobienne de g évaluée en $(x,y) \in D$. L'évaluation de cette matrice en $(x,y) = (k\theta, k(k+1)\theta^2) \in D$ et le calcul du produit matriciel apparaissant dans la variance asymptotique de l'estimateur des moments est laissé au soin du lecteur et n'est pas recommandé par l'auteur.

Exercice 7

(a) On observe, après calcul, que la log-vraisemblance de l'échantillon est donnée par

$$\ell_n(\theta) = \sum_{i=1}^n \log f(X_i;\theta) = n\log(2) - 2n\log(\theta) + \sum_{i=1}^n \log(X_i) - \frac{1}{\theta^2}\sum_{i=1}^n X_i^2.$$

Afin de maximiser cette fonction par rapport à $\theta > 0$, on cherche toutes les solutions de l'équation de premier ordre donnée par

$$\partial_\theta \ell_n(\theta) = 0.$$

On vérifie facilement que celle-ci possède une unique solution donnée par

$$\theta_0 := \sqrt{\overline{X^2}_n}.$$

De plus, on vérifie aisément que

$$\partial_\theta^2 \ell_n(\theta_0) < 0,$$

ce qui démontre que θ_0 est un maximum local de log-vraisemblance. Par le théorème de la note associée à ce TP, ce maximum est en réalité global et l'on conclut que

$$\hat{\theta}_{\mathrm{MLE}} = \theta_0 = \sqrt{\overline{X^2}_n}.$$

(b) La fonction $g : x \in (0,\infty) \mapsto g(x) := x^2 \in (0,\infty)$ étant bijective, le MLE de $\theta^2 = g(\theta)$ est donné par

$$g(\hat{\theta}_{\mathrm{MLE}}) = \overline{X^2}_n.$$

(c) Il suit de la loi des grands nombres que

$$\overline{X^2}_n \xrightarrow{p} \mathrm{E}[X_1^2] = \theta^2.$$

Ainsi, par le CMT appliqué avec g^{-1} où g est la fonction bijective du point (c) qui est en réalité un homéomorphisme, on a

$$\hat{\theta}_{\mathrm{MLE}} = g^{-1}(\overline{X^2}_n) \xrightarrow{p} g^{-1}(\theta^2) = \theta.$$

(d) Il suit du CLT que

$$\sqrt{n}(\overline{X^2}_n - \theta^2) \xrightarrow{d} \mathrm{N}(0, \mathrm{Var}[X_1^2]) = \mathrm{N}(0, \theta^4).$$

Par la méthode delta avec la fonction g^{-1} qui est en réalité un difféomorphisme, dont la dérivée en $x > 0$ est donnée par $(g^{-1})'(x) = 1/(2\sqrt{x})$, on trouve

$$\sqrt{n}(\hat{\theta}_{\mathrm{MLE}} - \theta) = \sqrt{n}(g^{-1}(\overline{X^2}_n) - g^{-1}(\theta^2)) \xrightarrow{d} (g^{-1})'(\theta^2)\mathrm{N}(0, \theta^4) = \mathrm{N}\left(0, \frac{\theta^2}{4}\right).$$

Exercice 8

(a) La vraisemblance d'un échantillon de taille $n \in \mathrm{N}$ est donnée par

$$L_n(\theta) = \prod_{i=1}^n \frac{1}{2\theta+1} \mathrm{I}(0 \leq X_i \leq 2\theta + 1) = (2\theta+1)^{-n} = \mathrm{I}\left(\theta \geq \frac{X_{(n)} - 1}{2}\right).$$

Ainsi, la vraisemblance vaut zéro si $\theta < (X_{(n)} - 1)/2$ et est décroissante pour les valeurs de θ pour lesquelles celle-ci ne s'annule pas. On conclut que

$$\hat{\theta}_{\mathrm{MLE}} = \frac{X_{(n)} - 1}{2}.$$

(b) La variance commune des $(X_i)_{i \in \mathrm{N}}$ est donnée par

$$\mathrm{Var}[X_1] = \frac{(2\theta+1)^2}{12} = g(\theta),$$

avec $g : x \in (0, \infty) \mapsto g(x) := (2x+1)^2/12 \in (0, \infty)$. Notons que g est bijective. Par conséquent, le MLE de la variance des $(X_i)_{i \in \mathrm{N}}$ est donnée par

$$g(\hat{\theta}_{\mathrm{MLE}}) = \frac{X_{(n)}^2}{12}.$$

Exercice 9

(a) Si Z désigne le nombre de boules bleues tirées dans l'urne, on a que le vecteur aléatoire (X, Y, Z) est distributée selon une loi multinomiale $\mathrm{Mult}_3(N, \pi)$ avec $\pi = (\pi_1, \pi_2, 1-\pi_1-\pi_2)$. Puisque $Z = N - X - Y$, on peut omettre la dernière composante dans la fonction de vraisemblance de l'échantillon et écrire

$$L_n(\pi_1, \pi_2) = \frac{(N!)^n}{\prod_{i=1}^n (X_i! Y_i! (N - X_i - Y_i)!)} \pi_1^{\sum_{i=1}^n X_i} \pi_2^{\sum_{i=1}^n Y_i} (1 - \pi_1 - \pi_2)^{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}.$$

Il suit alors que, à des constantes près par rapport aux paramètres, la log-vraisemblance est donnée par

$$\ell_n(\pi_1, \pi_2) \equiv \sum_{i=1}^n X_i \log(\pi_1) + \sum_{i=1}^n Y_i \log(\pi_2) + (nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i) \log(1 - \pi_1 - \pi_2).$$

Les conditions de première ordre associées peuvent alors s'écrire comme

$$\partial_{\pi_1} \ell_n(\pi_1, \pi_2) = \frac{\sum_{i=1}^n X_i}{\pi_1} - \frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{1 - \pi_1 - \pi_2} = 0$$
$$\partial_{\pi_2} \ell_n(\pi_1, \pi_2) = \frac{\sum_{i=1}^n Y_i}{\pi_2} - \frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{1 - \pi_1 - \pi_2} = 0.$$

La résolution de ce système donne une unique solution donnée par

$$(\pi_{1,0}, \pi_{2,0}) := \left(\frac{\overline{X}_n}{N}, \frac{\overline{Y}_n}{N}\right).$$

De plus, on vérifie aisément que la matrice Hessienne associée à la log-vraisemblance est donnée par

$$\nabla^2 \ell_n(\pi_1, \pi_2) = \begin{pmatrix} -\frac{\sum_{i=1}^n X_i}{\pi_1^2} - \frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{(1-\pi_1-\pi_2)^2} & -\frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{(1-\pi_1-\pi_2)^2} \\ -\frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{(1-\pi_1-\pi_2)^2} & -\frac{\sum_{i=1}^n Y_i}{\pi_2^2} - \frac{nN - \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{(1-\pi_1-\pi_2)^2} \end{pmatrix}.$$

Cette matrice est définie négative pour toute valeur de (π_1, π_2) puisque les dérivées sécondes non-croisées sont toutes deux strictement négatives et le déterminant est strictement positif. Cela permet de conclure que la log-vraisemblance est une fonction strictement concave et donc que le point stationnaire $(\pi_{1,0}, \pi_{2,0})$ trouvé grâce à la condition de premier ordre est l'unique maximum global de la log-vraisemblance et donc

$$(\hat{\pi}_{1\text{MLE}}, \hat{\pi}_{2\text{MLE}}) = (\pi_{1,0}, \pi_{2,0}) = \left(\frac{\overline{X}_n}{N}, \frac{\overline{Y}_n}{N}\right).$$

(b) Il suit du CLT multivarié que

$$\sqrt{n}\left\{(\overline{X}_n, \overline{Y}_n) - (N\pi_1, N\pi_2)\right\} \xrightarrow{d} \mathrm{N}_2((0,0), \Sigma),$$

où

$$\Sigma = \begin{pmatrix} N\pi_1(1-\pi_1) & -N\pi_1\pi_2 \\ -N\pi_1\pi_2 & N\pi_2(1-\pi_2) \end{pmatrix},$$

où l'on a utilisé le fait que $X \sim \mathrm{Bin}(N, \pi_1)$, $Y \sim \mathrm{Bin}(N, \pi_2)$ et $\mathrm{Cov}[X, Y] = -N\pi_1\pi_2$ (voir formulaire). La méthode delta garantit alors que

$$\sqrt{n}\left\{(\hat{\pi}_{1\text{MLE}}, \hat{\pi}_{2\text{MLE}}) - (\pi_1, \pi_2)\right\} \xrightarrow{d} \mathrm{N}_2((0,0), \Sigma/N^2).$$

Exercice 10

La log-vraisemblance d'un échantillon de taille $n \in \mathrm{N}$ est donné par

$$\ell_n(\mu, \sigma) = -n\log(2) - n\log(\sigma) - \frac{1}{\sigma}\sum_{i=1}^n |X_i - \mu|.$$

On ne peut clairement pas espérer maximiser cette fonction de manière conjointe sur $(\mu, \sigma) \in \mathbb{R} \times (0, \infty)$ puisque celle-ci n'est pas différentiable en μ à cause de la valeur absolue. Cependant, on peut procéder par maximisation de la log-vraisemblance profilée en σ donnée par

$$\ell_n^P(\sigma) = \ell_n(\hat{\mu}(\sigma), \sigma),$$

où $\hat{\mu}(\sigma)$ est la valeur de μ qui maximise $\ell_n(\mu, \sigma)$ pour n'importe quel valeur de σ. Il est aisé de voir que

$$\hat{\mu}(\sigma) = \operatorname*{argmin}_{\mu \in \mathbb{R}} \sum_{i=1}^n |X_i - \mu|,$$

peu importe la valeur de $\sigma > 0$. La solution de ce problème d'optimisation est bien connue et vaut

$$\hat{\mu}(\sigma) = \mathrm{median}(X_1, \ldots, X_n).$$

Une manière non-formelle de s'en convaincre est de se rappeler que dans un sens faible, la dérivée faible de la fonction $x \in \mathbb{R} \mapsto |x|$ est donnée par la fonction signe sgn définie pour tout $x \in \mathbb{R}$ par

$$\mathrm{sgn}(x) := \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ +1 & \text{if } x > 0 \end{cases},$$

ainsi, au sens faible, la condition de premier ordre pour ce problème d'optimisation s'écrit

$$\partial_\mu \sum_{i=1}^{n} |X_i - \mu| = \sum_{i=1}^{n} \operatorname{sgn}(X_i - \mu) = 0.$$

Cela revient à demander qu'il y ait autant de X_i plus petits que de X_i plus grands que μ, et donc à demander que $\mu = \text{median}(X_1, \dots, X_n)$. Cela ne constitue pas une preuve formelle mais permet d'expliquer d'où provient la médiane dans ce problème. On a donc montré que la vraisemblance profilée en σ est donnée par

$$\ell_n^P(\sigma) = -n\log(2) - n\log(\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n} |X_i - \text{median}(X_1, \dots, X_n)|$$

Cette fonction est différentiable en σ et la condition de premier ordre associée donnée par

$$\partial_\sigma \ell_n^P(\sigma) = 0$$

n'admet qu'une solution donnée par

$$\sigma_0 := \frac{1}{n}\sum_{i=1}^{n} |X_i - \text{median}(X_1, \dots, X_n)|.$$

Il est aisé de vérifier que

$$\partial_\sigma^2 \ell_n^P(\sigma_0) < 0$$

et ainsi, par le théorème de la note associée à ce TP, σ_0 est le maximum global de la fonction ℓ_n^P. En combinant nos résultats, on déduit que le MLE de (μ, σ) est donné par

$$(\hat{\mu}_{\text{MLE}}, \hat{\sigma}_{\text{MLE}}) = \left(\text{median}(X_1, \dots, X_n), \frac{1}{n}\sum_{i=1}^{n} |X_i - \text{median}(X_1, \dots, X_n)|\right).$$