PUBLICATIONS

MATHEMATIQUES

D'ORSAY

80.05

L'INÉGALITÉ DE TYPE CRAMER-RAO ET LA VITESSE DE CONVERGENCE POUR CERTAINS ESTIMATEURS SUPER-EFFICACES

LY HOANG TU

Université de Paris-Sud

Département de Mathématique

Bât. 425

91405 ORSAY France

code matière AMS (1980) : 62 F 12

Mots-clefs : Super-efficacité, vitesse de convergence

PUBLICATIONS

MATHEMATIQUES

D'ORSAY

80.05

L'INÉGALITÉ DE TYPE CRAMER-RAO ET LA VITESSE DE CONVERGENCE POUR CERTAINS ESTIMATEURS SUPER-EFFICACES

LY HOANG TU

Université de Paris-Sud

Département de Mathématique

Bât. 425

91405 **ORSAY** France

L'INEGALITE DE TYPE CRAMER-RAO ET LA VITESSE DE CONVERGENCE POUR CERTAINS ESTIMATEURS SUPER-EFFICACES

LY HOANG TU

INSTITUT MATHEMATIQUE DE HA NOI
et UNIVERSITE PARIS SUD - ORSAY

Parmi les inégalité de type Cramer-Rao, nous citons l'inégalité d'Hammersley-Chapman-Robbins, qui n'exige que de très faibles conditions. En 1972 Ibragimov et Haminskii ont présenté une nouvelle inégalité de l'information, dans le cas d'estimateurs sans biais, qui a la forme suivante :

$$\mathrm{Var}_\theta T(x_1,\dots,x_n)+\mathrm{Var}_\phi T(x_1,\dots,x_n) > \frac{1}{2}\,|\tau(\theta) - \tau(\phi)|^2 \times$$

$$\times \{n^{-1}\,|\int_{\mathbb{R}} (\sqrt{f(x,\theta)} - \sqrt{f(x,\phi)})^2\, \mu(dx)|^{-1} - 1\} \qquad (1)$$

om $X_1,\dots,X_n$ sont des variables aléatoires indépendantes, équidistribuées de densité $f(\cdot,\theta)$ par rapport à μ .

Dans cette communication nous présentons une nouvelle inégalité et la comparons avec celle d'Hammersley-Chapman-Robbins et avec (1). Nous examinons aussi la vitesse de convergence de certains estimateurs super-efficaces, un exemple séquentiel est présenté.

I - Soit $(\mathscr{X}, \mathfrak{A}, P_\theta)$, une structure statistique à paramètre $\theta \in \mathbb{R}$; P_θ est dominée par une mesure μ σ-finie et $\frac{dP_\theta}{d\mu} = f(x,\theta)$; soit $T(x)$ un estimateur sans biais d'une fonction numérique $\tau(\theta)$. Alors nous avons l'inégalité suivante :

$$\mathrm{Var}_\theta\, T(x) + c\cdot\mathrm{Var}_\phi(T(x) > |\tau(\theta) - \tau(\phi)|^2 \times$$

$$\times \{|\int \left(\frac{f(x,\theta) - f(x,\phi)}{f(x,\theta) + cf(x,\phi)}\right)^2 (f(x,\theta) + cf(x)\phi))\mu(dx)|^{-1} - c\} \qquad (2)$$

où c est une constante positive.

Si $c = 0$ (2) devant l'inégalité d'Hammersley-Chapman-Robbins. Quand $c = 1$, $X = (x_1,\ldots,x_n)$ et les variables X_i $(i = 1,\ldots,n)$ sont indépendantes et équidistribuées (2) ressemble à l'inégalité d'Ibragimov-Häminskii, mais nous pouvons montrer des exemples où (2) est plus forte que (1).

II - Soit $X_1,\ldots,X_n$, n variables aléatoires réelles, indépendantes, équidistribuées de loi $\frac{dP_\theta}{d\mu} = f(x,\theta)$ où $\theta \in (\alpha,\beta)$ $\mathbb{R}$ $(\alpha < \beta)$ et μ une mesure σ-finie sur $\mathbb{R}$ munie de sa tribu boréliennes

<u>Théorème 1</u> : Soit $T(X_1,\ldots,X_n)$ un estimateur sans biais de θ.

On suppose :

$$\rho(\theta,\phi) = \int_{\mathbb{R}} (\sqrt{f(x,\theta)} - \sqrt{f(x,\phi)})^2 \mu(dx) \leqslant a|\theta-\phi|^{\frac{1}{K}} \qquad (3)$$

où $\theta,\phi \in (\alpha,\beta)$ $K \geqslant \frac{1}{2}$ $a > 0$ $|\phi-\theta| < 1$

Alors il existe une $B > 0$ telle que :

$$\mathrm{Var}_\theta\, T(X_1,\ldots,X_n) + \mathrm{Var}_{\theta \pm \frac{1}{(B_n)^K}}\, T(X_1,\ldots,X_n) \gtrsim \underset{(n\to+\infty)}{n^{(-\frac{1}{2K})}} \qquad (4)$$

<u>Exemple 1</u> : Soit $X_1,\ldots,X_n$ des variables aléatoires indépendantes, équidistribuées de loi :

$$F_\theta(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 - \sqrt[K]{\theta - x} & \text{si} \quad 0 < x \leqslant \theta < 1 \qquad (K \geqslant 2) \\ 1 & \text{si} \quad x > \theta \end{cases}$$

et μ est la somme de la mesure de Lebesgue sur $(\mathbb{R}, \mathscr{B})$ et d'une mesure de Dirac en 0. Alors $\rho(\theta,\phi) < 2|\theta-\phi|^{\frac{1}{K}}$ et on peut appliquer le théorème 1 à l'estimateur de θ.

Si $K = 2$, $T = X_{max}$ notons $X_{max} = \max(X_1,\ldots,X_n)$, on a :

$$E_\theta X_{max} = \theta - \frac{2}{(n+2)(n+1)} + O(u^n) \ , \quad (u = 1 - \sqrt{\theta}); \text{ et } \ E_\theta(X_{max}-\theta)^2 = O(\frac{1}{n^4}) \ .$$

<u>Théorème 2</u> : Soit $T(X_1,\ldots,X_n)$ estimateur non biais de θ ; on suppose :

$$\rho(\theta,\phi) = \int_{\mathbb{R}} (\sqrt{f(x,\theta)} - \sqrt{f(x,\phi)})^2 \mu(dx) \leqslant a \frac{1}{|\ell_n|\phi-\theta||} \ , \quad (a > 0) \qquad (5)$$

Alors il existe une constante $B > 0$ telle que :

$$\mathrm{Var}_\theta\, T(X_1,\ldots,X_n) + \mathrm{Var}_{\theta\pm e^{-|B_n|}}\, T(X_1,\ldots,X_n) \geqslant \underset{(n\to+\infty)}{O(\frac{1}{e^n})} \qquad (6)$$

<u>Exemple 2</u> : Soit $X_1,\ldots,X_n$ les variables aléatoires indépendantes, équidistribuées de loi :

$$F_\theta(x) = \begin{cases} 0 & \text{si} \quad x \leqslant 0 \\ 1 + \frac{1}{\log(\theta-x)} & \text{si} \quad 0 < x < \theta < a < 1 \ , \quad (\ell_n\, a < 1) \\ 1 & \text{si} \quad x \geqslant \theta \end{cases}$$

et μ est la même que dans l'exemple 1. Dans ce cas $\rho(\theta,\phi) < 3 \frac{1}{|\ell_n|\phi-\theta||}$ pour $|\phi-\theta|$ petit.

Soit $T(X_1,\ldots,X_n) = X_{max}$, on peut vérifier directement

$$E_\theta X_{max} = \theta + O(\frac{1}{u^n}) \quad \text{où} \quad u = (1 + \frac{1}{\ell_n \theta})^{-1}$$

et $E_\theta(T(X_1,\ldots,X_n) - \theta)^2 = O(\frac{1}{u^n})$.

<u>Théorème 3</u> : Soit $T(X_1,\ldots,X_n)$-estimateur quelconque (biais ou non) de θ .

Si pour $\theta,\phi \in \mathbb{R}$, on a :

$$\int_{\mathbb{R}} \inf\{f(x,\theta) \ , \ f(x,\phi)\} \mu(dx) \geqslant \frac{1}{a} > 0 \qquad (7)$$

Alors :

$$E_\theta|T(X_1,\dots,X_n)-\theta|^2 + E_\phi|T(X_1,\dots,X_n)-\phi|^2 \geqslant \frac{1}{2}|\theta-\phi|^2 \cdot \frac{1}{a^n} \tag{8}$$

III - Maintenant nous étudions le cas sêquentiel. Soit $X_1,X_2,\dots$ une suite de variables aléatoires indépendantes, équidistribuées de loi P_θ , $\theta \in \mathbb{R}$, nous utilisons la définition du coefficient d'efficacité ε de l'estimateur sêquentiel T_τ de Y. Yu Linnik :

$$\varepsilon = \lim_{n\to\infty} \frac{E_\theta(T_\tau - \theta)^2}{E_\theta(\hat{\theta}_n - \theta)^2}$$

où T_τ est un estimateur sêquentiel sans biais de θ et $E_\theta\tau \leqslant n$, $\hat{\theta}_n$ est un estimateur sans biais de θ de variance minimum à n fixe.

En 1969 Chalit a présenté le cas que $\varepsilon = \frac{1}{3}$ [2] .

En 1972 Ibragimov et Haminskü ont présenté le cas général quand $\varepsilon < 1$ [3].

Ici nous présentons un estimateur sêquentiel pour lequel $\varepsilon = 0$ presque tout θ .

Exemple 3 : Soit $P_\theta(X_i = 0) = P_\theta(X_i = \theta) = \frac{1}{2}$, $0 < \theta < +\infty$, $i = 1,2,\dots$.
Il est facile de montrer que pour tout $\tilde{\theta}_n$, $E_\theta(\tilde{\theta}_n-\theta)^2$ ne peut être nul que pour un seul $\theta = \theta_o$. Posons $\{\mathscr{F}_n\} = \{\sigma(X_1,\dots,X_n)\}$ $n = 1,2,\dots$.
Soit l'estimateur T_τ choisi à l'instant :

$$\tau = \inf \{n : X_{(n)} > 0 \quad n \geqslant 1\}$$

om $X_{(n)} = \max(X_1,\dots,X_n)$ et $T_n = X_{(n)}$.
Alors $E_\theta T_\tau = \theta$ et $E_\theta(T_\tau-\theta)^2 = 0$ pour tout θ et pour $n > c$ $(c = \sum_{i=1}^{\infty} i \frac{1}{2^i})$:

$$\frac{E_\theta(T_\tau-\theta)^2}{E_\theta(\hat{\theta}_n-\theta)^2} = 0 \quad \text{et} \quad \varepsilon = 0 \text{ pour tout } \theta \neq \theta_o .$$

(1)

Exemple 1

$$F_\theta(x) = \begin{cases} 0 & \text{si} \quad x \leqslant 0 \\ 1 - \sqrt[K]{\theta - x} & \text{si} \quad 0 < x \leqslant \theta < 1\,, \qquad K \geqslant 2 \\ 1 & \text{si} \quad x > \theta \end{cases}$$

$$f(x,\theta) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 - \sqrt[K]{\theta} & \text{si} \quad x = 0 \\ \frac{1}{K}(\theta - x)^{\frac{1}{K} - 1} & \text{si} \quad 0 < x < \theta \\ 0 & \text{si} \quad x \geqslant \theta \end{cases}$$

Si $\phi < \theta$, la distance d'Hellinger vérifie :

$$\int \left| \sqrt{f(x,\theta)} - \sqrt{f(x,\phi)} \right|^2 \mu(dx) \leqslant 2\,\sqrt[K]{\theta - \phi} \;.$$

Examinons $X_{max} = \max(X_1, \ldots, X_n)$

$$H_\theta(x) = P_\theta(X_{max} < x) = [F_\theta(x)]^n$$

$$H'_\theta(x) = g(x,\theta) = \begin{cases} 0 & \text{si} \quad x < 0 \\ (1 - \sqrt[2]{\theta})^n & \text{si} \quad x = 0\,, \quad \text{ici} \quad K = 2 \\ n(1 - \sqrt{\theta - x})^{n-1} \times \dfrac{1}{2\sqrt{\theta - x}} & \quad 0 < x < \theta \\ 0 & \quad x \geqslant \theta \end{cases}$$

$$E_\theta\, X_{max} = \int_{\{0\}} 0 \times g(x,\theta)\mu(dx) + n \int_0^\theta x(1 - \sqrt{\theta - x})^{n-1} \times \frac{1}{2\sqrt{\theta - x}}\, dx$$

$$E_\theta\, X_{max} = \theta \cdot [1 - (1 - \sqrt{\theta})^n] + \int_0^\theta (x - \theta) g(x,\theta)\, dx$$

$$\int_0^\theta (x-\theta)\ g(x,\theta)dx = -\int_{1-\sqrt{\theta}}^1 n(1+z^2-2z)z^{n-1}\ dz$$

$$E_\theta\ X_{max} = \theta - \frac{2}{(n+2)(n+1)} + 0\ \left|(1-\sqrt{\theta})^n\right|$$

$$E_\theta(X_{max})^2 = n\int_0^\theta x^2\ \frac{(1-\sqrt{\theta-x})^{n-1}}{2\sqrt{\theta-x}}\ dx$$

$$= -\int_{1-\sqrt{\theta}}^1 n\ \left[\theta^2 + (1-z)^4 - 2\theta(1-z)^2\right]z^{n-1}\ dz$$

$$E_\theta(X_{max})^2 = \theta^2 - \frac{4\theta}{(n+1)(n+2)} + \frac{24}{(n+1)(n+2)(n+3)(n+4)} + 0((1-\sqrt{\theta})^n)$$

$$\Longrightarrow E_\theta(X_{max}-\theta)^2 = E_\theta(X_{max})^2 + \theta^2 - 2\theta\ E_\theta\ X_{max}$$

$$= \frac{24}{(n+1)(n+2)(n+3)(n+4)} + 0((1-\sqrt{\theta})^n)$$

$$E_\theta(X_{max}-\theta)^2 = 0(\frac{1}{n^4})$$

(2)

<u>Exemple 2</u>

$$F_\theta(x) = \begin{cases} \bullet \quad 0 & \text{si} \quad x \leqslant 0 \\ \bullet \quad 1 + \dfrac{1}{\log(\theta - x)} & 0 < x < \theta < a \\ & \quad \log a \leqslant -2 \quad * \\ \bullet \quad 1 & x \geqslant \theta \end{cases}$$

$$f_\theta(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 + \dfrac{1}{\log \theta} & x = 0 \\ \dfrac{1}{(\theta - x)\log^2(\theta - x)} & 0 < x < \theta \\ 0 & x > \theta \end{cases}$$

Nous calculons la distance d'Hellinger, si $\phi < 0$,

$$\rho = \int (\sqrt{f(x,\theta)} - \sqrt{f(x,\phi)})^2 \mu(dx) = \underset{(1)}{\int_{\{0\}}} + \underset{(2)}{\int_0^\phi} + \underset{(3)}{\int_\phi^\theta}$$

(1) $$\int_{\{0\}} = 1 + \frac{1}{\log \theta} + 1 + \frac{1}{\log \phi} - 2\sqrt{(1 + \frac{1}{\log \theta})(1 + \frac{1}{\log \phi})}$$

$$\leqslant 2 + \frac{1}{\log \theta} + \frac{1}{\log \phi} - 2(1 + \frac{1}{\log \theta})$$

$$\leqslant \frac{1}{\log \phi} - \frac{1}{\log \theta} = \left|\frac{1}{\log \theta} - \frac{1}{\log \phi}\right|$$

(remarquons que :

$$\left| \frac{1}{\log \theta} - \frac{1}{\log \phi} \right| < \frac{1}{|\log|\theta - \phi||} \quad \text{pour} \quad |\phi - \theta| \quad \text{assez petit} \quad (|\theta - \phi| \to 0)$$

car

$$\left| \frac{1}{\log \theta} - \frac{1}{\log \phi} \right| = \left| \frac{\log(1 + \frac{\phi-\theta}{\theta})}{\log \phi \cdot \log \theta} \right| \sim \left| \frac{\frac{\phi-\theta}{\theta}}{\log \phi \cdot \log \theta} \right|$$

et
$$\frac{|\phi-\theta|}{\frac{1}{\log|\theta-\phi|}} \longrightarrow 0 \;)$$

et nous avons :

$$\int_{\{0\}} \leqslant \left|\frac{1}{\log \theta} - \frac{1}{\log \phi}\right| \leqslant \frac{1}{|\log|\theta-\phi||}$$

$$(3) \quad \int_{\phi}^{\theta} = 1 - (1 + \frac{1}{\log(\theta-\phi)}) = - \frac{1}{\log(\theta-\phi)} = \frac{1}{|\log|\theta-\phi||}$$

maintenant nous calculons :

$$(2) \quad \int_0^{\phi} (\sqrt{f(x,\theta)} - \sqrt{f(x,\phi)})^2 \, \mu(dx) = \frac{1}{\log(\theta-x)} \Big|_0^{\phi} + \frac{1}{\log(\phi-x)} \Big|_0^{\phi}$$
$$- 2 \int_0^{\phi} \frac{dx}{\sqrt{(\theta-x)(\phi-x)\log^2(\theta-x) \, \log^2(\phi-x)}}$$

si $0 < x < \phi < \theta \leqslant a = e^{-2}$ on a :

$$(\theta-x)\log^2(\theta-x) \geqslant (\phi-x) \log^2(\phi-x)$$

et
$$\int_0^{\phi} < \frac{1}{\log(\theta-\phi)} - \frac{1}{\log \theta} - \frac{1}{\log \phi} - 2 \int_0^{\phi} \frac{dx}{(\theta-x)\log^2(\theta-x)}$$

$$\leqslant \frac{2}{|\log|\theta-\phi||}$$

Combinons (1), (2), (3) nous avons :

$$\rho(\theta,\phi) \leqslant \frac{3}{|\log|\theta-\phi||} \quad \text{pour} \quad |\theta-\phi| \quad \text{assez petit} .$$

(3)

Maintenant examinons X_{max}

$$H_\theta(x) = P_\theta(X_{max} < x) = [F_\theta(x)]^n$$

et

$$g(x,\theta) = \begin{cases} 0 & \text{si} \quad x < 0 \\ (1 + \frac{1}{\log\theta})^n & x = 0 \\ n(1 + \frac{1}{\log(\theta-x)})^{n-1} \times \frac{1}{(\theta-x)\log^2(\theta-x)} & 0 < x < \theta \\ 0 & x \geqslant \theta \end{cases}$$

$$+ \; E_\theta X_{max} = \int_{\{0\}} 0 \times g(x,\theta)\mu(dx) + \int_0^\theta x \, g(x,\theta)dx$$

$$E_\theta X_{max} = \int_0^\theta [\theta + (x-\theta)]^n (1 + \frac{1}{\log(\theta-x)})^{n-1} \frac{1}{(\theta-x)\log^2(\theta-x)} dx$$

$$= \theta[(1 - (1 + \frac{1}{\log\theta})^n] + n \int_{\ell_n\theta}^{-\infty} e^t(1 + \frac{1}{t})^{n-1} \times \frac{1}{t^2} dt$$

$$\left|n \int_{\ell_n\theta}^{-\infty} e^t(1 + \frac{1}{t})^{n-1} \times \frac{1}{t^2} dt\right| \leqslant \left|(1 + \frac{1}{t})^n e^t \Big|_{-\infty}^{\ell_n\theta}\right| + \left|\int_{-\infty}^{\ell_n\theta} (1 + \frac{1}{t})^n e^t \, dt\right|$$

$$\leqslant (1 + \frac{1}{\ell_n\theta})^n + (1 + \frac{1}{\ell_n\theta})^n \times \left|\int_{-\infty}^{\ell_n\theta} e^t \, dt\right|$$

$$\leqslant 2(1 + \frac{1}{\ell_n\theta})^n$$

et

$$E_\theta X_{max} = \theta + O[(1 + \frac{1}{\log\theta})^n]$$

$$E_\theta(X_{max})^2 = \int_0^\theta x^2 \, g(x,\theta)dx = \int_0^\theta (x^2 - \theta^2 + \theta^2)g(x,\theta)dx$$

$$E_\theta(X_{max})^2 = \theta^2\left[1 - (1 + \frac{1}{\log\theta})^n\right] + \int_0^\theta (x-\theta)(x+\theta)g(x,\theta)dx$$

$$\left|\int_0^\theta (x-\theta)(x+\theta)g(x,\theta)dx\right| \leqslant 2\left|\int_0^\theta (\theta-x)g(x,\theta)dx\right|$$

On sait que $$\left|\int_0^\theta (\theta-x)g(x,\theta)dx\right| \leqslant 2(1 + \frac{1}{\log\theta})^n$$

Alors :

$$E_\theta\, X^2_{max} = \theta^2 + 0\left((1 + \frac{1}{\log\theta})^n\right)$$

Et

$$E_\theta(X_{max}-\theta)^2 = \theta^2 + \theta^2 - 2\theta\cdot\theta + 0\left((1 + \frac{1}{\log\theta})^n\right)$$

$$E_\theta(X_{max}-\theta)^2 = 0\left((1 + \frac{1}{\log\theta})^n\right)$$

Ce que nous voulons démontrer.

BIBLIOGRAPHIE

(1) IBRAGIMOV, HAMINSKII (*en russe*).
Une inégalité sur les estimateurs super-efficaces.
Comptes Rendus Ac. Sc. URSS, 1972 T. 204 n° 6.

(2) CHALIT.
Quelques résultats sur la vitesse de convergence d'un estimateur.
Comptes Rendus Ac. Sc. URSS, 1969 T. 189 n° 1.

(3) IBRAGIMOV, HAMINSKII.
Sur l'estimation d'un paramètre.
Comptes Rendus Ac. Sc. URSS, 1972 T. 204 n° 1.

N° D'IMPRESSION 460
4ÈME TRIMESTRE 1980