



HAL
open science

Vitesse de convergence presque sûre des estimateurs du mode de Chernoff et de Venter

Joséphine Leclerc

► **To cite this version:**

Joséphine Leclerc. Vitesse de convergence presque sûre des estimateurs du mode de Chernoff et de Venter. *Annales de l'ISUP*, 2001, XXXXV (1), pp.93-104. hal-03649040

HAL Id: hal-03649040

<https://hal.science/hal-03649040>

Submitted on 22 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
 XXXXV, fasc. 1, 2001, 93 à 104

Vitesse de convergence presque sûre des estimateurs du mode de Chernoff et de Venter

Joséphine LECLERC

L. S. T. A. de l'Université Paris VI, Boîte 158, 4 place Jussieu, 75252 Paris cedex 05

Résumé : Dans ce travail, nous obtenons un majorant de la vitesse de convergence presque sûre de deux estimateurs du mode, les estimateurs de Chernoff et de Venter.

AMS classification : 62G05.

Mots clefs : estimation non paramétrique du mode, processus empiriques.

1 Introduction

Soit X_1, X_2, \dots, X_n une suite de variables aléatoires réelles indépendantes identiquement distribuées, de fonction de répartition F absolument continue (par rapport à la mesure de Lebesgue) admettant une fonction de densité f inconnue qui possède un unique mode θ . Nous nous intéressons au problème particulier de l'estimation de θ , et considérons pour cela les estimateurs proposés par Chernoff (1964) et Venter (1967).

L'estimateur proposé par Chernoff (1964), noté θ_n^* , est défini comme le centre de l'intervalle de longueur $2h_n$, qui contient le plus d'observations, où $\{h_n, n \geq 1\}$ est une suite de réels strictement positifs telle que $\lim_{n \rightarrow \infty} h_n = 0$. Chernoff (1964) (voir également Eddy (1982)) obtient que la distribution asymptotique de son estimateur est identique à celle de la variable qui maximise un mouvement Brownien standard bilatéral, à dérive quadratique. Wegman (1971) a démontré la convergence presque sûre de l'estimateur de Chernoff. Chanda et Ruymgaart (1990) obtiennent une vitesse de convergence presque sûre de θ_n^* , lorsque la densité est deux fois continûment dérivable sur \mathbb{R} . Enfin, Kim (1994) a obtenu la distribution asymptotique de θ_n^* dans le cas multivarié.

L'estimateur proposé par Venter (1967), noté $\hat{\theta}_n$, est défini comme la médiane des observations appartenant à l'intervalle le plus court, contenant au moins $n(1-2a_n)$ des observations, où (a_n) est une suite de réels de $]0, 1/2[$ telle que $\lim_{n \rightarrow +\infty} a_n = 1/2$. Venter (1967) a obtenu la convergence presque sûre de $\hat{\theta}_n$ et établi sa distribution asymptotique pour différents choix de a_n , et sous différentes hypothèses concernant le développement en série de Taylor de f au voisinage de θ . Par une méthode analogue à celle de Chernoff

(1964), Venter (1967) montre que la loi limite est la même que celle d'une variable qui minimise une fonction d'un processus gaussien. Finalement, Venter (1967) obtient des vitesses de convergence presque sûre, pour son estimateur, qui dépendent de l'allure de la densité au voisinage de θ . Plus le mode est prononcé, plus la vitesse de convergence est rapide (si la suite $\{a_n, n \geq 1\}$ est convenablement choisie). Sager (1975) a montré que la convergence presque sûre de $\hat{\theta}_n$, pouvait s'obtenir sous des conditions plus générales sur la distribution. Enfin, Sager (1975) obtient des vitesses de convergence qui, dans certains cas, sont plus rapides que celles trouvées par Venter (1967).

Remarquons que les estimateurs de Chernoff et de Venter sont deux des trois principaux estimateurs directs du mode, avec l'estimateur de Grenander (1965) (voir aussi Andriano (1980) et Hall (1982)). Les estimateurs indirects, qui sont basés sur l'estimation préalable de la densité, sont également considérés dans la littérature. Parmi ces estimateurs, citons l'estimateur à noyau du mode proposé par Parzen (1962), qui repose sur l'estimateur à noyau de la densité introduit par Rosenblatt (1956) (voir Eddy (1980, 1982), Romano (1988), Vieu (1996), Leclerc (2000a) et Leclerc et Pierre-Loti-Viaud (2000)).

Dans ce travail, nous nous intéressons au comportement limite fort des estimateurs de Chernoff et de Venter. En considérant différentes hypothèses de régularité de la densité au voisinage du mode, nous obtenons un majorant de la vitesse de convergence presque sûre (p.s.) de ces deux estimateurs. Précisons que ces résultats complètent et améliorent ceux relevés dans la littérature. Compte tenu du travail effectué et des résultats analogues obtenus pour d'autres estimateurs du mode (voir Leclerc (1997, 2000b)), nous conjecturons que nos résultats donnent la vitesse de convergence p.s. exacte. Ils sont donc une étape indispensable pour établir la loi du logarithme itéré pour ces deux estimateurs.

Dans la Section 2, nous présentons les résultats que nous obtenons pour la vitesse de convergence presque sûre de θ_n^* et $\hat{\theta}_n$. Nous exposons la démonstration de ces résultats dans la Section 3 (voir également Leclerc (2000a)).

2 Résultats

Introduisons tout d'abord les notations suivantes. Soient $G(s) = \inf\{x : F(x) \geq s\}$, $0 < s < 1$, la fonction des quantiles associée à F , $F_n(x) = \frac{1}{n} \#\{X_i \leq x : 1 \leq i \leq n\}$, $x \in \mathbb{R}$, la fonction de répartition empirique, et $G_n(s) = \inf\{x : F_n(x) \geq s\}$, $0 < s \leq 1$, la fonction empirique des quantiles. Enfin, nous notons $\log_2 n = \log^+(\log^+ n)$ avec $\log^+ n = \log(\max(n, e))$.

2.1 L'estimateur de Chernoff

Définissons tout d'abord l'estimateur de Chernoff θ_n^* .

$$\theta_n^* \text{ est toute valeur de } T \in \mathbb{R} \text{ satisfaisant :} \\ F_n(h_n+T) - F_n(-h_n+T) = \sup_{y \in \mathbb{R}} (F_n(h_n+y) - F_n(-h_n+y)). \quad (1)$$

Notons que l'estimateur de Chernoff θ_n^* n'est pas défini de manière unique par (1). Ceci n'est pas un problème car les résultats présentés ci-après sont établis pour toute valeur θ_n^* satisfaisant (1).

Pour cette étude, nous avons besoin des hypothèses suivantes :

Sur la densité

(C.1) f admet un unique mode $\theta \in \mathbb{R}$, qui vérifie

$$\sup_{\{t \in \mathbb{R} : |t - \theta| > \gamma\}} f(t) < f(\theta), \text{ pour tout } \gamma > 0.$$

(C.2) f est continue dans un voisinage de θ .

Sur h_n

(C.3) $\{h_n, n \geq 1\}$ est une suite décroissante de réels strictement positifs vérifiant

$$0 < h_n < 1, \text{ pour } n \geq 1, \text{ et } \lim_{n \rightarrow \infty} h_n = 0.$$

(C.4) $n^{-1/2}(\log_2 n)^{1/2} = o(h_n)$, quand $n \rightarrow +\infty$.

Les résultats que nous obtenons sont présentés dans les deux théorèmes suivants.

Théorème 1 Pour $n \geq 1$, soit θ_n^* défini par (1). Sous les hypothèses (C.1) ... (C.4), nous avons p.s. quand $n \rightarrow +\infty$

$$|\theta_n^* - \theta| = o(1).$$

Théorème 2 Pour $n \geq 1$, soit θ_n^* défini par (1). En plus des hypothèses du Théorème 1, nous supposons qu'il existe un entier $k \geq 2$ tel que

(i) f est C^k au voisinage de θ avec $f^{(1)}(\theta) = f^{(3)}(\theta) = \dots = f^{(j)}(\theta) = 0$

$$\text{où } j = \begin{cases} k-2 & , \text{ si } k \text{ est impair} \\ k-1 & , \text{ si } k \text{ est pair} \end{cases}$$

et que

(ii) $f^{(2)}(\theta) < 0$.

Alors, p.s. quand $n \rightarrow +\infty$, nous avons

$$|\theta_n^* - \theta| = \begin{cases} \mathcal{O}\left(\frac{n^{-1/3}(\log_2 n)^{1/3}}{h_n^{2/3}}\right) & , \text{ si } h_n \leq \hat{H}_{n,k} \\ \mathcal{O}(h_n^{k-1}) & , \text{ si } h_n \geq \hat{H}_{n,k} \end{cases}$$

où $n^{1/(3k-1)}(\log_2 n)^{-1/(3k-1)} \times \hat{H}_{n,k}$ tend vers une constante strictement positive, lorsque $n \rightarrow +\infty$.

La meilleure vitesse de convergence p.s. de $|\theta_n^* - \theta|$, que nous obtenons, est donc

$$|\theta_n^* - \theta| = \mathcal{O}\left(n^{-(k-1)/(3k-1)}(\log_2 n)^{(k-1)/(3k-1)}\right), \text{ p.s. quand } n \rightarrow +\infty,$$

vitesse obtenue pour un h_n de l'ordre de $n^{-1/(3k-1)}(\log_2 n)^{1/(3k-1)}$.

Remarques : 1) Comme θ est le mode de f qui est au minimum C^1 au voisinage de θ , il en découle que $f^{(1)}(\theta) = 0$ et que la première dérivée non nulle en θ est paire. Nous présentons ici le cas $f^{(2)}(\theta) \neq 0$, qui donne la vitesse la plus rapide.

2) Cas particulier : $k = 3$. Nous supposons que la densité f est C^3 au voisinage de θ et qu'elle vérifie $f^{(1)}(\theta) = 0$ et $f^{(2)}(\theta) < 0$. Le théorème 2 établit alors que la vitesse p.s. de $|\theta_n^* - \theta|$ est en $\mathcal{O}\left(n^{-1/4}(\log_2 n)^{1/4}\right)$, vitesse obtenue pour un h_n de l'ordre de $n^{-1/8}(\log_2 n)^{1/8}$. Ce résultat complète donc celui de Chernoff (1964), en donnant une majoration presque sûre. Notons également que nous nous affranchissons de l'hypothèse de Chernoff $f^{(3)}(\theta) \neq 0$.

3) Cas particulier : $k = 2$. Si la densité f est C^2 au voisinage de θ , le théorème 2 établit alors que la vitesse p.s. de $|\theta_n^* - \theta|$ est en $\mathcal{O}(n^{-1/5}(\log_2 n)^{1/5})$, vitesse obtenue pour h_n de l'ordre de $n^{-1/5}(\log_2 n)^{1/5}$. Ce résultat améliore donc celui de Chanda et Ruymgaart (1990), pour lequel la vitesse p.s. de $|\theta_n^* - \theta|$ est en $o(n^{-(\frac{1}{5}-\varepsilon)})$ pour $\varepsilon > 0$.

2.2 L'estimateur de Venter

Nous désignons par $[u]$, respectivement $\lceil u \rceil$, pour $u \in \mathbb{R}$, la partie entière inférieure, respectivement supérieure, de u , i.e. les entiers définis par $u-1 < [u] \leq u$ et $u \leq \lceil u \rceil < u+1$.

Définissons tout d'abord l'estimateur de Venter $\hat{\theta}_n$. Pour cela, posons $N_{a_n} = \lfloor n(1-2a_n) \rfloor$.

$$\hat{\theta}_n = \begin{cases} X_{k_n + \frac{N_{a_n}}{2}, n} & , \text{ si } N_{a_n} \text{ est un entier pair} \\ \frac{1}{2} \left[X_{k_n + \frac{N_{a_n}+1}{2}, n} + X_{k_n + \frac{N_{a_n}+1}{2}, n} \right] & , \text{ si } N_{a_n} \text{ est un entier impair} \end{cases} \quad (2)$$

où $k_n \in \{1, \dots, n - N_{a_n}\}$ est toute valeur de $k \in \{1, \dots, n - N_{a_n}\}$ telle que

$$X_{k+N_{a_n}, n} - X_{k,n} = \min_{k \in \{1, \dots, n - N_{a_n}\}} (X_{k+N_{a_n}, n} - X_{k,n}).$$

Notons que k_n et par conséquent $\hat{\theta}_n$ sont, p. s. quand $n \rightarrow \infty$, définis de manière unique.

Pour cet estimateur, nous avons besoin des hypothèses suivantes :

Sur la densité

(V.1) la densité f admet un support de la forme $S =]a, b[$, avec $-\infty \leq a < b \leq \infty$, i.e. $f(x) > 0$ pour tout $x \in S$, et $f(x) = 0$ sinon.

(V.2) f admet un unique mode $\theta \in S$, qui vérifie

$$\sup_{\{t \in S: |t - \theta| > \gamma\}} f(t) < f(\theta), \text{ pour tout } \gamma > 0.$$

(V.3) f est continue sur son support S .

Sur a_n

En posant $b_n = \frac{1}{2} - a_n$

(V.4) $\{b_n, n \geq 1\}$ est une suite décroissante de réels strictement positifs vérifiant

$$\lim_{n \rightarrow \infty} b_n = 0.$$

(V.5) $n^{-1/2}(\log_2 n)^{1/2} = o(b_n)$, quand $n \rightarrow +\infty$.

Les résultats que nous obtenons sont présentés dans les deux théorèmes suivants.

Théorème 3 *Sous les hypothèses (V.1) ... (V.5), nous obtenons que, p.s. quand $n \rightarrow +\infty$*

$$|\hat{\theta}_n - \theta| = o(1).$$

Théorème 4 *En plus des hypothèses du Théorème 3, nous supposons qu'il existe un entier $k \geq 2$ tel que*

(i) f est C^k au voisinage de θ avec $f^{(1)}(\theta) = f^{(3)}(\theta) = \dots = f^{(j)}(\theta) = 0$

$$\text{où } j = \begin{cases} k-2 & , \text{ si } k \text{ est impair} \\ k-1 & , \text{ si } k \text{ est pair} \end{cases}$$

et que

(ii) $f^{(2)}(\theta) < 0$.

Alors, p.s. quand $n \rightarrow +\infty$, nous avons

$$|\hat{\theta}_n - \theta| = \begin{cases} \mathcal{O}\left(\frac{n^{-1/3}(\log_2 n)^{1/3}}{b_n^{2/3}}\right) & , \text{ si } b_n \leq \hat{H}_{n,k} \\ \mathcal{O}(b_n^{k-1}) & , \text{ si } b_n \geq \hat{H}_{n,k} \end{cases}$$

où $n^{1/(3k-1)}(\log_2 n)^{-1/(3k-1)} \times \hat{H}_{n,k}$ tend vers une constante strictement positive, lorsque $n \rightarrow +\infty$.

La meilleure vitesse de convergence p.s. de $|\hat{\theta}_n - \theta|$, que nous obtenons est

$$|\hat{\theta}_n - \theta| = \mathcal{O}\left(n^{-(k-1)/(3k-1)}(\log_2 n)^{(k-1)/(3k-1)}\right), \text{ p.s. quand } n \rightarrow +\infty,$$

vitesse obtenue pour un b_n de l'ordre de $n^{-1/(3k-1)}(\log_2 n)^{1/(3k-1)}$.

Remarques : 1) Venter (1967) considère une densité f de mode θ admettant le développement : $f(x) = \gamma_0 - \frac{1}{2}\gamma(x-\theta)^2 + o(|x-\theta|^2)$ quand x est proche de θ , avec $\gamma, \gamma_0 > 0$. Il obtient que $|\hat{\theta}_n - \theta| = o(n^{-1/5}(\log n)^{1/2})$, p.s. quand $n \rightarrow +\infty$. Le résultat du Théorème 4 améliore donc celui de Venter, puisqu'il établit que la vitesse p.s. de $|\hat{\theta}_n - \theta|$ est en $\mathcal{O}(n^{-1/5}(\log_2 n)^{1/5})$ (vitesse obtenue pour un b_n de l'ordre de $n^{-1/5}(\log_2 n)^{1/5}$).

2) Venter (1967) considère également le cas d'une densité f de mode θ admettant le développement : $f(x) = \gamma_0 - \frac{1}{2}\gamma(x-\theta)^2 + \frac{1}{6}\gamma_3(x-\theta)^3 + o(|x-\theta|^3)$ quand x est proche de θ , avec $\gamma, \gamma_0 > 0$. Il obtient que $|\hat{\theta}_n - \theta| = o(n^{-1/5}(\log n)^{1/2})$, p.s. quand $n \rightarrow +\infty$. Le résultat du Théorème 4 améliore donc celui de Venter, puisqu'il établit que la vitesse p.s. de $|\hat{\theta}_n - \theta|$ est en $\mathcal{O}(n^{-1/4}(\log_2 n)^{1/4})$ (vitesse obtenue pour un b_n de l'ordre de $n^{-1/8}(\log_2 n)^{1/8}$).

3 Démonstrations

Introduisons quelques notations : soit, sans perte de généralité, $X_i = G(U_i)$, $i = 1, 2, \dots$ où les U_i sont des variables aléatoires indépendantes identiquement distribuées de loi uniforme sur $[0,1]$. Posons $U_n(x) = \frac{1}{n} \#\{U_i \leq x : 1 \leq i \leq n\}$, $x \in \mathbb{R}$, la fonction de répartition empirique uniforme et $V_n(s) = \inf\{x \geq 0 : U_n(x) \geq s\}$, $0 \leq s \leq 1$, la fonction empirique uniforme des quantiles.

L'estimateur de Chernoff, respectivement Venter, a la particularité d'être en relation avec la solution d'un problème de minimisation d'un processus, fonction des accroissements du processus empirique uniforme, respectivement du processus empirique des quantiles. Plus précisément, nous nous ramenons à l'étude d'une variable aléatoire T_n , respectivement \hat{T}_n , qui p.s. quand $n \rightarrow +\infty$, minimise un processus de la forme

$$E_n(T) = D_n(T) + A_n(T), \text{ pour } T \text{ dans un voisinage de } 0, \quad (3)$$

où D_n est une fonction déterministe qui admet un développement asymptotique de type polynomial d'ordre deux en T , et A_n est une fonction des accroissements du processus empirique uniforme, respectivement du processus empirique des quantiles. Cette dernière partie aléatoire se traite en faisant appel à des résultats de type loi du logarithme itéré.

3.1 Résultat préliminaire

Nous établissons tout d'abord un résultat simple d'analyse, qui est nécessaire au traitement du problème de minimisation de (3).

Proposition 1 Soient g, f et E trois fonctions définies sur \mathbb{R}^+ , à valeurs réelles, et vérifiant

$$(i) \quad g(T) \leq E(T) \leq f(T), \quad \forall T \in \mathbb{R}^+.$$

(ii) $g(T), E(T)$ et $f(T)$ sont continues sur \mathbb{R}^+ .

$$(iii) \quad \lim_{T \rightarrow +\infty} g(T) = +\infty.$$

Alors, un majorant des valeurs de T qui minimisent $E(T)$ est défini par

$$T^* = \max\{T \geq 0 / g(T) = \inf_{y \geq 0} [f(y)]\}.$$

La preuve de la Proposition 1 se trouve dans Leclerc (2000a).

Remarque : L'analogie de la Proposition 1 sur \mathbb{R}^- s'énonce de la même manière.

3.2 Preuve du Théorème 1

La démonstration s'inspire de celle du résultat de Romano (1988) qui établit la convergence presque sûre de l'estimateur à noyau du mode.

Posons $P_n(t) = F_n(t+h_n) - F_n(t-h_n)$ et $p_n(t) = F(t+h_n) - F(t-h_n)$, sur $t \in \mathbb{R}$. Comme θ_n^* maximise $P_n(t)$ sur $t \in \mathbb{R}$, il est clair que

$$\begin{aligned} \liminf_{n \rightarrow +\infty} \left(\frac{P_n(\theta_n^*)}{2h_n} \right) &\geq \liminf_{n \rightarrow +\infty} \left(\frac{P_n(\theta)}{2h_n} \right) \\ &= \liminf_{n \rightarrow +\infty} \left(\frac{P_n(\theta)}{2h_n} - f(\theta) \right) + f(\theta) \end{aligned}$$

En supposant

$$\begin{cases} (i) \quad \liminf_{n \rightarrow +\infty} \left(\frac{P_n(\theta)}{2h_n} - f(\theta) \right) = 0 \text{ p.s.}, \\ (ii) \quad \text{pour tout } \sigma > 0, \limsup_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| \geq \sigma\}} \left(\frac{P_n(t)}{2h_n} \right) < f(\theta) \text{ p.s.}, \end{cases}$$

nous obtenons alors, pour tout $\sigma > 0$, que

$$\liminf_{n \rightarrow +\infty} \left(\frac{P_n(\theta_n^*)}{2h_n} \right) > \limsup_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| \geq \sigma\}} \left(\frac{P_n(t)}{2h_n} \right) \text{ p.s.},$$

ce qui implique que p.s. quand $n \rightarrow +\infty$: $|\theta_n^* - \theta| < \sigma$.

Il nous reste donc à montrer que les deux assertions (i) et (ii) sont bien vérifiées. Observons pour commencer que (C.4) combinée avec la loi du logarithme itéré de Chung

(1949) et Smirnov (1944) (voir par exemple le théorème 5.1.1 de Csörgő et Révész (1981) page 157) implique que, p.s. quand $n \rightarrow +\infty$

$$\begin{aligned} \sup_{t \in \mathbb{R}} \left| \frac{P_n(t) - p_n(t)}{2h_n} \right| &\leq \frac{1}{h_n} \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \\ &= \mathcal{O} \left(\frac{n^{-1/2} (\log_2 n)^{1/2}}{h_n} \right) \\ &= o(1). \end{aligned} \quad (4)$$

Vérifions maintenant (i). Deux développements de Taylor-Lagrange à l'ordre 0 de F en θ combinés avec (C.2) et (C.3) donnent que

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \left| \frac{p_n(\theta)}{2h_n} - f(\theta) \right| &\leq \lim_{n \rightarrow +\infty} \frac{1}{2} \{ |f(\theta + \mu_1 h_n) - f(\theta)| + |f(\theta - \mu_2 h_n) - f(\theta)| \} \\ &= 0, \end{aligned} \quad (5)$$

où $0 < \mu_i = \mu_i(\theta, h_n) < 1$, $i = 1, 2$.

En combinant maintenant (4) avec (5), il est clair que l'assertion (i) est donc bien vérifiée.

Vérifions maintenant (ii). Deux développements de Taylor-Lagrange à l'ordre 0 de F en t combinés avec (C.3) et (C.1) donnent que, pour tout $\gamma > 0$

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| > 2\gamma\}} \left(\frac{p_n(t)}{2h_n} \right) &\leq \sup_{\{t: |t-\theta| > \gamma\}} f(t) \\ &< f(\theta), \end{aligned} \quad (6)$$

Soit $\sigma > 0$ une constante. En combinant (4) avec (6) (avec $\gamma < \sigma/2$), nous obtenons que

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| \geq \sigma\}} \left(\frac{P_n(t)}{2h_n} \right) &\leq \lim_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| \geq \sigma\}} \left| \frac{P_n(t) - p_n(t)}{2h_n} \right| + \limsup_{n \rightarrow +\infty} \sup_{\{t: |t-\theta| \geq \sigma\}} \left(\frac{p_n(t)}{2h_n} \right) \\ &< f(\theta) \text{ p.s.} \end{aligned}$$

L'assertion (ii) est donc bien vérifiée, ce qui termine la preuve du Théorème 1.

3.3 Preuve du Théorème 2

Nous utilisons la transformation suivante de l'estimateur de Chernoff. Nous posons $\theta_n^* = T_n + \theta$ dans (1) et écrivons que T_n est défini par

$$\begin{aligned} T_n \text{ est toute valeur de } T \in \mathbb{R} \text{ qui minimise la fonction} \\ L_n^*(T) = F_n(\theta - h_n + T) - F_n(\theta + h_n + T). \end{aligned} \quad (7)$$

Notons que T_n n'est pas défini de manière unique. Ceci n'est pas un problème car les résultats présentés ci-après sont établis pour toute valeur T_n satisfaisant (7).

Pour la démonstration du Théorème 2, nous avons besoin d'établir, au préalable, le résultat suivant.

Lemme 1 Soit f une densité de mode θ . Supposons les hypothèses (C.2) et (C.3) vérifiées, et posons

$$\begin{cases} 0 < m = \inf_{-\varepsilon < x < \varepsilon} f(\theta+x) < +\infty, \text{ pour } \varepsilon > 0, \\ a_n = \frac{d \log_2 n}{n} \text{ pour } d \geq 0.236\dots \end{cases}$$

Alors, nous obtenons que p.s. quand $n \rightarrow +\infty$

$$L_n^*(T) = [F(\theta-h_n+T) - F(\theta-h_n)] - [F(\theta+h_n+T) - F(\theta+h_n)] + C_n + \mu_n(T), \quad (8)$$

avec C_n indépendant de T et

$$|\mu_n(T)| = \sqrt{|T|} \times n^{-1/2} (\log_2 n)^{1/2} \times \mathcal{O}(1), \quad (9)$$

où le $\mathcal{O}(1)$ est uniforme en $T \in \mathcal{D} = [-\eta, -\frac{a_n}{m}] \cup [\frac{a_n}{m}, \eta]$, avec $\eta > 0$ une constante suffisamment petite.

Remarques : 1) Le fait que f admet un mode θ et l'hypothèse (C.2), nous assurent de l'existence d'un $\varepsilon > 0$ tel que $m > 0$.

2) Le terme C_n n'intervient pas dans la minimisation de $L_n^*(T)$.

Preuve du Lemme 1 :

En combinant (7) avec le fait que, sans perte de généralité, $U_n(F(x)) = F_n(x)$, remarquons que $L_n^*(T)$ s'écrit comme (8) avec $C_n = U_n(F(\theta-h_n)) - U_n(F(\theta+h_n))$ et

$$\begin{aligned} \mu_n(T) &= \{[U_n(F(\theta-h_n+T)) - F(\theta-h_n+T)] - [U_n(F(\theta-h_n)) - F(\theta-h_n)]\} - \\ &\quad \{[U_n(F(\theta+h_n+T)) - F(\theta+h_n+T)] - [U_n(F(\theta+h_n)) - F(\theta+h_n)]\} \\ &= A - B \end{aligned}$$

Les termes A et B se traitent de la même manière. Pour obtenir le $\sqrt{|T|}$ de (9), nous utilisons le résultat suivant qui découle du théorème de Csaki (théorème 4.5.4 page 148 de Csörgö et Révész (1981)).

Si $a_n = \frac{d \log_2 n}{n}$ où $d \geq 0.236\dots$ et $\tilde{U}_n(s)$ désigne la fonction empirique associée à $\tilde{U}_1, \dots, \tilde{U}_n$ des variables aléatoires indépendantes identiquement distribuées de loi uniforme sur $[0,1]$, alors, pour $P > 2$ fixé, et pour presque tout ω , il existe $R = R(\omega) < +\infty$, tel que pour tout $n \geq R$

$$\sup_{x \in [a_n, 1-a_n]} \frac{|\tilde{U}_n(x) - x|}{\sqrt{x(1-x)}} \leq P \times n^{-1/2} (\log_2 n)^{1/2}. \quad (10)$$

Afin de pouvoir utiliser ce résultat, nous effectuons le changement de variables suivant. Soient $\tilde{U}_i, i = 1 \dots n$ les variables aléatoires indépendantes uniformément distribuées sur $[0, 1]$, définies par

$$\tilde{U}_i = \begin{cases} U_i - F(\theta - h_n) & , \text{ si } F(\theta - h_n) \leq U_i \leq 1 \\ U_{i+1} - F(\theta - h_n) & , \text{ si } 0 \leq U_i \leq F(\theta - h_n) \end{cases}$$

Nous obtenons alors que la fonction de répartition empirique associée à $\tilde{U}_1, \dots, \tilde{U}_n$ est

$$\tilde{U}_n(x) = \begin{cases} U_n(F(\theta - h_n) + x) - U_n(F(\theta - h_n)) & , \text{ si } 0 \leq x \leq 1 - F(\theta - h_n) \\ 1 - U_n(F(\theta - h_n)) + U_n(F(\theta - h_n) + x - 1) & , \text{ si } 1 - F(\theta - h_n) \leq x \leq 1 \end{cases} \quad (11)$$

Considérons maintenant $T \in [\frac{2n}{m}, \eta]$ et posons $x = F(\theta - h_n + T) - F(\theta - h_n)$. Comme (C.3), un développement de Taylor-Lagrange à l'ordre 0 de F en $(\theta - h_n)$ donne, pour η suffisamment petit, et quand $n \rightarrow +\infty$, que

$$\begin{cases} x = T f(\theta - h_n + \mu_1 T) , \text{ où } 0 < \mu_1 = \mu_1(T, h_n) < 1, \\ a_n \leq x \leq 1 - a_n. \end{cases}$$

Nous pouvons donc combiner (11) avec (10), ce qui implique que, p.s. quand $n \rightarrow +\infty$

$$\begin{aligned} |A| &= |\tilde{U}_n(x) - x| \\ &\leq P\sqrt{x} \times n^{-1/2} (\log_2 n)^{1/2} \\ &= \sqrt{T} \times n^{-1/2} (\log_2 n)^{1/2} \times \mathcal{O}(1), \end{aligned}$$

où $\mathcal{O}(1)$ est uniforme en $T \in [\frac{2n}{m}, \eta]$.

Le cas $T \in [-\eta, -\frac{2n}{m}]$ se traite de la même manière en utilisant le terme $\sqrt{1-x}$ de (10). Finalement, nous obtenons donc que, pour tout $T \in \mathcal{D} = [-\eta, -\frac{2n}{m}] \cup [\frac{2n}{m}, \eta]$, p.s. quand $n \rightarrow +\infty$

$$|A| = \sqrt{|T|} \times n^{-1/2} (\log_2 n)^{1/2} \times \mathcal{O}(1),$$

où $\mathcal{O}(1)$ est uniforme en $T \in \mathcal{D}$.

Une démarche analogue permettant de traiter le terme B , la preuve du Lemme 1 est donc terminée.

Preuve du Théorème 2 :

Dans cette démonstration, nous allons effectuer des développements de Taylor de F , ce qui nous amène à préciser l'hypothèse (i) du Théorème 2. Nous ne considérons ici que le cas où f est C^k au voisinage de θ , avec $k \geq 3$ un entier impair tel que

$$\begin{cases} f^{(1)}(\theta) = f^{(3)}(\theta) = \dots = f^{(k-2)}(\theta) = 0 \\ f^{(k)}(\theta) \neq 0 \end{cases}$$

Les trois autres cas de figure se traitent de la même manière.

Nous allons, dans un premier temps, combiner le Lemme 1 avec des développements de Taylor de F .

Soit $\eta > 0$ une constante suffisamment petite, $0 < m = \inf_{-\varepsilon < x < \varepsilon} f(\theta+x) < +\infty$, pour $\varepsilon > 0$ et $a_n = \frac{d \log_2 n}{n}$ pour $d \geq 0.236\dots$ D'après le théorème 1, p.s. quand $n \rightarrow +\infty$, T_n minimise $L_n^*(T)$ pour $T \in [-\eta, \eta]$. Remarquons que si $T \in [\frac{-a_n}{m}, \frac{a_n}{m}]$, alors le résultat du Théorème 2 est immédiat, ce qui nous permet de considérer $L_n^*(T)$ pour $T \in \mathcal{D} = [-\eta, -\frac{a_n}{m}] \cup [\frac{a_n}{m}, \eta]$.

Posons $\Phi(v) = F(\theta-h_n+vT) - F(\theta+h_n+vT)$. Le Lemme 1 combiné avec la formule de MacLaurin, à l'ordre $(k-1)$, appliquée à Φ au point 1, permet d'écrire que, p. s. quand $n \rightarrow +\infty$, T_n minimise

$$\begin{aligned} & [F(\theta-h_n+T) - F(\theta-h_n)] - [F(\theta+h_n+T) - F(\theta+h_n)] + \mu_n(T) \\ &= [\Phi(1) - \Phi(0)] + \mu_n(T) \\ &= \left[T\varphi_1(1) + \frac{T^2}{2}\varphi_2(1) + \frac{T^3}{3!}\varphi_3(1) + \dots + \frac{T^{(k-1)}}{(k-1)!}\varphi_{k-1}(1) + \frac{T^k}{k!}\varphi_k(1) \right] + \mu_n(T), \end{aligned} \quad (12)$$

où

$$\begin{cases} \varphi_s(v) = f^{(s-1)}(\theta-vh_n) - f^{(s-1)}(\theta+vh_n), & s = 1, 2, \dots, k-1 \\ \varphi_k(v) = f^{(k-1)}(\theta-vh_n+\alpha T) - f^{(k-1)}(\theta+vh_n+\alpha T), & \text{pour un } 0 < \alpha = \alpha(v) < 1 \\ |\mu_n(T)| = \sqrt{|T|} n^{-1/2} (\log_2 n)^{1/2} \times \mathcal{O}(1), & \text{où le } \mathcal{O}(1) \text{ est uniforme en } T \in \mathcal{D} \end{cases}$$

Nous allons maintenant appliquer la formule de MacLaurin, à l'ordre $(k-s)$, aux $\varphi_s(v)$ et au point 1, pour $s = 1, 2, \dots, k$.

Comme (C.3) et f est C^k au voisinage de θ , avec k impair, la formule de MacLaurin, à l'ordre $(k-1)$, appliquée à $\varphi_1(v)$ au point 1, donne, quand $n \rightarrow +\infty$

$$\varphi_1(1) = -2h_n \frac{f^{(1)}(\theta)}{1!} - 2h_n^3 \frac{f^{(3)}(\theta)}{3!} - \dots - 2h_n^{k-2} \frac{f^{(k-2)}(\theta)}{(k-2)!} - \frac{h_n^k}{k!} [f^{(k)}(\theta - \beta_1 h_n) + f^{(k)}(\theta + \beta_1 h_n)],$$

où $0 < \beta_1 < 1$.

En combinant ceci avec l'hypothèse (i) et (C.3), nous obtenons quand $n \rightarrow +\infty$, que

$$T\varphi_1(1) = -2Th_n^k \frac{f^{(k)}(\theta)}{k!} \times (1+o(1)), \quad (13)$$

où le $o(1)$ est uniforme en $T \in \mathcal{D}$.

Nous procédons maintenant de la même manière pour $\varphi_2(1)$. Comme (C.3) et f est C^k au voisinage de θ avec k impair, la formule de MacLaurin, à l'ordre $(k-2)$, appliquée à $\varphi_2(v)$ au point 1, donne, quand $n \rightarrow +\infty$

$$\varphi_2(1) = -2h_n \frac{f^{(2)}(\theta)}{1!} - 2h_n^3 \frac{f^{(4)}(\theta)}{3!} - \dots - 2h_n^{k-2} \frac{f^{(k-1)}(\theta)}{(k-2)!} + \frac{h_n^{k-1}}{(k-1)!} [f^{(k)}(\theta - \beta_2 h_n) - f^{(k)}(\theta + \beta_2 h_n)],$$

En combinant ceci avec les hypothèses (i), (ii) et (C.3), nous obtenons quand $n \rightarrow +\infty$, que

$$\frac{T^2}{2} \varphi_2(1) = -T^2 h_n f^{(2)}(\theta) \times (1+o(1)), \tag{14}$$

où le $o(1)$ est uniforme en $T \in \mathcal{D}$.

Intéressons-nous maintenant à $\varphi_s(1)$, pour $3 \leq s \leq k-1$. Comme (C.3) et f est C^k au voisinage de θ , la formule de MacLaurin, à l'ordre $(k-s)$, appliquée à $\varphi_s(v)$ au point 1, donne, quand $n \rightarrow +\infty$

$$\begin{aligned} \varphi_s(1) = & -2h_n \frac{f^{(s)}(\theta)}{1!} - 2h_n^3 \frac{f^{(s+2)}(\theta)}{3!} - \dots + \frac{(-1)^{k-s}}{(k-s)!} h_n^{k-s} [f^{(k-1)}(\theta) - (-1)^{k-s} f^{(k-1)}(\theta)] \\ & + \frac{(-1)^{k-s+1}}{(k-s+1)!} h_n^{k-s+1} [f^{(k)}(\theta - \beta_s h_n) - (-1)^{k-s+1} f^{(k)}(\theta + \beta_s h_n)], \end{aligned}$$

où $0 < \beta_s < 1$.

En combinant ceci avec l'hypothèse (i) et (C.3), nous obtenons quand $n \rightarrow +\infty$, que

$$\frac{T^s}{s!} \varphi_s(1) = \begin{cases} T^2 h_n \times o(1) & , \text{ si } s \text{ est impair} \\ T^3 h_n \times \mathcal{O}(1) & , \text{ si } s \text{ est pair} \end{cases} \tag{15}$$

où les $o(1)$ et $\mathcal{O}(1)$ sont uniformes en $T \in \mathcal{D}$.

Intéressons-nous enfin à $\varphi_k(1)$. Comme (C.3) et f est C^k au voisinage de θ avec k impair, la formule de MacLaurin, à l'ordre 0, appliquée à $\varphi_k(v)$ au point 1, donne, quand $n \rightarrow +\infty$

$$\varphi_k(1) = -h_n [f^{(k)}(\theta - \beta_k h_n + \alpha T) + f^{(k)}(\theta + \beta_k h_n + \alpha T)],$$

où $0 < \beta_k < 1$.

En combinant ceci avec l'hypothèse (i) et (C.3), nous obtenons quand $n \rightarrow +\infty$, que

$$\frac{T^k}{k!} \varphi_k(1) = T^3 h_n \times \mathcal{O}(1), \tag{16}$$

où le $\mathcal{O}(1)$ est uniforme en $T \in \mathcal{D}$.

Posons maintenant $r = -f^{(2)}(\theta) + T\mathcal{O}(1)$, $l = -2 \frac{f^{(k)}(\theta)}{k!} \neq 0$ et $v_n = \frac{n^{-1/2} (\log_2 n)^{1/2}}{h_n}$. En combinant (12) avec (13), (14), (15) et (16), nous obtenons finalement que p.s. quand $n \rightarrow +\infty$, T_n minimise

$$E_n(T) = rT^2 \times (1+o(1)) + lTh_n^{k-1} \times (1+o(1)) + \sqrt{|T|} v_n \times \mathcal{O}(1),$$

où $r > 0$ dès que la constante η est suffisamment petite et les $o(1)$ et $\mathcal{O}(1)$ sont uniformes en $T \in \mathcal{D}$.