Math-Net.Ru

Общероссийский математический портал

A. Mokkadem, M. Pelletier, J. Worms, A large deviations upper bound for the kernel mode estimator, *Теория вероятн. и ее примен.*, 2005, том 50, выпуск 1, 189–200

DOI: <https://doi.org/10.4213/tvp168>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.91.87.62

26 апреля 2023 г., 13:31:14

26. Samuel-Cahn E. Asymptotic distributions for occupancy and waiting time problems with positive probability of falling through the cells. — Ann. Probab., 1974, v. 2, p. 515–521.
27. Сенета Е. Правильно меняющиеся функции. М.: Наука, 1985, 141 с.

Поступила в редакцию
16.IX.2001

© 2005 г. MOKKADEM A.* , PELLETIER M.* , WORMS J.*

A LARGE DEVIATIONS UPPER BOUND FOR THE KERNEL MODE ESTIMATOR¹⁾

Доказывается верхняя граница больших уклонений для ядерной оценки моды, соответствующей строгому максимуму плотности вероятностей. Скорость сходимости неожиданно оказывается более быстрой, чем можно было ожидать, судя по скорости слабой сходимости изучаемой оценки.

Ключевые слова и фразы: плотность, мода, ядерная оценка, принцип больших уклонений.

1. Introduction. Let $X_1, \dots, X_n$ be a sequence of independent and identically distributed real-valued random variables with unknown continuous probability density f . A mode of $f(x)$ is a value of x which maximizes f . More precisely, we define the functional

$$L: g \mapsto \inf \left\{ y, g(y) = \sup_{x \in \mathbf{R}} g(x) \right\} \quad \text{for } g: \mathbf{R} \rightarrow \mathbf{R}$$

and assume $|L(f)| < +\infty$; since f is continuous, the real number θ which satisfies $\theta = L(f)$ is a mode of f .

The kernel estimator of θ , introduced by Parzen [9], is defined as

$$\theta_n = L(f_n),$$

where f_n is the well-known Rosenblatt–Parzen kernel estimator of the density f ; recall that, for all $x \in \mathbf{R}$,

$$f_n(x) = \frac{1}{nh_n} \sum_{k=1}^n K\left(\frac{X_k - x}{h_n}\right),$$

where $(h_n)_{n \geq 1}$ is a sequence of positive real numbers such that $\lim_{n \rightarrow +\infty} h_n = 0$, $\lim_{n \rightarrow +\infty} nh_n = +\infty$, and the kernel K is a continuous function satisfying $\lim_{|x| \rightarrow \infty} K(x) = 0$, $\int K(x) dx = 1$.

The properties of K ensure the existence of θ_n . Moreover, K being continuous, $\sup_{x \in I} f_n(x)$ is a random variable for any interval I ; the fact that, for any $\beta \in \mathbf{R}$,

$$\{\theta_n \leq \beta\} = \left\{ \sup_{x \in]-\infty, \beta]} f_n(x) = \sup_{x \in \mathbf{R}} f_n(x) \right\}$$

ensures θ_n to be a random variable.

The convergence in probability of (θ_n) to θ was established by Parzen [9] and Yamato [13], its almost sure convergence was proved by Nadaraja [7], Van Ryzin [12], and

* Université de Versailles-Saint-Quentin, Département de Mathématiques, 45, Avenue des Etats-Unis, 78035 Versailles Cedex, France; e-mail: mokkadem@math.uvsq.fr, pelletier@math.uvsq.fr, worms@math.uvsq.fr

Romano [10]. The L^p error in mode estimation is studied by Eddy [1] and Grund and Hall [3]; Novak [8] gives the optimal kernels with respect to mean square error.

The weak convergence rate of (θ_n) to θ was studied by Parzen [9], Konakov [4], Eddy [1], [2], Romano [10], and Mokkadem and Pelletier [6]: under some regularity assumptions on f and K , and for a suitable choice of (h_n) , it is known that $(\theta_n - \theta)$ fulfills the central limit theorem

$$\sqrt{nh_n^3}(\theta_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mu, \frac{f(\theta)}{[f''(\theta)]^2} \int K'^2(x) dx\right), \quad (1)$$

where $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution, $\mathcal{N}$ the Gaussian distribution, and $\mu \in \mathbf{R}$ is eventually nonzero.

Our aim in this paper is to establish a large deviations upper bound for $(\theta_n - \theta)$. More precisely, we prove that for any $\varepsilon > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\}] \leq -c(\varepsilon), \quad (2)$$

where c is a strictly positive function explicitly defined in Section 2.

In view of the central limit theorem (1), one would have expected $(\theta_n - \theta)$ to fulfill a large deviations principle with rate nh_n^3 . Thus, our result is quite amazing since the rate nh_n we obtain in the upper bound (2) is faster than the rate nh_n^3 suggested by the weak convergence rate of $(\theta_n - \theta)$.

2. Assumptions and main result. Our assumptions are the following.

(A1) *Conditions on the probability density f :*

- i) there exists $\theta \in \mathbf{R}$ such that $f(x) < f(\theta)$ for all $x \neq \theta$;
- ii) f is uniformly continuous on $\mathbf{R}$;
- iii) there exists $\alpha > 0$ such that $\int_{\mathbf{R}} |x|^\alpha f(x) dx < +\infty$.

(A2) *Conditions on the kernel K :*

- i) $\int_{\mathbf{R}} K(x) dx = 1$;
- ii) K is Hölder continuous with exponent β ;
- iii) there exists $\gamma > 0$ such that $x \mapsto |x|^\gamma K(x)$ is bounded.

(A3) *Conditions on the sequence (h_n) :*

- i) $\lim_{n \rightarrow \infty} h_n = 0$;
- ii) $\lim_{n \rightarrow \infty} nh_n [\ln(1/h_n)]^{-1} = +\infty$.

R e m a r k s. 1. (A1) i) means that f has a unique mode θ , which corresponds to a strict maximum of f (and $\|f\|_\infty = f(\theta)$).

2. Note that K is not assumed to be nonnegative.

3. (A3) ii) is the sharp condition that ensures the almost sure convergence of θ_n to θ (see Romano [10]). Note that the upper bound (2) implies that, for n large enough,

$$\mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\} \leq \exp\left(-\frac{nh_n c(\varepsilon)}{2}\right).$$

Now, in view of assumption (A3) ii), we have, for n sufficiently large, $nh_n c(\varepsilon)/2 \geq 3 \ln(1/h_n)$, $\exp(-nh_n c(\varepsilon)/2) \leq h_n^3$, and thus

$$\mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\} \leq \begin{cases} n^{-3/2} & \text{if } h_n \leq n^{-1/2}, \\ \exp\left(-\frac{c(\varepsilon)\sqrt{n}}{2}\right) & \text{if } h_n \geq n^{-1/2}. \end{cases}$$

So, (2) implies the strong consistency of θ_n .

Let us also mention that (A3) is a necessary and sufficient condition for the uniform convergence of f_n to f (see [11]).

Let us now introduce some notation.

We first define

$$\begin{aligned} \mathcal{H}_r &= \left\{ H \mid \exists x_0, s \text{ such that } \theta < x_0 < s \text{ and } f(x_0) < f(s) = H \right\}, \\ \mathcal{H}_l &= \left\{ H \mid \exists x_0, s \text{ such that } s < x_0 < \theta \text{ and } f(x_0) < f(s) = H \right\}. \end{aligned}$$

Note that if $\mathcal{H}_r \neq \emptyset$ (respectively, $\mathcal{H}_l \neq \emptyset$), then f has at least one local maximum on $] \theta, +\infty[$ (respectively, on $] -\infty, \theta[$); however, the converse is false (see Fig. 1). Note also that $\mathcal{H}_r = \emptyset$ (respectively, $\mathcal{H}_l = \emptyset$) if and only if f is nonincreasing on $[\theta, +\infty[$ (respectively, nondecreasing on $] -\infty, \theta]$).

Fig. 1. $H_1 \notin \mathcal{H}_r$, $H_2 \in \mathcal{H}_r$, $H_l = \emptyset$.

Then, we set

$$\begin{cases} M_r = 0 \text{ if } \mathcal{H}_r = \emptyset \text{ and } M_r = \sup \{H \in \mathcal{H}_r\} \text{ otherwise,} \\ M_l = 0 \text{ if } \mathcal{H}_l = \emptyset \text{ and } M_l = \sup \{H \in \mathcal{H}_l\} \text{ otherwise} \end{cases}$$

and

$$\begin{cases} a_r = \sup \{u \geq 0 \mid f(x) \geq M_r \forall x \in [\theta, \theta + u]\}, \\ a_l = \sup \{u \geq 0 \mid f(x) \geq M_l \forall x \in [\theta - u, \theta]\} \end{cases}$$

(see Fig. 2).

Fig. 2.

Note that $a_r = +\infty$ (respectively, $a_l = +\infty$) when $M_r = 0$ (respectively, $M_l = 0$). Moreover, the uniform continuity and the integrability of f imply that $\lim_{|x| \rightarrow \infty} f(x) = 0$; it follows that $M_r < f(\theta)$ and $a_r > 0$ (respectively, $M_l < f(\theta)$ and $a_l > 0$). Let us also

mention that, in view of the definitions of M_r and a_r (respectively, of M_l and a_l), f is nonincreasing on $[\theta, \theta + a_r]$ (respectively, nondecreasing on $[\theta - a_l, \theta]$).

Finally, for u , t , and δ in $\mathbf{R}$, we define the functions ψ , I , g , and G as

$$\psi : u \mapsto \int_{\mathbf{R}} (e^{uK(z)} - 1) dz, \quad (3)$$

$$I : t \mapsto \sup_{u \in \mathbf{R}} \{ut - \psi(u)\}, \quad (4)$$

$$g : \delta \mapsto f(\theta) I \left(1 + \frac{\delta}{f(\theta)} \right), \quad (5)$$

$$G : \delta \mapsto \min\{g(\delta); g(-\delta)\}. \quad (6)$$

We shall give in Section 3 the main properties of these functions. (Note that G is even.)

We now state our main result.

Theorem 1. Assume (A1)–(A3) hold. Then, for any $\varepsilon > 0$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}] \leq -c_r(\varepsilon)$$

with

$$c_r(\varepsilon) = \begin{cases} \frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - f(\theta + \varepsilon)}{2}\right) & \text{if } \varepsilon \in]0; a_r[, \\ \frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - M_r}{2}\right) & \text{if } \varepsilon \geq a_r \end{cases}$$

and

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{\theta - \theta_n \geq \varepsilon\}] \leq -c_l(\varepsilon)$$

with

$$c_l(\varepsilon) = \begin{cases} \frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - f(\theta - \varepsilon)}{2}\right) & \text{if } \varepsilon \in]0; a_l[, \\ \frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - M_l}{2}\right) & \text{if } \varepsilon \geq a_l. \end{cases}$$

The result we have claimed in the introduction is the following corollary.

Corollary 1. Assume (A1)–(A3) hold. Then, for any $\varepsilon > 0$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\}] \leq -c(\varepsilon) \quad \text{with} \quad c(\varepsilon) = \frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - m(\varepsilon)}{2}\right),$$

where

$$m(\varepsilon) = \begin{cases} \max\{f(\theta + \varepsilon); f(\theta - \varepsilon)\} & \text{if } \varepsilon \in]0; \min\{a_r; a_l\}], \\ \max\{M_r; f(\theta - \varepsilon)\} & \text{if } a_r < a_l \text{ and } \varepsilon \in]a_r; a_l], \\ \max\{M_l; f(\theta + \varepsilon)\} & \text{if } a_l < a_r \text{ and } \varepsilon \in]a_l; a_r], \\ \max\{M_r; M_l\} & \text{if } \varepsilon \in]\max\{a_r; a_l\}; +\infty[. \end{cases}$$

In Section 3 we prove in particular that G is an increasing function on $[0, +\infty[$ such that $G(0) = 0$. The fact that $f(\theta) > f(\theta + \varepsilon)$ for any $\varepsilon > 0$ and $f(\theta) > M_r$ then ensures that $c_r(\varepsilon) > 0$ for any $\varepsilon > 0$. Moreover, since the function $\varepsilon \mapsto (f(\theta) - f(\theta + \varepsilon))/2$ on $]0, a_r[$ and $\varepsilon \mapsto (f(\theta) - M_r)/2$ on $[a_r, +\infty[$ is nondecreasing on $]0, +\infty[$, the function c_r is nondecreasing on $]0, +\infty[$. In the same way, c_l and c are positive and nondecreasing functions on $]0, +\infty[$.

Two straightforward corollaries can be deduced from Theorem 1.

(i) First, in the case f is nonincreasing on $[\theta, +\infty[$ (that is, when $\mathcal{H}_r = \emptyset$), we have, for any $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}] \leq -\frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - f(\theta + \varepsilon)}{2}\right)$$

with $\lim_{\varepsilon \rightarrow +\infty} G([f(\theta) - f(\theta + \varepsilon)]/2) = G(f(\theta)/2)$. And, if the support of f is bounded, there exists ε_0 such that $G([f(\theta) - f(\theta + \varepsilon)]/2) = G(f(\theta)/2)$ for all $\varepsilon \geq \varepsilon_0$.

(ii) On the other hand, when all the moments of the random variable X_i are finite, the upper bound given by Theorem 1 holds for any $\alpha > 0$; we thus obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}] \leq \begin{cases} -G\left(\frac{f(\theta) - f(\theta + \varepsilon)}{2}\right) & \text{if } \varepsilon \in]0; a_r[, \\ -G\left(\frac{f(\theta) - M_r}{2}\right) & \text{if } \varepsilon \geq a_r. \end{cases}$$

Of course, similar properties hold for c_l and c as well.

3. Considerations on the functions ψ , I , and G . Since $|e^t - 1| \leq |t|e^{|t|}$ for all $t \in \mathbf{R}$, and thanks to the boundedness and integrability of K , we have

$$\int_{\mathbf{R}} |e^{uK(z)} - 1| dz \leq |u| e^{|u| \|K\|_\infty} \int_{\mathbf{R}} |K(z)| dz < \infty$$

which ensures that ψ is a finite function on $\mathbf{R}$. Moreover, it follows from classical analysis considerations that ψ is twice continuously differentiable, and

$$\psi'(u) = \int_{\mathbf{R}} K(z) e^{uK(z)} dz; \quad \psi''(u) = \int_{\mathbf{R}} K^2(z) e^{uK(z)} dz.$$

Since $\psi''(u) > 0$ for any $u \in \mathbf{R}$, ψ' is an increasing function on $\mathbf{R}$, and ψ is strictly convex. However, the range of ψ' depends on the nature of the kernel K ; to study the properties of I , and thus of g and G , we consider two cases.

Case 1: $K(x) \geq 0$ for all $x \in \mathbf{R}$. We have $\lim_{u \rightarrow -\infty} \psi'(u) = 0$ and $\lim_{u \rightarrow +\infty} \psi'(u) = +\infty$, so that the range of ψ' is $]0, +\infty[$. Moreover, $\lim_{u \rightarrow -\infty} \psi(u) = -|S|$, where $|S|$ denotes the Lebesgue measure of the support of K (eventually $|S| = +\infty$). This implies that $I(0) = |S|$ and $I(t) = +\infty$ when $t < 0$. Finally, for all $t > 0$,

$$I(t) = t(\psi')^{-1}(t) - \psi((\psi')^{-1}(t)). \quad (7)$$

This implies that $I'(t) = (\psi')^{-1}(t)$ for every $t > 0$, and I is thus strictly convex on $]0, +\infty[$. Since $\psi'(0) = 1$, we have $I'(1) = 0$, and $I(1) = 0$ is the unique strict global minimum of I on $]0, +\infty[$. It follows that the function g is finite on $] -f(\theta), +\infty[$ and $g \equiv +\infty$ on $] -\infty, -f(\theta)[$. Moreover, it is an increasing function on $]0, +\infty[$, a decreasing function on $] -f(\theta), 0[$, and $g(0) = 0$ is the unique strict global minimum of g . Since K is nonnegative, we have $G(\delta) = g(|\delta|)$ for all $\delta \in \mathbf{R}$ (see [5] or [14]); thus, G is continuous, increasing on $]0, +\infty[$ and decreasing on $] -\infty, 0[$, with a unique global minimum in 0, $G(0) = 0$.

Case 2: there exists $x \in \mathbf{R}$ such that $K(x) < 0$. Set $S_+ = \{z \in \mathbf{R}; K(z) > 0\}$ and $S_- = \{z \in \mathbf{R}; K(z) < 0\}$. Rewriting ψ' as

$$\psi'(u) = \int_{S_+} K(z) e^{uK(z)} dz + \int_{S_-} K(z) e^{uK(z)} dz,$$

we note that $\lim_{u \rightarrow -\infty} \psi'(u) = -\infty$ and $\lim_{u \rightarrow +\infty} \psi'(u) = +\infty$, so that the range of ψ' is $\mathbf{R}$ in this case, and formula (7) holds for every $t \in \mathbf{R}$. In this case, the rate I is strictly convex on $\mathbf{R}$ and $I(1) = 0$ is the unique strict global minimum of I on $\mathbf{R}$. It follows that the function g is finite on $\mathbf{R}$, decreasing on $] -\infty, 0[$, increasing on $]0, +\infty[$ and $g(0) = 0$ is its unique strict global minimum. Thus, in this case also, G is continuous, increasing on $]0, +\infty[$ and decreasing on $] -\infty, 0[$, with a unique global minimum in 0, $G(0) = 0$.

4. Proof of the results. The key idea to prove Theorem 1 is to relate the behavior of $\theta - \theta_n$ with the one of $f(\theta) - f(\theta_n)$. Then we establish a large deviations upper bound for $f(\theta) - f(\theta_n)$. More precisely, we need the following two lemmas, which are proved in Sections 4.1 and 4.2, respectively.

Lemma 1. *Set $\varepsilon > 0$; we have*

$$\mathbf{P}\{\theta_n - \theta \geq \varepsilon\} \leq \begin{cases} 2\mathbf{P}\{f(\theta) - f(\theta_n) \geq f(\theta) - f(\theta + \varepsilon)\} & \text{if } \varepsilon < a_r, \\ \mathbf{P}\{f(\theta) - f(\theta_n) \geq f(\theta) - M_r\} & \text{if } \varepsilon \geq a_r \end{cases}$$

and

$$\mathbf{P}\{\theta - \theta_n \geq \varepsilon\} \leq \begin{cases} 2\mathbf{P}\{f(\theta) - f(\theta_n) \geq f(\theta) - f(\theta - \varepsilon)\} & \text{if } \varepsilon < a_l, \\ \mathbf{P}\{f(\theta) - f(\theta_n) \geq f(\theta) - M_l\} & \text{if } \varepsilon \geq a_l. \end{cases}$$

Lemma 2. *For any $\delta > 0$,*

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln [\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\}] \leq -\frac{\alpha}{1+\alpha} G\left(\frac{\delta}{2}\right).$$

Let us now conclude the proof of Theorem 1. Set $\varepsilon \in]0, a_r[$; in view of Lemma 1, we have

$$\ln[\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}] \leq \ln 2 + \ln [\mathbf{P}\{f(\theta) - f(\theta_n) \geq f(\theta) - f(\theta + \varepsilon)\}].$$

The application of Lemma 2 then gives

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}] \leq -\frac{\alpha}{1+\alpha} G\left(\frac{f(\theta) - f(\theta + \varepsilon)}{2}\right).$$

The upper bounds of $\mathbf{P}\{\theta_n - \theta \geq \varepsilon\}$ for $\varepsilon \geq a_r$ on the one hand, and of $\mathbf{P}\{\theta - \theta_n \geq \varepsilon\}$ for $\varepsilon > 0$ on the other hand, are deduced from Lemmas 1 and 2 in the same way. Theorem 1 is proved.

We now prove Corollary 1. Note that

$$\begin{aligned} \mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\} &= \mathbf{P}\{\theta_n - \theta \geq \varepsilon\} + \mathbf{P}\{\theta - \theta_n \geq \varepsilon\} \\ &\leq 2 \max \left\{ \mathbf{P}\{\theta_n - \theta \geq \varepsilon\}; \mathbf{P}\{\theta - \theta_n \geq \varepsilon\} \right\}. \end{aligned}$$

Thus,

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln [\mathbf{P}\{|\theta_n - \theta| \geq \varepsilon\}] \leq \max\{-c_r(\varepsilon); -c_l(\varepsilon)\} \leq -\min\{c_r(\varepsilon); c_l(\varepsilon)\}.$$

Corollary 1 follows from the fact that G is an increasing function on $]0; +\infty[$ and from the definitions of c_r and c_l .

It remains to prove Lemmas 1 and 2.

4.1. Proof of Lemma 1. The definitions of M_r and a_r imply that f is nonincreasing on $[\theta; \theta + a_r]$; moreover, note that

$$\begin{cases} \theta_n \in [\theta; \theta + a_r] \implies f(\theta_n) \geq M_r, \\ \theta_n \geq \theta + a_r \implies f(\theta_n) \leq M_r \implies f(\theta) - f(\theta_n) \geq f(\theta) - M_r. \end{cases}$$

For $\varepsilon \in]0, a_r[$, we thus have

$$\begin{aligned} \mathbf{P}\{\theta_n - \theta \geq \varepsilon\} &= \mathbf{P}\{\theta_n - \theta \geq \varepsilon \text{ and } \theta_n \in [\theta; \theta + a_r]\} \\ &\quad + \mathbf{P}\{\theta_n - \theta \geq \varepsilon \text{ and } \theta_n \notin [\theta; \theta + a_r]\} \\ &\leq \mathbf{P}\{\theta_n \geq \theta + \varepsilon \text{ and } \theta_n \in [\theta; \theta + a_r]\} + \mathbf{P}\{\theta_n \geq \theta + a_r\} \end{aligned}$$

$$\begin{aligned}
&\leq \mathbf{P}\left\{f(\theta_n) \leq f(\theta + \varepsilon) \text{ and } \theta_n \in [\theta; \theta + a_r]\right\} \\
&\quad + \mathbf{P}\left\{f(\theta) - f(\theta_n) \geq f(\theta) - M_r\right\} \\
&\leq \mathbf{P}\left\{f(\theta) - f(\theta_n) \geq f(\theta) - f(\theta + \varepsilon)\right\} + \mathbf{P}\left\{f(\theta) - f(\theta_n) \geq f(\theta) - M_r\right\}.
\end{aligned}$$

Since $f(\theta) - f(\theta + \varepsilon) \leq f(\theta) - M_r$, it follows that

$$\mathbf{P}\{\theta_n - \theta \geq \varepsilon\} \leq 2\mathbf{P}\left\{f(\theta) - f(\theta_n) \geq f(\theta) - f(\theta + \varepsilon)\right\}.$$

On the other hand, for $\varepsilon \geq a_r$, we have

$$\mathbf{P}\{\theta_n - \theta \geq \varepsilon\} \leq \mathbf{P}\{\theta_n \geq \theta + a_r\} \leq \mathbf{P}\left\{f(\theta) - f(\theta_n) \geq f(\theta) - M_r\right\}$$

which concludes the proof of the first part of Lemma 1.

To prove the second part of Lemma 1, we note that f is nondecreasing on $[\theta - a_l; \theta]$, and that

$$\begin{cases} \theta_n \in [\theta - a_l; \theta] \implies f(\theta_n) \geq M_l, \\ \theta_n \leq \theta - a_l \implies f(\theta_n) \leq M_l \implies f(\theta) - f(\theta_n) \geq f(\theta) - M_l \end{cases}$$

and we proceed exactly as previously.

4.2. Proof of Lemma 2. The proof of Lemma 2 requires several steps. We first prove the following result.

Lemma 3. For any w_n such that $|\theta| \leq w_n$, and any $\delta > 0$, we have

$$\begin{aligned}
&\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\} \\
&\leq 3\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \frac{\delta}{2}\right\} + \mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2}\right\}.
\end{aligned}$$

Then we establish the following lemma.

Lemma 4. For any $b > 0$, set $w_n = \exp(bnh_n)$; we have

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln \left[\mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2}\right\} \right] \leq -b\alpha.$$

In order to obtain a large deviations upper bound for $(\sup_{|x| \leq w_n} |f_n(x) - f(x)|)$, we prove the following lemma.

Lemma 5. For any $\delta > 0$, we have

$$\limsup_{n \rightarrow +\infty} \frac{1}{nh_n} \ln \sup_{x \in \mathbf{R}} \mathbf{P}\{|f_n(x) - f(x)| \geq \delta\} \leq -G(\delta).$$

Then we deduce the following result from Lemma 5.

Lemma 6. For any $b > 0$, set $w_n = \exp(bnh_n)$; for any $\delta > 0$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln \left[\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} \right] \leq b - G(\delta).$$

It follows from Lemmas 3, 4, and 6 that, for any $b > 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\}] \leq \max \left\{ -b\alpha; b - G\left(\frac{\delta}{2}\right) \right\}.$$

Thus

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln[\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\}] \leq \inf_{b > 0} \left(\max \left\{ -b\alpha; b - G\left(\frac{\delta}{2}\right) \right\} \right).$$

Since $\inf_{b > 0} (\max\{-b\alpha; b - G(\delta/2)\})$ is achieved by choosing $b = G(\delta/2)/(1 + \alpha)$ and equals $-\alpha G(\delta/2)/(1 + \alpha)$, the proof of Lemma 2 is completed.

It remains to prove Lemmas 3, 4, 5, and 6.

4.2.1. Proof of Lemma 3. If $\delta > f(\theta)$, then $\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\} = 0$, and Lemma 3 is straightforward; from now on, we set $\delta \in]0, f(\theta)]$. Since $f_n(\theta_n) \geq f_n(x)$ for any $x \in \mathbf{R}$, we have

$$f(\theta) - f(\theta_n) = f(\theta) - f_n(\theta) + f_n(\theta) - f(\theta_n) \leq |f(\theta) - f_n(\theta)| + f_n(\theta_n) - f(\theta_n).$$

It follows that

$$\begin{aligned} \mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\} &\leq \mathbf{P}\left\{|f(\theta) - f_n(\theta)| \geq \frac{\delta}{2}\right\} + \mathbf{P}\left\{f_n(\theta_n) - f(\theta_n) \geq \frac{\delta}{2}\right\} \\ &\leq \mathbf{P}\left\{|f(\theta) - f_n(\theta)| \geq \frac{\delta}{2}\right\} + \mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \frac{\delta}{2}\right\} \\ &\quad + \mathbf{P}\left\{f_n(\theta_n) - f(\theta_n) \geq \frac{\delta}{2} \text{ and } |\theta_n| \geq w_n\right\} \end{aligned}$$

and, since $|\theta| \leq w_n$, we find

$$\mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\} \leq 2\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \frac{\delta}{2}\right\} + \mathbf{P}\{|\theta_n| \geq w_n\}.$$

However, since $\delta \leq f(\theta)$ and $f_n(\theta_n) \geq f_n(\theta)$, we have

$$\begin{aligned} \mathbf{P}\{|\theta_n| \geq w_n\} &\leq \mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta) \geq \frac{f(\theta)}{2}\right\} + \mathbf{P}\left\{f_n(\theta) \leq \frac{f(\theta)}{2}\right\} \\ &\leq \mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta) \geq \frac{f(\theta)}{2}\right\} + \mathbf{P}\left\{f(\theta) - f_n(\theta) \geq \frac{f(\theta)}{2}\right\} \\ &\leq \mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2}\right\} + \mathbf{P}\left\{|f(\theta) - f_n(\theta)| \geq \frac{\delta}{2}\right\}. \end{aligned}$$

Thus,

$$\begin{aligned} \mathbf{P}\{f(\theta) - f(\theta_n) \geq \delta\} &\leq 3\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \frac{\delta}{2}\right\} \\ &\quad + \mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2}\right\}. \end{aligned}$$

4.2.2. Proof of Lemma 4. Set $\kappa = \sup_{x \in \mathbf{R}} |x|^\gamma |K(x)|$; we have

$$\begin{aligned} f_n(\theta_n) \geq \frac{f(\theta)}{2} &\implies \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{X_i - \theta_n}{h_n}\right) \geq \frac{f(\theta)}{2} \\ &\implies \exists i \in \{1, \dots, n\} \text{ such that } K\left(\frac{X_i - \theta_n}{h_n}\right) \geq h_n \frac{f(\theta)}{2} \\ &\implies \exists i \in \{1, \dots, n\} \text{ such that } \kappa h_n^\gamma \geq h_n |X_i - \theta_n|^\gamma \frac{f(\theta)}{2} \\ &\implies \exists i \in \{1, \dots, n\} \text{ such that } |X_i| - |\theta_n| \leq \left(\frac{2\kappa h_n^{\gamma-1}}{f(\theta)}\right)^{1/\gamma} \\ &\implies \exists i \in \{1, \dots, n\} \text{ such that } |X_i| \geq |\theta_n| - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)}\right)^{1/\gamma}. \end{aligned}$$

It follows that

$$\mathbf{P}\left\{|\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2}\right\} \leq \mathbf{P}\left\{\max_{1 \leq i \leq n} |X_i| \geq w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)}\right)^{1/\gamma}\right\}.$$

Note that

$$w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)}\right)^{1/\gamma} = w_n \left[1 - \left(\frac{2\kappa}{f(\theta)}\right)^{1/\gamma} \exp\left(-bnh_n \left[1 - \left(1 - \frac{1}{\gamma}\right) \frac{\ln h_n}{bnh_n}\right]\right)\right]$$

and assumption (A3) ensures that

$$\lim_{n \rightarrow \infty} \frac{1}{w_n} \left[w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)} \right)^{1/\gamma} \right] = 1. \quad (8)$$

We deduce in particular that, for n large enough, we have $w_n - h_n^{1-1/\gamma} (2\kappa/f(\theta))^{1/\gamma} > 0$ and thus

$$\begin{aligned} \mathbf{P} \left\{ |\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2} \right\} &\leq \mathbf{P} \left\{ \sum_{i=1}^n |X_i|^\alpha \geq \left[w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)} \right)^{1/\gamma} \right]^\alpha \right\} \\ &\leq n \left(w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)} \right)^{1/\gamma} \right)^{-\alpha} \mathbf{E}|X_1|^\alpha. \end{aligned}$$

It follows that

$$\begin{aligned} \ln \left[\mathbf{P} \left\{ |\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2} \right\} \right] \\ \leq -\alpha \ln \left[w_n - h_n^{1-1/\gamma} \left(\frac{2\kappa}{f(\theta)} \right)^{1/\gamma} \right] + \ln n + \ln[\mathbf{E}|X_1|^\alpha]. \end{aligned}$$

Now, since (A3) ensures $\lim_{n \rightarrow \infty} (\ln n)/(nh_n) = 0$ and since $w_n = \exp(bnh_n)$, we obtain, in view of (8),

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln \left[\mathbf{P} \left\{ |\theta_n| \geq w_n \text{ and } f_n(\theta_n) \geq \frac{f(\theta)}{2} \right\} \right] \leq -b\alpha.$$

4.2.3. Proof of Lemma 5. Large deviations principles for the kernel density estimator have been first proved in [5] in the independent case, then generalized in [14] for Markov chains with a unique invariant probability measure. However, the assumptions made by the two former authors (in particular K nonnegative and f Lipschitz continuous) are more restrictive than our conditions (A1)–(A2).

For any $x \in \mathbf{R}$, $u \in \mathbf{R}$, and $n \geq 1$, the moment generating function of the estimator $f_n(x)$ is defined by

$$\Lambda_{n,x}(u) = \frac{1}{nh_n} \ln[\mathbf{E} \exp(u nh_n f_n(x))].$$

Set $\Delta_n(u) = \sup_{x \in \mathbf{R}} |\Lambda_{n,x}(u) - f(x) \psi(u)|$, where ψ is defined by (3); to prove Lemma 5, we first establish that

$$\lim_{n \rightarrow \infty} \Delta_n(u) = 0. \quad (9)$$

Since

$$\begin{aligned} \Lambda_{n,x}(u) &= \frac{1}{h_n} \ln \left[\mathbf{E} \exp \left(uK \left(\frac{X_1 - x}{h_n} \right) \right) \right] = \frac{1}{h_n} \ln \left[1 + \int_{\mathbf{R}} (e^{uK((y-x)/h_n)} - 1) f(y) dy \right] \\ &= \frac{1}{h_n} \ln \left[1 + h_n \int_{\mathbf{R}} (e^{uK(z)} - 1) f(x + h_n z) dz \right], \end{aligned}$$

we have

$$\begin{aligned} |\Lambda_{n,x}(u) - f(x) \psi(u)| &\leq \frac{1}{h_n} \left| h_n \Lambda_{n,x}(u) - h_n \int_{\mathbf{R}} (e^{uK(z)} - 1) f(x + h_n z) dz \right| \\ &\quad + \left| \int_{\mathbf{R}} (e^{uK(z)} - 1) f(x + h_n z) dz - f(x) \psi(u) \right|. \end{aligned}$$

Now, since $|\ln(1+t) - t| \leq t^2$ for t sufficiently small, and noting that $|\int_{\mathbf{R}} (e^{uK(z)} - 1) f(x + h_n z) dz| \leq f(\theta) \int_{\mathbf{R}} |e^{uK(z)} - 1| dz < +\infty$, it follows that, for n large enough,

$$\begin{aligned} |\Lambda_{n,x}(u) - f(x) \psi(u)| &\leq h_n \left| \int_{\mathbf{R}} (e^{uK(z)} - 1) f(x + h_n z) dz \right|^2 \\ &+ \int_{\mathbf{R}} |e^{uK(z)} - 1| |f(x + h_n z) - f(x)| dz \leq h_n \|f\|_{\infty}^2 \left(\int_{\mathbf{R}} |e^{uK(z)} - 1| dz \right)^2 \\ &+ \int_{\mathbf{R}} |e^{uK(z)} - 1| \left(\sup_{x \in \mathbf{R}} |f(x + h_n z) - f(x)| \right) dz. \end{aligned}$$

Since f is bounded and uniformly continuous and since $e^{uK(\cdot)} - 1$ is integrable, the application of Lebesgue's convergence theorem gives (9).

Now, set $\delta > 0$; for any $u > 0$ and any $x \in \mathbf{R}$, we have

$$\begin{aligned} \mathbf{P}\{f_n(x) - f(x) \geq \delta\} &= \mathbf{P}\left\{\exp[nh_n u(f_n(x) - f(x))] \geq \exp[nh_n u\delta]\right\} \\ &\leq \exp[-nh_n u\delta] \mathbf{E}\left[\exp[nh_n u(f_n(x) - f(x))]\right] \\ &\leq \exp\left[-nh_n(u\delta - \Lambda_{n,x}(u) + u f(x))\right] \\ &\leq \exp\left[-nh_n(u\delta - f(x)(\psi(u) - u))\right] \\ &\quad \times \exp\left[nh_n(\Lambda_{n,x}(u) - f(x)\psi(u))\right]. \end{aligned}$$

Since $\psi(u) - u = \int (e^{uK(z)} - 1 - uK(z)) dz > 0$ for every $u \neq 0$, it follows that, for all $u > 0$,

$$\sup_{x \in \mathbf{R}} \mathbf{P}\{f_n(x) - f(x) \geq \delta\} \leq \exp\left(-nh_n\{u\delta - f(\theta)(\psi(u) - u)\}\right) \exp(nh_n \Delta_n(u)). \quad (10)$$

The supremum of the function

$$\phi(u) = u\delta - f(\theta)(\psi(u) - u) = f(\theta) \left[u \left(1 + \frac{\delta}{f(\theta)} \right) - \psi(u) \right]$$

is achieved at the point $u_{\delta} = \psi'^{-1}(1 + \delta/f(\theta))$ and, by (5),

$$\phi(u_{\delta}) = f(\theta) I \left(1 + \frac{\delta}{f(\theta)} \right) = g(\delta).$$

Since ψ' is increasing and $\psi'(0) = 1$, we have $u_{\delta} > 0$, and the application of (10) at u_{δ} gives

$$\sup_{x \in \mathbf{R}} \mathbf{P}\{f_n(x) - f(x) \geq \delta\} \leq \exp(-nh_n\{g(\delta) + \Delta_n(u_{\delta})\}).$$

In view of (9), we obtain

$$\limsup_{n \rightarrow +\infty} \frac{1}{nh_n} \ln \sup_{x \in \mathbf{R}} \mathbf{P}\{f_n(x) - f(x) \geq \delta\} \leq -g(\delta). \quad (11)$$

In the same way, we prove that, for any $u < 0$, we have

$$\sup_{x \in \mathbf{R}} \mathbf{P}\{f_n(x) - f(x) \leq -\delta\} \leq \exp\left(-nh_n\{-u\delta - f(\theta)(\psi(u) - u)\}\right) \exp(nh_n \Delta_n(u)).$$

Noting that the supremum of the function

$$\chi(u) = -u\delta - f(\theta)(\psi(u) - u)$$

is achieved at the point $u_{-\delta} = \psi'^{-1}(1 - \delta/f(\theta)) < 0$ and that $\chi(u_{-\delta}) = g(-\delta)$, we obtain

$$\limsup_{n \rightarrow +\infty} \frac{1}{nh_n} \ln \sup_{x \in \mathbf{R}} \mathbf{P}\{f(x) - f_n(x) \geq \delta\} \leq -g(-\delta). \quad (12)$$

The combination of (6), (11), and (12) then gives Lemma 5.

4.2.4 Proof of Lemma 6. Set $\rho \in]0, \delta/2[$, $l(n) = (\rho h_n^{1+\beta}/2)^{1/\beta}$, and let $N(n)$ be the integer part of $1 + w_n/l(n)$; there exist $x_1^{(n)}, \dots, x_{N(n)}^{(n)}$ such that

$$\{x \in \mathbf{R}; |x| \leq w_n\} \subset \bigcup_{i=1}^{N(n)} I_i^{(n)}, \quad I_i^{(n)} = [x_i^{(n)} - l(n), x_i^{(n)} + l(n)].$$

Thus, we have

$$\begin{aligned} \mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} &\leq \sum_{i=1}^{N(n)} \mathbf{P}\left\{\sup_{|x| \in I_i^{(n)}} |f_n(x) - f(x)| \geq \delta\right\} \\ &\leq N(n) \max_{1 \leq i \leq N(n)} \mathbf{P}\left\{\sup_{|x| \in I_i^{(n)}} |f_n(x) - f(x)| \geq \delta\right\}. \end{aligned} \quad (13)$$

Set $x \in I_i^{(n)}$, and note that

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_n(x_i^{(n)})| + |f_n(x_i^{(n)}) - f(x_i^{(n)})| \\ &\quad + |f(x_i^{(n)}) - f(x)|. \end{aligned}$$

Since f is uniformly continuous, and since $\lim_{n \rightarrow \infty} l(n) = 0$, we have, for n large enough,

$$|f(x_i^{(n)}) - f(x)| \leq \frac{\rho}{2}$$

and, since K is Hölder continuous with exponent β ,

$$|f_n(x) - f_n(x_i^{(n)})| \leq \frac{1}{h_n} \left| \frac{x - x_i^{(n)}}{h_n} \right|^\beta \leq \frac{\rho}{2}.$$

It follows that, for n large enough,

$$\mathbf{P}\left\{\sup_{|x| \in I_i^{(n)}} |f_n(x) - f(x)| \geq \delta\right\} \leq \mathbf{P}\{|f_n(x_i^{(n)}) - f(x_i^{(n)})| \geq \delta - \rho\}$$

and therefore, in view of (13),

$$\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} \leq N(n) \sup_{x \in \mathbf{R}} \mathbf{P}\{|f_n(x) - f(x)| \geq \delta - \rho\}$$

which gives

$$\begin{aligned} &\frac{1}{nh_n} \ln \left[\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} \right] \\ &\leq \frac{\ln[N(n)]}{nh_n} + \frac{1}{nh_n} \ln \left[\sup_{x \in \mathbf{R}} \mathbf{P}\{|f_n(x) - f(x)| \geq \delta - \rho\} \right]. \end{aligned}$$

Now, in view of assumption (A3) ii), we have $\lim_{n \rightarrow \infty} \ln[N(n)]/(nh_n) = b$. Thus the application of Lemma 5 gives

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln \left[\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} \right] \leq b - G(\delta - \rho).$$

Since this last upper bound holds for any $\rho \in]0, \delta/2[$, the continuity of G ensures that

$$\limsup_{n \rightarrow \infty} \frac{1}{nh_n} \ln \left[\mathbf{P}\left\{\sup_{|x| \leq w_n} |f_n(x) - f(x)| \geq \delta\right\} \right] \leq b - G(\delta).$$

REFERENCES

1. Eddy W. F. Optimum kernel estimators of the mode. — Ann. Statist., 1980, v. 8, № 4, p. 870–882.
2. Eddy W. F. The asymptotic distributions of kernel estimators of the mode. — Z. Wahrscheinlichkeitstheor. Verw. Geb., 1982, v. 59, № 3, p. 279–290.
3. Grund B., Hall P. On the minimisation of L^p error in mode estimation. — Ann. Statist., 1995, v. 23, № 6, p. 2264–2284.
4. Конаков В. Д. Об асимптотической нормальности моды многомерных распределений. — Теория вероятн. и ее примен., 1973, т. 18, в. 4, с. 836–842.
5. Louani D. Large deviations limit theorems for the kernel density estimator. — Scand. J. Statist., 1998, v. 25, № 1, p. 243–253.
6. Mokkadem A., Pelletier M. The law of the iterated logarithm for the multivariable kernel mode estimator. — ESAIM Probab. Statist., 2003, v. 7, p. 1–21.
7. Надарая Э. А. О непараметрических оценках плотности вероятности и регрессии. — Теория вероятн. и ее примен., 1965, т. 10, в. 1, p. 199–203.
8. Новак С. Ю. О моде неизвестного вероятностного распределения. — Теория вероятн. и ее примен., 1999, т. 44, в. 1, с. 119–123.
9. Parzen E. On estimation on of a probability density function and mode. — Ann. Math. Statist., 1962, v. 33, p. 1065–1076.
10. Romano J. On weak convergence and optimality of kernel density estimates of the mode. — Ann. Statist., 1988, v. 16, № 2, p. 629–647.
11. Silverman B. W. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. — Ann. Statist., 1978, v. 6, № 1, p. 177–184.
12. Van Ryzin J. On strong consistency of density estimates. — Ann. Math. Statist., 1969, v. 40, p. 1765–1772.
13. Yamato H. Sequential estimation of a continuous probability density function and mode. — Bull. Math. Statist., 1971, v. 14, p. 1–12; corrections: ibid., 1972, v. 15, p. 133.
14. Worms J. Moderate deviations of some dependent variables. II: Some kernel estimators. — Math. Methods Statist., 2001, v. 10, № 2, p. 161–193.

Поступила в редакцию
20.VIII.2002