SHARP LARGE DEVIATIONS IN NONPARAMETRIC ESTIMATION

Cyrille JOUTARD[1]

Abstract: Large deviation results for nonparametric estimators have been given by Louani (in particular for the kernel density estimator and the Nadaraya-Watson estimator). We complete these works by proving large deviation principles for these two estimators under less restrictive assumptions. Moreover we give sharp large deviation results. Two cases depending on the support of the kernel are considered. For the proof we need an Edgeworth expansion. This last one is proved using a version of Cramer's condition.

Key words: kernel density estimator, Nadaraya-Watson estimator, large deviation principle, Cramer's condition, Edgeworth expansion.

Résumé: Des résultats de grandes déviations ont été obtenus par Louani (en particulier pour l'estimateur à noyau de la densité et l'estimateur de Nadaraya-Watson). Nous complétons ces travaux en prouvant des principes de grandes déviations pour ces deux estimateurs sous des conditions moins restrictives. De plus nous donnons des résultats de grandes déviations précises. Deux cas dépendant du support du noyau sont considérés. Pour la démonstration nous avons besoin d'un développement de Edgeworth. Ce dernier est prouvé en utilisant une version de la condition de Cramer.

Mots-clés: estimateur à noyau de la densité, estimateur de Nadaraya-Watson, principe de grandes déviations, condition de Cramer, développement de Edgeworth.

[1]CREST, Laboratoire de Statistique, 3 avenue Pierre Larousse, 92240 Malakoff Cedex, France. E-mail: joutard@ensae.fr.
Université Paul Sabatier, Laboratoire de Statistique et Probabilités, 118 route de Narbonne, 31062 Toulouse Cedex, France.

1 Introduction

Before presenting our results, we recall some generalities on large deviations and nonparametric estimation. One of the most classical examples in large deviation theory involves the empirical mean of a n-sample. More precisely, let $(Y_1, \ldots, Y_n)$ be independent and identically distributed (i.i.d.) real random variables with zero mean. Consider the empirical mean

$$\bar{Y}_n = \frac{1}{n} \sum_{j=1}^{n} Y_j.$$

By the weak law of large numbers, as $n \to \infty$, $\bar{Y}_n$ converges towards $\mathbb{E}Y_1 = 0$ in probability. Hence if A is a measurable set in $\mathbb{R}$ and $0 \notin A$, $\mathbb{P}(\bar{Y}_n \in A)$ converges to 0. Large deviation theory gives a logarithmic equivalent for this deviation probability. We recall the basic definition of a Large Deviation Principle (LDP) on $\mathbb{R}$ (see, for example, [6] or [7]).

Definition 1.1. *Let (ϵ_n) be a sequence of positive real numbers such that $\epsilon_n \searrow 0$ as $n \to \infty$. We say that the sequence of real random variables (Z_n) satisfies a large deviation principle (LDP) with speed (ϵ_n) and rate function $I : \mathbb{R} \to [0, \infty]$ (which means that I is a lower semi-continuous function) if*

1) For any open set $G \subset \mathbb{R}$,

$$\liminf_{n \to \infty} \epsilon_n \log \mathbb{P}(Z_n \in G) \geq - \inf_{x \in G} I(x). \quad \text{(lower bound)}$$

2) For any closed set $F \subset \mathbb{R}$,

$$\limsup_{n \to \infty} \epsilon_n \log \mathbb{P}(Z_n \in F) \leq - \inf_{x \in F} I(x). \quad \text{(upper bound)}$$

We say that I is a good rate function if its level sets are compact subsets of $\mathbb{R}$. In general, it could be more interesting to have a full expansion of the tail probabilities (for example for the tests, see [9]). Bahadur and Rao [1] studied this kind of problem for the empirical mean. In particular under suitable assumptions, they obtained the following expansions:

$$\mathbb{P}(\bar{Y}_n \geq c) = \frac{\exp(-nI(c))}{(2\pi n)^{1/2} \sigma_c \tau_c} [1 + o(1)], \quad c > 0 \tag{1}$$

or

$$\mathbb{P}(\bar{Y}_n \leq c) = \frac{\exp(-nI(c))}{(2\pi n)^{1/2} \sigma_c \tau_c} [1 + o(1)], \quad c < 0 \tag{2}$$

where $\tau_c > 0$ and $\sigma_c > 0$ are parameters depending on c.
If (1) or (2) holds for $c \in \mathbb{R}$, we say that $\bar{Y}_n$ satisfies a Local Sharp Large Deviation

Principle (LSLDP) of order 0 at c (for more precision see [3]).

In this paper we study large deviation properties of nonparametric estimators. Louani gave large deviation results for the kernel density estimator (see [10]). The frame is the following: Let $(X_1, \ldots, X_n)$ be an i.i.d. sample with common density function f on the real line. Let K be a positive kernel and h_n a sequence of positive real numbers (h_n is called the bandwidth). Then we define for any $n \in \mathbb{N}^*$, the kernel density estimator (see [14]) of $f(x)$, $x \in \mathbb{R}$, by

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right).$$

Louani also studied large deviation probabilities for the kernel regression estimator (Nadaraya-Watson estimator) (see [11]). The frame is the following:
Let $((X_1, Y_1), \ldots, (X_n, Y_n))$ be an i.i.d bivariate sample such that $\mathbb{E}|Y_1| < \infty$. Then the Nadaraya-Watson estimator ([13]; [15]) of the regression function $r(z) = \mathbb{E}(Y_1|X_1 = z)$, $z \in \mathbb{R}$, is defined for any $n \in \mathbb{N}^*$ by

$$\hat{r}_n(z) = \frac{\sum_{j=1}^{n} Y_j K\left(\frac{z - X_j}{h_n}\right)}{\sum_{j=1}^{n} K\left(\frac{z - X_j}{h_n}\right)}. \tag{3}$$

Menneteau [12] studied uniform large deviations for local empirical processes and his results can be used to obtain those of Louani.

In this paper, we prove large deviation principles for the two estimators under more general assumptions than in [10] and [11]. We also give a LSLDP for both of them (this means that we establish expansions of the same kind as (1) and (2)). To obtain the LSLDP for the kernel density estimator $\hat{f}_n(x)$, $x \in \mathbb{R}$, we use an Edgeworth expansion for a distribution function ([2]; [9]). First we assume $f(x) > 0$ and we consider two cases depending on the support of the kernel. The first case concerns kernels with compact support and the second one kernels with unbounded support. The Edgeworth expansion has been given by Hall [8] in the first case and we give the proof in the second case. A version of Cramer's condition [5] for each case is needed to get the Edgeworth expansion. We also give a LSLDP for $\hat{f}_n(x)$ when $f(x) = 0$. We follow the same approach to prove the LSLDP for the Nadaraya-Watson estimator.

The paper is organized as follows: In Section 2, we deal with the large deviation results for the two estimators. In Section 3, we give the LSLDP satisfied by the kernel density estimator. Finally in Section 4, the case of the Nadaraya-Watson estimator is treated quicker, the proof being quite similar.

2 Large Deviations in Nonparametric Estimation

In the paper, we consider a positive kernel K such that $\int_{\mathbb{R}} K(z)dz = 1$ and a sequence of real numbers $h_n > 0$ satisfying $h_n \to 0$ as $n \to \infty$.

2.1 Large deviations for the kernel density estimator

2.1.1 Notations and assumptions

Before stating our result, we introduce some notations and assumptions.

Let X be a real random variable with density function f. Let $(X_1, \ldots, X_n)$ be an i.i.d. sample having the same distribution as X.

For the kernel K and the bandwidth h_n, we introduce the notation

$$K_{h_n}(x) = h_n^{-1} K(x/h_n).$$

The Lebesgue measure on $\mathbb{R}$ is noted λ and μ_K is the distribution function such that $d\mu_K(z) = K(z)dz$. We also consider the function introduced in [10]

$$I(t) = \int_{\mathbb{R}} (\exp(tK(z)) - 1)dz \tag{4}$$

which is used to define the rate function associated with the LDP.

Now we consider two assumptions on K and f (compare with [10]):

(A.1) $\varphi_0(t) = \int_{\mathbb{R}} K(z) \exp(tK(z))dz$ is defined on $]-\infty, a]$, $a > 0$.

(A.2) f is continuous on $\mathbb{R}$ and $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)| < \infty$.

Remark 2.1

(i) Using assumption $(A.1)$, $I \in C^\infty(]-\infty, a[)$. Indeed $I(t) = \int_0^t \varphi_0(s)ds$ $(t \in]-\infty, a])$ where $\varphi_0 \in C^\infty(]-\infty, a[)$ (as it is the Laplace transform of $\mu_K \circ K^{-1}$). We note I' the derivative of I.

(ii) We can prove the large deviation result for $\hat{f}_n(x), x \in \mathbb{R}$, if f is continuous only on a neighborhood of x but for reasons of simplicity, we assume the continuity on $\mathbb{R}$. The same kind of remark concerning the regularity of the density function is also valid for the other results in Section 2.2, Section 3 and Section 4.

2.1.2 Result and proof

Large deviation results for the kernel density estimator $\hat{f}_n$ have been given by Louani [10] under the restrictive assumption that I is finite everywhere. In this case an exponential change of probability can be used to prove the lower bound. We give the LDP in a more general case using the assumptions above. Indeed, we have the following theorem

Theorem 2.1. *Assume that* $(A.1)-(A.2)$ *hold and that* $\lim_{n\to\infty} nh_n = \infty$. *Then, for any* $x \in \mathbb{R}$, $\hat{f}_n(x) - f(x)$ *satisfies a LDP with speed* $(nh_n)^{-1}$ *and good rate function* $\Lambda^*(u) = \sup_{t\in\mathbb{R}}\{t(u + f(x)) - f(x)I(t)\}$.

Remark 2.2 As we have to establish a LDP on $\mathbb{R}$, it is sufficient to prove the convergence for the following tail probabilities (see, for example, [6]): For any $x \in \mathbb{R}$, any $\alpha > 0$, we have

$$\lim_{n\to\infty} \frac{1}{nh_n} \log \mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) = -\Lambda^*(\alpha) \tag{5}$$

where $\Lambda^*(\alpha) = \sup_{t\in]0,a]}\{t(\alpha + f(x)) - f(x)I(t)\}$,
and

$$\lim_{n\to\infty} \frac{1}{nh_n} \log \mathbb{P}(\hat{f}_n(x) - f(x) < -\alpha) = -\Lambda^*(-\alpha) \tag{6}$$

where $\Lambda^*(-\alpha) = \sup_{t\in]0,\infty[}\{t(\alpha - f(x)) - f(x)I(-t)\}$.
For the left tail probability (6), the rate function is equal to infinity if $\alpha > f(x)$. For the right tail probability (5), it is equal to $a\alpha$ if $f(x) = 0$.

Proof. For the left tail probability (6), the proof is the same as in [10] since I is defined on $\mathbb{R}_-$.
So we only give the proof of the lower bound for the right tail probability (5) (for the upper bound, we use the Chebycheff inequality).
The lower bound is usually proved in the large deviation theory by an exponential change of probability using a parameter (see, for example, [6]).
Let $x \in \mathbb{R}$ such that $f(x) > 0$ and let $\alpha > 0$.
If we consider the rate function Λ^* evaluated at α, the existence of a parameter $\tau \in \mathbb{R}$ satisfying

$$I'(\tau) = \varphi_0(\tau) = (\alpha + f(x))/f(x), \tag{7}$$

is not always guaranteed.
Indeed, with assumption $(A.1)$, we know that

$$\varphi_0(t) = \begin{cases} \int_{\mathbb{R}} K(z)\exp(tK(z))dz & \text{if } t \in]-\infty, a] \\ \infty & \text{elsewhere} \end{cases}$$

Then it is clear that there exists α^* $((\alpha^* + f(x))/f(x) = I'(a))$ such that for all $\alpha > \alpha^*$, no τ satisfying (7) exists.

4

To deal with this problem, we shall use the same arguments as in the proof of the lower bound in [6, Theorem 2.2.3]:

First we suppose that K is bounded everywhere by a constant $M > 0$. In this case the function I is finite and differentiable on $\mathbb{R}$ (see Remark 2.1 (i)).

Then we know that a parameter τ satisfying (7) exists. We use assumption $(A.2)$ to prove (5) as in [10].

Now, as in [6], we turn to the general case when K is not necessary bounded. We consider

$$
\begin{aligned}
\Lambda_{M,\alpha}(t) &= \liminf_{n\to\infty} \frac{1}{h_n} \log \left(\int_{z:K\left(\frac{x-z}{h_n}\right)\leq M} \exp\left(tK\left(\frac{x-z}{h_n} \right) \right) f(z)dz \right) - t(\alpha + f(x)) \\
&= \liminf_{n\to\infty} \frac{1}{h_n} \log \left(1 - h_n \int_{z:K(z)>M} f(x-h_n z)dz \right. \\
&\quad + h_n \int_{z:K(z)\leq M} (\exp(tK(z)) - 1)f(x-h_n z)dz \bigg) - t(\alpha + f(x)).
\end{aligned}
$$

Using the assumptions $(A.1) - (A.2)$ and Remark 2.1 (i), we get

$$
\Lambda_{M,\alpha}(t) = -f(x) \int_{z:K(z)>M} dz + f(x) \int_{z:K(z)\leq M} (\exp(tK(z)) - 1)dz - t(\alpha + f(x))
$$

which is finite everywhere for any $M > 0$.

We fix M large enough to have $\lambda\{\{z : K(z) > 0\} \cap \{z : K(z) \leq M\}\} > 0$. Hence $\lim_{|t|\to\infty} \Lambda_{M,\alpha}(t) = \infty$ for M large enough.

For the rest of the proof we follow the same approach as in [6] by using the law of $\hat{f}_n(x) - f(x) - \alpha$ conditioned on $\{K((x - X_j)/h_n) \leq M, j = 1,.., n\}$. For this last one, the lower bound is proved as in the case where K is bounded everywhere. $\square$

2.2 Large deviations for the Nadaraya-Watson estimator

2.2.1 Notations and assumptions

Let (X, Y) be a couple of real random variables with joint density function $f(x, y)$. Consider the regression model:

$$
Y_i = r(X_i) + \epsilon_i, \quad i = 1,.., n.
$$

$((X_1, Y_1), \ldots, (X_n, Y_n))$ is an i.i.d. bivariate sample having the same distribution as (X, Y). $(\epsilon_i)_{i=1,..,n}$ are i.i.d. random variables.

For each $i \in \{1, \cdots, n\}$, ϵ_i is assumed to be independent of X_i and to have $\mathbb{R}$ as support.

Let z_0 be a fixed point in $\mathbb{R}$. Let f_X be the marginal density function of X, f_1' the derivative of f with respect to the first variable and $f(y|z_0)$ the conditional density function of Y when $X = z_0$.

We suppose $\mathbb{E}|Y| < \infty$. Then the Nadaraya-Watson estimator of the regression function $r(z) = \mathbb{E}(Y_1|X_1 = z)$, $z \in \mathbb{R}$, is defined by (3).

We also define for any $\alpha \in \mathbb{R}$, the random variables

$$W_{jn,\alpha} = (Y_j - r(z_0) - \alpha) K\left(\frac{z_0 - X_j}{h_n}\right), \quad j = 1, .., n,$$

and we consider the function introduced in [11]

$$I_{r,\alpha}(t) = \int_{\mathbb{R}^2} [\exp(t(y - r(z_0) - \alpha)K(x)) - 1]f(z_0, y)dxdy. \tag{8}$$

We give the following assumptions for any $\alpha \in \mathbb{R}$:

$(A_r.1)$ $\varphi_{r,\alpha}(t) = \int_{\mathbb{R}^2} |y - r(z_0) - \alpha|K(x) \exp[t(y - r(z_0) - \alpha)K(x)]f(z_0, y)dxdy$ is defined on $[-a_1, a_2]$ where $a_1 > 0$ and $a_2 > 0$ do not depend on α.

$(A_r.2)$ $\|f_X\|_\infty < \infty$ and $\sup_{u \in \mathbb{R}} \int_{\mathbb{R}} |y|f(u, y)dy < \infty$.

$(A_r.3)$ f is continuous with respect to the first variable on $\mathbb{R}$.

$(A_r.4)$ For all $t \in [-a_1, a_2]$,
$\sup_{u \in \mathbb{R}} \int_{\mathbb{R}^2} |\exp[t(y - r(z_0) - \alpha)K(x)] - 1|f(u, y)dxdy < \infty$.

$(A_r.5)$ $K(x) > 0$ for all $x \in \mathbb{R}$.

Remark 2.3

(i) The last assumption $(A_r.5)$ is only needed to write the left tail probability (9) by using the sum of the random variables $W_{jn,\alpha}$, $j = 1, .., n$ (see [11]). This assumption is not necessary for the right tail probability (10).

(ii) We have assumed that a_1 and a_2 do not depend on α to obtain a LDP in the sense of Definition 1.1. If they do, then we only have large deviation results for the tail probabilities.

(iii) As in Remark 2.1 (i), assumption $(A_r.1)$ assures that $I_{r,\alpha} \in C^\infty(]-a_1, a_2[)$. Indeed $\varphi_{r,\alpha}$ is proportional to the Laplace transform of $(U - r(z_0) - \alpha)K(V)$ where (U, V) is a couple of random variables admitting for density with respect to the Lebesgue measure on $\mathbb{R}^2$

$$\frac{|v - r(z_0) - \alpha|K(u)f(z_0, v)}{\mu_\alpha(z_0)}dudv,$$

with $\mu_\alpha(z_0) = \int_{\mathbb{R}^2} |v - r(z_0) - \alpha|K(u)f(z_0, v)dudv = \varphi_{r,\alpha}(0) < \infty$. We note $I_{r,\alpha}'$ the derivative of $I_{r,\alpha}$.

6

(iv) If we suppose $\varphi_{r,\alpha}$ (and therefore $I_{r,\alpha}$) defined on $]-a_1, a_2[$ (as in Section 4) and $(A_r.4)$ satisfied for all $t \in]-a_1, a_2[$, we only need the assumptions $(A_r.3)$ and $(A_r.5)$. Then the proof is the same as the one of Louani (who assumes $a_1 = a_2 = \infty$) in [11].

In this case notice that the existence of a finite value for $\inf_{t \in [-a_1, a_2]} I_{r,\alpha}(t)$ ($\alpha \in \mathbb{R}$) is guaranteed since ϵ_1 is independent of X_1 and its support is $\mathbb{R}$.

2.2.2 Result and proof

We give the theorem. As in Section 2.1, we prove it in a more general case than in [11] (Remark 2.3 (iv)) with the assumptions above.

Theorem 2.2. *Let the conditions $(A_r.1) - (A_r.5)$ hold for any $\alpha \in \mathbb{R}$. Moreover assume that $\lim_{n \to \infty} n h_n = \infty$. Then $\hat{r}_n(z_0) - r(z_0)$ satisfies a LDP with speed $(n h_n)^{-1}$ and rate function $\Lambda_r^*(\alpha) = -\inf_{t \in \mathbb{R}} I_{r,\alpha}(t)$.*

Remark 2.4 As in Theorem 2.1, it is sufficient to prove the convergence for the following tail probabilities: For any $\alpha > 0$, we have

$$\lim_{n \to \infty} \frac{1}{n h_n} \log \mathbb{P}(\hat{r}_n(z_0) - r(z_0) > \alpha) = \inf_{t \in]0, a_2]} I_{r,\alpha}(t). \tag{9}$$

and

$$\lim_{n \to \infty} \frac{1}{n h_n} \log \mathbb{P}(\hat{r}_n(z_0) - r(z_0) < -\alpha) = \inf_{t \in]0, a_1]} I_{r,-\alpha}(-t). \tag{10}$$

If $f_X(z_0) = 0$ then $I_{r,\alpha}(t) = 0$ for all $\alpha \in \mathbb{R}$ and all $t \in [-a_1, a_2]$.

Proof. We quickly give some elements of proof for (9) as it is very similar to the one of Theorem 2.1. Let $\alpha > 0$ and assume that $f_X(z_0) > 0$.

For the upper bound, we use, as in [11], the assumptions $(A_r.1)$ and $(A_r.3) - (A_r.4)$.

For the lower bound, $t \to I'_{r,\alpha}(t)$ is non-decreasing and $I'_{r,\alpha}(0) = -\alpha$.

Then with assumption $(A_r.1)$, the existence of τ such that $I'_{r,\alpha}(\tau) = 0$, is not guaranteed for α large enough.

Therefore we use the same kind of proof as in [6, Theorem 2.2.3]. We are interested in the probability

$$\mathbb{P}(\hat{r}_n(z_0) - r(z_0) > \alpha) = \mathbb{P}\left(\frac{1}{n h_n} \sum_{j=1}^{n} W_{jn,\alpha} > 0\right).$$

For $t \in \mathbb{R}$, we have

$$
\begin{aligned}
\Lambda_{M,\alpha}(t) &= \liminf_{n \to \infty} h_n^{-1} \log \mathbb{E}\left\{\exp(t W_{1n,\alpha}) \mathbb{1}_{|W_{1n,\alpha}| \leq M}\right\} \\
&= -\int_{|y - r(z_0) - \alpha|K(x) > M} f(z_0, y) dx dy \\
&\quad + f_X(z_0) \int_{|y - r(z_0) - \alpha|K(x) \leq M} (\exp(t(y - r(z_0) - \alpha)K(x)) - 1) f(y|z_0) dx dy
\end{aligned}
$$

7

where we have used the assumptions $(A_r.1) - (A_r.3)$.

$\Lambda_{M,\alpha}$ is finite everywhere for all $M > 0$ and $\lim_{|t| \to \infty} \Lambda_{M,\alpha}(t) = \infty$ for M large enough. We conclude as in the proof of Theorem 2.1. $\quad\square$

3 Sharp Large Deviations for the kernel density estimator

3.1 Notations and assumptions

In this section, we are interested in the LSLDP satisfied by $\hat{f}_n(x)$ when $x \in \mathbb{R}$ is such that $f(x) > 0$. We will see in Section 3.4 a similar result when $f(x) = 0$.

We consider two cases depending on the support of the kernel: The first one concerns kernels with compact support and the second one kernels with unbounded support.

The proof of the LSLDP is essentially the same in the two cases. The main difficulty lies in the Edgeworth expansion (see Lemma 3.2). This expansion has been given by Hall in the first case [8, Theorem 5.5] and the proof is not so different in the second case. Of course the assumptions which have to be satisfied by the kernel K in this last case are stronger.

We use the same notations as in Section 2.2.1 and introduce a new function:

$$J(t) = \int_{\mathbb{R}} z(\exp(tK(z)) - 1)dz.$$

We consider new conditions. Indeed for a sharper development of the tail probability, we need the existence of the parameter τ used for the exponential change of probability.

Then we have to introduce an assumption on the definition domain of the function I (through φ_0) less general than $(A.1)$ to guarantee this existence. We also need more regularity for the density function f. Other conditions upon K and f are also given for each case. We first give the common assumptions for the two cases:

$(B.1)$ $\varphi_0(t) = \int_{\mathbb{R}} K(z) \exp(tK(z))dz$ is defined on the open interval $] - \infty, a[$, $a > 0$.

$(B.2)$ $\varphi_1(t) = \int_{\mathbb{R}} |z|K(z) \exp(tK(z))dz$ is defined on the open interval $] - \infty, b[$, $b \geq a > 0$.

$(B.3)$ f is continuously differentiable on $\mathbb{R}$, $\|f\|_\infty < \infty$ and $\|f'\|_\infty < \infty$.

Remark 3.1

(i) If $\lim_{z\to 0} K(z) < \infty$ then $a = b$ and if $\lim_{z\to 0} K(z) = \infty$ then $b > a$. For most classical kernels, we have $a = b = \infty$ but for example if $K(z) = -C_1 \log(z)1_{[0,1/2[} + C_2 \exp(-z)1_{[1/2,\infty[}$, where C_1 and C_2 are normalization constants, we get $a < b < \infty$.

(ii) As in Remark 2.1, $I \in C^\infty(]-\infty, a[)$ by using $(B.1)$ and $J \in C^\infty(]-\infty, b[)$ follows from $(B.2)$. We note I'' (respectively J'') the second derivative of I (respectively J).
Notice that $(B.2)$ implies $\int_{\mathbb{R}} |z|K(z)dz < \infty$ and reciprocally a kernel such that $\int_{\mathbb{R}} |z|K(z)dz = \infty$ cannot satisfy $(B.2)$.

3.1.1   Assumptions for kernels with compact support

We give the assumptions satisfied by K in the first case:

$(B.4)$ K has a compact support $[A, B]$.

$(B.5)$ $\mu_{A,B} \circ K^{-1} << d\lambda$ where $\mu_{A,B}$ is the uniform measure on $[A, B]$.

Remark 3.2   The assumptions $(B.4)$ and $(B.5)$ are needed to derive a version of Cramer's condition for the random variable $h_n K_{h_n}(x - X)$. This condition used by Hall [8, Lemma 5.6] to prove the Edgeworth expansion is the following: For each $k > 0$, there exists $C(k, x) > 0$ such that

$$\sup_{|t|>k} |\mathbb{E}\exp(ith_n K_{h_n}(x - X))| < 1 - C(k, x)h_n \tag{11}$$

for n large enough.
Indeed the characteristic function will be considered over a new distribution. There is no important difference between the proof under the original distribution and the one under the new distribution. Then we will not give this last one. The following assumption on K which appears in [8] is equivalent to $(B.5)$: There exists a partition of the interval $[A, B]$: $A = y_0 < y_1 < .. < B = y_n$ such that K' exists and on each segment $]y_{j-1}, y_j[$, K' is bounded and is either strictly positive or negative.
For example, the uniform kernel does not satisfy condition $(B.5)$ while the kernel $K(t) = -2^{-1} \log |t|1_{[-1,1]}(t)$ does.

3.1.2   Assumptions for kernels with unbounded support

We also give the assumptions satisfied by K and f in the second case:

$(B.6)$ K has an unbounded support.

$(B.7)$ For all $\tau \in]-\infty, a[$ and all $p \in \mathbb{N}^*$, $\int_{\mathbb{R}} K^{1/p}(z) \exp(\tau K(z)) dz < \infty$.

$(B.8)$ $\sup_{n \in \mathbb{N}^*} \sup_{t>0} | \int_{\mathbb{R}} \sin(tK(z)) f(x - h_n z) dz | < \infty$.

Remark 3.3 The assumptions $(B.7) - (B.8)$ are used to prove a version of Cramer's condition for $h_n K_{h_n}(x - X)$ when K has an unbounded support (see Lemma 3.3). This version is more general than the one for kernels with compact support (11).
These assumptions are quite restrictive. For example, the Cauchy kernel does not satisfy $(B.7)$.
We verify that $(B.8)$ holds for the exponential kernel $K(z) = 2^{-1} \exp(-|z|)$: For all $t > 0$, we have

$$\int_{\mathbb{R}} \sin(tK(z)) f(x - h_n z) dz = \int_0^\infty \sin(2^{-1} t \exp(-z)) f(x - h_n z) dz$$
$$+ \int_0^\infty \sin(2^{-1} t \exp(-z)) f(x + h_n z) dz$$

For the first integral (it is the same proof for the second), we make the change of variable $u = t \exp(-z)$. For all $t > 1$, we get

$$\left| \int_0^t \frac{\sin(2^{-1} u)}{u} f\left(x + h_n \log\left(\frac{u}{t}\right)\right) du \right| \leq \left| \int_1^t \frac{\sin(2^{-1} u)}{u} f\left(x + h_n \log\left(\frac{u}{t}\right)\right) du \right|$$
$$+ 2^{-1} \|f\|_\infty.$$

We use an integration by parts for the integral above and finally $(B.8)$ follows from condition $(B.3)$. Notice that this last one is needed to prove $(B.8)$.
We can also prove it for the Gaussian kernel with the same arguments and more generally for kernels of the following type: $K(z) = C \exp(-\varphi(z))$ where C is the normalization constant and φ a convex function such that $\varphi(z)/|z|^p \to \infty$ as $|z| \to \infty$ for all $p < 1$.
Of course these kernels also satisfy $(B.7)$

3.2 Main Theorem

Under the preceding notations and assumptions, we have the following theorem,

Theorem 3.1. *Let the conditions $(B.1) - (B.3)$ hold.*
Assume conditions $(B.4) - (B.5)$ in the first case or $(B.6) - (B.8)$ in the second case. Moreover suppose that $\lim_{n\to\infty} nh_n^2 = c \geq 0$ where c is a constant.
Then for any $x \in \mathbb{R}$ such that $f(x) > 0$, any $\alpha > 0$ and n large enough,

$$\mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) = \frac{\exp[-nh_n \Lambda^*(\alpha) + cH(\tau)]}{\tau (2\pi nh_n f(x) I''(\tau))^{1/2}} (1 + o(1)) \qquad (12)$$

10

where $\tau \in]0,a[$ and $H(\tau) = -(f^2(x)I^2(\tau)/2 + f'(x)J(\tau))$.
Similarly, for any α and x such that $0 < \alpha < f(x)$ and n large enough, we have

$$\mathbb{P}(\hat{f}_n(x) - f(x) < -\alpha) = \frac{\exp[-nh_n\Lambda^*(-\alpha) + cH(-\tau)]}{\tau(2\pi nh_n f(x)I''(-\tau))^{1/2}}(1 + o(1)) \qquad (13)$$

where $\tau > 0$.

Remark 3.4 It is possible to obtain a LSLDP of higher order (see [3]) but the calculation is more complicated and we will not make it here.

Proof. In order to prove (12) (the same arguments are used to prove (13)), we follow the same approach as in [1] and more recently in [3].
Let $x \in \mathbb{R}$ such that $f(x) > 0$ and let $\alpha > 0$. We consider the normalized cumulant generating function of $\hat{f}_n(x) - f(x)$, for $t \in]-\infty, a[$,

$$\begin{aligned}
\Lambda_n(t) &= \frac{1}{nh_n}\log\mathbb{E}\{\exp(tnh_n(\hat{f}_n(x) - f(x)))\} \\
&= \frac{1}{nh_n}\log\mathbb{E}\{\exp(tnh_n\hat{f}_n(x))\} - tf(x).
\end{aligned}$$

We want a sharper development of this function than in the proof of the large deviation theorem (Theorem 2.1). Then for the function $\theta : \mathbb{R} \to [0,1]$ and $n \in \mathbb{N}^*$, we set $L_{n,\theta}(t) = \int_{\mathbb{R}} z(\exp\{tK(z)\} - 1)f'(x - \theta(z)h_n z)dz$.
This function is defined on $]-\infty, b[$ by using the assumptions $(B.2)$ and $(B.3)$. Moreover the dominate convergence theorem allows us to write $L_{n,\theta}$ as:

$$L_{n,\theta}(t) = f'(x)J(t) + o(1). \qquad (14)$$

Notice that $L_{n,\theta} \in C^\infty(]-\infty, b[)$ (Remark 3.1 (ii)) and for $k \geq 1$, its derivatives can be written :

$$L_{n,\theta}^{(k)}(t) = \int_{\mathbb{R}} zK^k(z)\exp(tK(z))f'(x - \theta(z)h_n z)dz = f'(x)J^{(k)}(t) + o(1). \qquad (15)$$

Noting ψ_n the Laplace transform of $h_n K_{h_n}(x - X)$, we have, for $t \in]-\infty, a[$,

$$\begin{aligned}
\psi_n(t) &= \mathbb{E}\{\exp(th_n K_{h_n}(x - X))\} \\
&= 1 + h_n f(x)I(t) - h_n^2 L_{n,\theta}(t). \qquad (16)
\end{aligned}$$

As $nh_n^2 = c + o(1)$ and using the expression (14) of $L_{n,\theta}$, we obtain

$$\Lambda_n(t) = f(x)I(t) - tf(x) + \frac{c}{nh_n}H(t) + o\left(\frac{1}{nh_n}\right) \qquad (17)$$

where $H(t) = -(f^2(x)I^2(t)/2 + f'(x)J(t))$.
I is a convex and differentiable function on $]-\infty, a[$. Moreover $\lim_{t\to-\infty} I'(t) = 0$ and $\lim_{t\to a_-} I'(t) = \infty$. Then there exists a parameter $\tau \in]0,a[$ such that

$$\alpha + f(x) = f(x)I'(\tau) \qquad (18)$$

11

and
$$\Lambda^*(\alpha) = \tau(\alpha + f(x)) - f(x)I(\tau). \tag{19}$$

This parameter τ is used for the exponential change of probability which allows us to prove (12).

Consider the exponential change of probability:

$$\frac{d\mathbb{P}_n}{d\mathbb{P}_{n,0}} = \exp(\tau h_n K_{h_n}(x - X) - \log \psi_n(\tau))$$

where $\mathbb{P}_n$ (respectively $\mathbb{P}_{n,0}$) is the new (respectively original) distribution of $h_n K_{h_n}(x - X)$.

We note $\mathbb{E}_n$ and Var_n respectively the expectation and the variance over the distribution $\mathbb{P}_n$. These notations will be used in the rest of the proof.

Now, we can express the tail probability

$$
\begin{aligned}
\mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) &= \mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{n} K_{h_n}(x - X_j) > f(x) + \alpha\right) \\
&= \int_{f(x)+\alpha}^{\infty} \exp[-\tau n h_n u + n h_n \Lambda_n(\tau) + \tau n h_n f(x)] dQ_n(u)
\end{aligned}
$$

where Q_n is the distribution of $\hat{f}_n(x) = n^{-1}\sum_{j=1}^{n} K_{h_n}(x - X_j)$ over $\mathbb{P}_n$.

The expressions (17) and (19) yield

$$
\begin{aligned}
\mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) &= \exp[-n h_n \Lambda^*(\alpha) + cH(\tau) + o(1)] \\
&\quad \times \int_{f(x)+\alpha}^{\infty} \exp[-\tau n h_n(u - (f(x) + \alpha))] dQ_n(u).
\end{aligned}
$$

In order to evaluate the integral, we consider, using (18),

$$U_n = \frac{(n h_n)^{1/2}(\hat{f}_n(x) - f(x)I'(\tau))}{(f(x)I''(\tau))^{1/2}} = \frac{(n h_n)^{1/2}(\hat{f}_n(x) - (f(x) + \alpha))}{(f(x)I''(\tau))^{1/2}}.$$

The following lemma enables us to show the asymptotic normality of U_n under the distribution $\mathbb{P}_n$.

Lemma 3.1. *Let $(B.1) - (B.3)$ hold and assume that $\lim_{n \to \infty} n h_n^2 = c \geq 0$. Then the distribution of U_n under $\mathbb{P}_n$ converges to the standard Gaussian distribution.*

Proof. This result was found by Parzen [14] for $\hat{f}_n$ considered over the original distribution. Here $\hat{f}_n$ is considered over $\mathbb{P}_n$.

We first give some expressions that will also be useful further:

From classical results on exponential models, (14), (15) and (16), we have

$$
\begin{aligned}
\mathbb{E}_n \hat{f}_n &= \mathbb{E}_n\{K_{h_n}(x - X)\} \\
&= h_n^{-1}\frac{\psi_n'(\tau)}{\psi_n(\tau)} \\
&= f(x)I'(\tau) - h_n(f'(x)J'(\tau) + f^2(x)I(\tau)I'(\tau)) + o(h_n) \tag{20}
\end{aligned}
$$

and

$$\begin{aligned}
Var_n \hat{f}_n &= n^{-1} Var_n \{K_{h_n}(x - X)\} \\
&= (nh_n^2)^{-1} \frac{\psi_n''(\tau)}{\psi_n(\tau)} - n^{-1}(\mathbb{E}_n \hat{f}_n)^2 \\
&= (nh_n)^{-1} \left(f(x) I''(\tau) + o(1) \right).
\end{aligned} \qquad (21)$$

We use the Lyapounov central limit theorem for triangular arrays (see [4]) to prove the lemma.

As $\lim_{n \to \infty} nh_n^2 = c$, the convergence in distribution of U_n will be assured if the condition of Lyapounov is satisfied.

Noting $W_{kn} = n^{-1}(K_{h_n}(x - X_k) - \mathbb{E}_n K_{h_n}(x - X_k))$, for $k = 1, .., n$, the condition of Lyapounov is: There exists $\delta > 0$ such that

$$\lim_{n \to \infty} \frac{\sum_{k=1}^n \mathbb{E}_n |W_{kn}|^{2+\delta}}{(\sum_{k=1}^n \mathbb{E}_n W_{kn}^2)^{1+\delta/2}} = 0.$$

This is equivalent to

$$\lim_{n \to \infty} \frac{\mathbb{E}_n |K_{h_n}(x - X)|^{2+\delta}}{n^{\delta/2}(Var_n K_{h_n}(x - X))^{1+\delta/2}} = 0.$$

Then, we can easily verify this condition and prove the lemma by using (15), (16), (20), (21), the assumptions $(B_1) - (B_3)$ and $\lim_{n \to \infty} nh_n^2 = c$. $\qquad \square$

Hence we can write

$$\begin{aligned}
\mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) &= \exp[-nh_n \Lambda^*(\alpha) + cH(\tau) + o(1)] \\
&\times \int_0^\infty \exp[-\tau(nh_n f(x) I''(\tau))^{1/2} u] f_{U_n}(u) du, \quad (22)
\end{aligned}$$

where f_{U_n} is the density function of U_n.

For every integer $k > 1$, we note

$$\mu_{k,n} = h_n^{-1} \mathbb{E}_n |K((x - X)/h_n) - \mathbb{E}_n K((x - X)/h_n)|^k \qquad (23)$$

and we consider $U_n^* = (nh_n)^{1/2}(\hat{f}_n(x) - \mathbb{E}_n \hat{f}_n(x))/\mu_{2,n}^{1/2}$.

Then, with F_{U_n} (respectively $F_{U_n^*}$) standing for the distribution function of U_n (respectively U_n^*), we have, clearly, $F_{U_n}(u) = F_{U_n^*}((u - B_n)/A_n)$ where

$$A_n = \left(\frac{\mu_{2,n}}{f(x) I''(\tau)} \right)^{1/2} \text{ and } B_n = \frac{(nh_n)^{1/2}(\mathbb{E}_n \hat{f}_n(x) - f(x) I'(\tau))}{(f(x) I''(\tau))^{1/2}} . \qquad (24)$$

13

One can easily verify by using (20), (21) and (23) that $\lim_{n\to\infty} A_n = 1$ and $\lim_{n\to\infty} B_n = 0$.
Set

$$
\begin{aligned}
\alpha_n &= \tau(nh_n f(x)I''(\tau))^{1/2} A_n \\
&= \tau(nh_n f(x)I''(\tau))^{1/2}(1 + o(1))
\end{aligned}
\tag{25}
$$

and

$$
\begin{aligned}
\beta_n = \tau(nh_n f(x)I''(\tau))^{1/2} B_n &= \tau n h_n (\mathbb{E}_n \hat{f}_n(x) - f(x)I'(\tau)) \\
&= -\tau c(f'(x)J'(\tau) + f^2(x)I(\tau)I'(\tau)) + o(1)
\end{aligned}
\tag{26}
$$

where we have used (20) and the assumption $\lim_{n\to\infty} nh_n^2 = c$.
Therefore, we write (22) as

$$
\begin{aligned}
\mathbb{P}(\hat{f}_n(x) - f(x) > \alpha) &= \exp[-nh_n \Lambda^*(\alpha) + cH(\tau) + o(1)] \\
&\quad \times \ \exp(-\beta_n) \int_{-\frac{B_n}{A_n}}^{\infty} \exp(-\alpha_n u) f_{U_n^*}(u) du \\
&= \frac{\exp[-nh_n \Lambda^*(\alpha) + cH(\tau) + o(1)]}{\alpha_n (2\pi)^{1/2}} \times I_n \exp(-\beta_n)
\end{aligned}
$$

where $I_n = \alpha_n (2\pi)^{1/2} \int_{-\beta_n}^{\infty} \exp(-u)(F_{U_n^*}(u/\alpha_n) - F_{U_n^*}(-B_n/A_n)) du$.
We have used an integration by parts, two changes of variable (see the proof of the Bahadur-Rao Theorem in [6]) and noticed that the lower bound of the integral is $-\alpha_n B_n/A_n = -\beta_n$.
Now it remains to prove that $\lim_{n\to\infty} \exp(-\beta_n) I_n = 1$.
The following lemma yields an Edgeworth expansion. Using this last one, the end of the proof will be nearly the same as the one of the Bahadur-Rao Theorem in [6].

Lemma 3.2. *Under the preceding notations (in particular (23)) and the assumptions of Theorem 3.1,*

$$
\lim_{n\to\infty} (nh_n)^{1/2} \sup_{x\in\mathbb{R}} \left| F_{U_n^*}(x) - \Phi(x) - \frac{\mu_{3,n}}{6(nh_n)^{1/2}(\mu_{2,n})^{3/2}}(1 - x^2)\phi(x) \right| = 0, \tag{27}
$$

where $\Phi(x)$ is the standard normal distribution and $\phi(x)$ the standard normal density.

Proof. This result has been proved by Hall [8, Theorem 5.5] in the case of kernels with compact support (and under the original distribution but it is nearly the same proof under the new one). For kernels with unbounded support, we can also apply Hall's proof by using the assumptions $(B.7)$ and $(B.8)$. In fact the main difficulty is to obtain a version of Cramer's condition for $h_n K_{h_n}(x - X)$ which

guarantees the uniformity of the expansion.

More precisely, to prove his result, Hall uses a lemma [8, Lemma 5.7]. This lemma yields an uniform expansion for the density function of $U_n^* + n^{-d}Z$ where Z is a $\mathcal{N}(0,1)$ random variable independent of X_j, $j = 1, .., n$ and d a positive constant. For the lemma and the rest of the proof of (27), the constant $d > 0$ has to be large enough.

In the proof of his lemma, Hall needs a version of Cramer's condition for kernels satisfying the assumptions $(B.4)$ and $(B.5)$. Indeed this condition (11) yields the following result used in his proof: For all $a > 0$,

$$\sup_{|t|>k_0(nh_n)^{1/2}} \left| v_n \left(t/(nh_n)^{1/2} \right) \right|^n = O(n^{-a})$$

for n large enough, where $k_0 > 0$ and $v_n(t) = \mathbb{E} \exp(ith_n K_{h_n}(x - X))$.

Hall's proof remains valid for kernels with unbounded support if we show a more general result with d chosen large enough: For all $a > 0$,

$$\sup_{k_0(nh_n)^{1/2}<|t|<n^{2d}} \left| v_n^* \left(t/(nh_n)^{1/2} \right) \right|^n = \sup_{k_0<|t|<n^{2d-1/2}h_n^{-1/2}} |v_n^*(t)|^n = O(n^{-a}) \quad (28)$$

for n large enough, where $v_n^*(t) = \mathbb{E}_n \exp(ith_n K_{h_n}(x - X))$.

For this, we use the following version of Cramer's condition:

Lemma 3.3. *Let the conditions* $(B_1) - (B.3)$ *and* $(B.6) - (B.8)$ *hold. Assume that* $\lim_{n\to\infty} nh_n^2 = c \geq 0$. *Then for each* $k > 0$ *and each* $q > 0$,

$$\sup_{k<|t|<n^q} |\mathbb{E}_n \exp(ith_n K_{h_n}(x - X))| \leq 1 - C(k,x)h_n \quad (29)$$

for n large enough, where $C(k,x) > 0$.

Remark 3.5 We give the proof for the characteristic function over the original distribution for reasons of simplicity in the notations. Indeed there are very few differences with the proof of (29). Condition $(B.7)$ yields the result $\int_{\mathbb{R}} K^{1/p}(z)dz < \infty$ which is used below.

Proof. We have, under the original distribution, for $t \in \mathbb{R}$,

$$|v_n(t)|^2 - 1 = -4h_n A_n(t) \times B_n(t) + h_n^2 C_n^2(t)$$

where $A_n(t) = \int_{\mathbb{R}} \sin^2(tK(z)/2)f(x - h_n z)dz$, $C_n(t) = \int_{\mathbb{R}} \sin(tK(z))f(x - h_n z)dz$ and $B_n(t) = h_n \int_{\mathbb{R}} \cos^2(tK(z)/2)f(x - h_n z)dz$.

Let $k > 0$ and $q > 0$. Then we have to prove

$$\sup_{n\in\mathbb{N}^*} \sup_{t>0} \left| \int_{\mathbb{R}} \sin(tK(z))f(x - h_n z)dz \right| < \infty \quad (30)$$

15

and

$$\inf_{k<|t|<n^q} \left(\int_{\mathbb{R}} \sin^2 \left(\frac{tK(z)}{2} \right) f(x - h_n z) dz \right)$$

$$\times \inf_{k<|t|<n^q} \left(h_n \int_{\mathbb{R}} \cos^2 \left(\frac{tK(z)}{2} \right) f(x - h_n z) dz \right) > \varepsilon(k, x) \qquad (31)$$

for n large enough, where $\varepsilon(k, x) > 0$.

In fact (30) is the assumption $(B.8)$.

For the first term in (31), as $f(x) > 0$ and f is continuous on $\mathbb{R}$, there exists $A_x > 0$ and $B_x > 0$ such that $f > B_x$ on $]x - A_x, x + A_x[$. Then for all $t \in]k, n^q[$ and n large enough, we have

$$
\begin{aligned}
\int_{\mathbb{R}} \sin^2 \left(\frac{tK(z)}{2} \right) f(x - h_n z) dz &\geq B_x \int_{-\frac{A_x}{h_n}}^{\frac{A_x}{h_n}} \sin^2 \left(\frac{tK(z)}{2} \right) dz \\
&\geq 2^{-1}(\beta - \alpha) B_x \left(1 - \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} \cos(tK(z)) dz \right) \\
&\geq 2^{-1}(\beta - \alpha) B_x \left(1 - |\mathbb{E} \exp(itK(U))| \right)
\end{aligned}
$$

where U follows an uniform distribution on the interval $]\alpha, \beta[$ chosen such that K is finite and strictly positive there.

As $|\mathbb{E} \exp(itK(U))| = 1$ if and only if $t = 0$ (the distribution function of $K(U)$ being non-lattice), we have $\int_{\mathbb{R}} \sin^2(tK(z)/2) f(x - h_n z) dz > 2\varepsilon(k, x)$ for all $t \in]k, n^q[$ and n large enough, with $\varepsilon(k, x) > 0$.

For the second term in (31), notice that $\cos^2(u) \geq 1 - 2p|u|^{1/p}$ for all $u \in \mathbb{R}$ and all $p > 1$. Then for all $t \in]k, n^q[$,

$$
\begin{aligned}
h_n \int_{\mathbb{R}} \cos^2 \left(\frac{tK(z)}{2} \right) f(x - h_n z) dz &\geq 1 - p2^{1-1/p} h_n |t|^{1/p} \int_{\mathbb{R}} K^{1/p}(z) f(x - h_n z) dz \\
&\geq 1 - p2^{1-1/p} \left(n h_n^{\frac{p}{q}} \right)^{\frac{q}{p}} \int_{\mathbb{R}} K^{1/p}(z) f(x - h_n z) dz.
\end{aligned}
$$

If $2q < p$, then $\lim_{n \to \infty} n h_n^{\frac{p}{q}} = 0$. Choosing p large enough, the assumptions $(B.7)$ and $(B.3)$ yield $h_n \int_{\mathbb{R}} \cos^2(tK(z)/2) f(x - h_n z) dz \geq 1/2$ for all $t \in]k, n^q[$ and n large enough. This completes the proof of (31).

In conclusion, for each $k > 0$ and each $q > 0$, we have

$$\sup_{k<|t|<n^q} |v_n(t)|^2 \leq 1 - 2C(k, x) h_n \Rightarrow \sup_{k<|t|<n^q} |v_n(t)| \leq 1 - C(k, x) h_n$$

for n large enough, where $C(k, x) > 0$. $\qquad \square$

We apply this lemma with $n^{2d-1/2}h_n^{-1/2}$ (d chosen large enough) in place of n^q. We get for all $a > 0$ and n large enough,

$$n^a \sup_{k_0 < |t| < n^{2d-1/2}h_n^{-1/2}} |v_n^*(t)|^n \leq n^a \exp(-C(k_0, x)nh_n).$$

The assumption $\lim_{n \to \infty} nh_n^2 = c$ which implies $\lim_{n \to \infty} nh_n / \log n = \infty$, yields formula (28). Therefore we can apply Hall's proof to obtain the Edgeworth expansion (27). $\qquad\square$

Now as in the proof of the Bahadur-Rao Theorem in [6], using the equality $\alpha_n B_n / A_n = \beta_n$, we consider

$$
\begin{aligned}
J_n &= (2\pi)^{1/2} \int_{-\beta_n}^{\infty} \alpha_n \exp(-u) \left(\Phi\left(\frac{u}{\alpha_n}\right) + \frac{\mu_{3,n}}{6(nh_n)^{1/2}(\mu_{2,n})^{3/2}} \left[1 - \left(\frac{u}{\alpha_n}\right)^2\right] \right. \\
&\quad \cdot \phi\left(\frac{u}{\alpha_n}\right) - \Phi\left(\frac{-B_n}{A_n}\right) - \frac{\mu_{3,n}}{6(nh_n)^{1/2}(\mu_{2,n})^{3/2}} \left[1 - \left(\frac{-B_n}{A_n}\right)^2\right] . \phi\left(\frac{-B_n}{A_n}\right) \right) du
\end{aligned}
$$

with $\phi(t) = \Phi'(t) = (2\pi)^{-1/2} \exp(-t^2/2)$.
Clearly, (27) yields $\lim_{n \to \infty} |I_n - J_n| = 0$. After a Taylor expansion of $\Phi(u/\alpha_n)$ and using (24), (25), (26), we obtain

$$
\begin{aligned}
\lim_{n \to \infty} J_n &= \lim_{n \to \infty} (2\pi)^{1/2} \int_{-\beta_n}^{\infty} \alpha_n \exp(-u) \left(\Phi\left(\frac{u}{\alpha_n}\right) - \Phi\left(-\frac{B_n}{A_n}\right) \right) du \\
&= \lim_{n \to \infty} \exp(\beta_n),
\end{aligned}
$$

completing the proof of the theorem. $\qquad\square$

3.3 The case where f(x) = 0

In the case where $f(x) = 0$, $\hat{f}_n(x)$ also satisfies a LSLDP but with a change of speed. Indeed, we are interested in the probability $\mathbb{P}(\hat{f}_n(x) > \alpha h_n^2)$ where $\alpha > 0$. We use notations and assumptions very similar to the case where $f(x) > 0$ and one can make the same kind of remarks.
We consider the functions

$$J_2(t) = \int_{\mathbb{R}} z^2 (\exp(tK(z)) - 1) dz$$

and

$$J_3(t) = \int_{\mathbb{R}} z^3 (\exp(tK(z)) - 1) dz.$$

Let K have an unbounded support. Then we suppose that the following conditions are satisfied:

(Bo_1) $f(x) = f'(x) = 0$ and $f''(x) > 0$.

(Bo_2) $\varphi_2(t) = \int_\mathbb{R} z^2 K(z) \exp(tK(z))dz$ and $\varphi_3(t) = \int_\mathbb{R} |z|^3 K(z) \exp(tK(z))dz$ are defined respectively on $]-\infty, a[$ and $]-\infty, b[$ with $b \geq a > 0$.

(Bo_3) $\lim_{n\to\infty} nh_n^4 = c \geq 0$.

(Bo_4) f is three times continuously differentiable on $\mathbb{R}$, $\|f''\|_\infty < \infty$ and $\|f^{(3)}\|_\infty < \infty$.

(Bo_5) For all $\tau \in]-\infty, a[$ and all $p \in \mathbb{N}^*$, $\int_\mathbb{R} z^2 K^{1/p}(z) \exp(\tau K(z))dz < \infty$.

(Bo_6) $\sup_{n\in\mathbb{N}^*} \sup_{t>0} \left| \int_\mathbb{R} z^2 \sin(tK(z)) f''(x - \theta(z)h_n z)dz \right| < \infty$ where $\theta(z) \in]0, 1[$.

(Bo_7) α and x are such that $\alpha > J_2'(0)f''(x)/2$.

The last condition is used to have a parameter $\tau > 0$. We will not study here the case $0 < \alpha \leq J_2'(0)f''(x)/2$ which seems to be more complicated.

Notice that if K has a compact support, we need assumption $(B.5)$ from section 3.2 in place of $(Bo_5) - (Bo_6)$. We give only the result for the right tail probability (it is the same for the left one).

Theorem 3.2. *Under the preceding notations and assumptions, we have, for n large enough,*

$$\mathbb{P}(\hat{f}_n(x) > \alpha h_n^2) = \frac{\exp[-nh_n^3 \Lambda_2^*(\alpha) - cf^{(3)}(x)J_3(\tau)/6]}{\tau(\pi nh_n^3 f''(x)J_2''(\tau))^{1/2}}(1 + o(1))$$

with $\tau \in]0, a[$ and $\Lambda_2^*(\alpha) = \tau\alpha - f''(x)J_2(\tau)/2$.

Remark 3.6 We have a change of speed in this particular case, $(nh_n^3)^{-1}$ in place of $(nh_n)^{-1}$. The expansion depends on the second and third derivatives of f as the Taylor development has to be sharper.

Proof. We give a few elements of proof in this particular case: Considering

$$W_n(x) = \frac{1}{nh_n^3} \sum_{j=1}^n K\left(\frac{x - X_j}{h_n}\right),$$

we can write $\mathbb{P}(\hat{f}_n(x) > \alpha h_n^2) = \mathbb{P}(W_n(x) > \alpha)$.

Then with the assumptions $(Bo_1) - (Bo_4)$, we have, for $t \in]-\infty, a[$,
$\Lambda_{2n}(t) = (nh_n^3)^{-1} \log \mathbb{E} \exp(tnh_n^3 W_n(x)) = f''(x)J_2(t)/2 - h_n f^{(3)}(x)J_3(t)/6 + o(h_n)$
and $\Lambda_2(t) = \lim_{n\to\infty} \Lambda_{2n}(t) = f''(x)J_2(t)/2$.

Using assumption (Bo_7), the rate function evaluated at α is

$$\begin{aligned}
\Lambda_2^*(\alpha) &= \sup_{t\in]0,a[} \{t\alpha - (f''(x)J_2(t))/2\} \\
&= \tau\alpha - f''(x)J_2(\tau)/2,
\end{aligned}$$

18

with $\tau \in]0, a[$ such that $J_2'(\tau)f''(x) = 2\alpha$.

For the rest of the proof, we use the same arguments as in the case where $f(x) > 0$.

$\square$

Remark 3.7 More generally the speed of the LSLDP increases with the number of successive derivatives equal to zero at x: Let $k > 1$ be an integer. We suppose $f(x) = .. = f^{(2k-1)}(x) = 0$, $f^{(2k)}(x) > 0$, $\lim_{n\to\infty} nh_n^{2k+2} = c \geq 0$, and others assumptions on f and K which can be easily deduced from the case $k = 1$. Then we obtain the following result for any $\alpha > J_{2k}'(0)f^{(2k)}(x)/(2k)!$ and n large enough:

$$\mathbb{P}(\hat{f}_n(x) > \alpha h_n^{2k}) = \frac{\exp\left[-nh_n^{2k+1}\Lambda_{2k}^*(\alpha) - cf^{(2k+1)}(x)\frac{J_{2k+1}(\tau)}{(2k+1)!}\right]}{\tau(\frac{2\pi}{(2k)!}nh_n^{2k+1}f^{(2k)}(x)J_{2k}''(\tau))^{1/2}}(1 + o(1))$$

where $\Lambda_{2k}^*(\alpha) = \tau\alpha - f^{(2k)}(x)J_{2k}(\tau)/(2k)!$ and $J_{2k}(t) = \int_{\mathbb{R}} z^{2k}(\exp(tK(z)) - 1)dz$.

4 Sharp Large Deviations for the Nadaraya-Watson estimator

We use the same notations as in Section 2.2 and consider assumptions for kernels with unbounded support (for kernels with compact support, we only need the assumptions $(B_r.1) - (B_r.3)$ below). They are quite similar to those regarding the kernel density estimator and one can make the same kind of remarks. Let z_0 be a fixed point in $\mathbb{R}$ such that $f_X(z_0) > 0$. Then we introduce the new function

$$J_{r,\alpha}(t) = \int_{\mathbb{R}^2} x[\exp(t(y - r(z_0) - \alpha)K(x)) - 1]f_1'(z_0, y)dxdy. \tag{32}$$

We assume that the following assumptions are fulfilled for any $\alpha \in \mathbb{R}$:

$(B_r.1)$ f is continuously differentiable with respect to the first variable on $\mathbb{R}$ and f_X is continuous on $\mathbb{R}$.

$(B_r.2)$ $\varphi_{r,\alpha}(t) = \int_{\mathbb{R}^2} |y - r(z_0) - \alpha|K(x)\exp[t(y - r(z_0) - \alpha)K(x)]f(z_0, y)dxdy$ is defined on the open interval $] - a_1, a_2[$, with $a_1 \in]0, \infty]$ and $a_2 \in]0, \infty]$.

$(B_r.3)$ For all $m \in \mathbb{N}^*$ and all $t \in] - b_1, b_2[\supset] - a_1, a_2[$,

$$\sup_{u\in\mathbb{R}} \int_{\mathbb{R}^2} |x|(|y - r(z_0) - \alpha|K(x))^m \exp[t(y - r(z_0) - \alpha)K(x)]|f_1'(u, y)|dxdy < \infty.$$

(B_r.4) For all $p \in \mathbb{N}^*$ and all $t \in]-a_1, a_2[$,

$$\sup_{u \in \mathbb{R}} \int_{\mathbb{R}^2} (|y - r(z_0) - \alpha|K(x))^{1/p} \exp[t(y - r(z_0) - \alpha)K(x)]f(u, y)dxdy < \infty.$$

(B_r.5) $\sup_{n \in \mathbb{N}^*} \sup_{t>0} \left| \int_{\mathbb{R}^2} \sin(t(y - r(z_0) - \alpha)K(x))f(z_0 - h_n x, y)dxdy \right| < \infty.$

Remark Assumption (B_r.3) assures that $(Y - r(z_0) - \alpha)K((x - X)/h_n)$ has finite moments (under the new distribution obtained after an exponential change of probability) of all orders.

We have the following theorem, very similar to Theorem 3.1.

Theorem 4.1. *Let the conditions* $(B_r.1) - (B_r.5)$ *hold for any* $\alpha \in \mathbb{R}$. *Moreover suppose that* $\lim_{n \to \infty} nh_n^2 = c \geq 0$ *where* c *is a constant. Then, for any* $\alpha > 0$ *and* n *large enough,*

$$\mathbb{P}(\hat{r}_n(z_0) - r(z_0) > \alpha) = \frac{\exp[nh_n I_{r,\alpha}(\tau) + cH_{r,\alpha}(\tau)]}{\tau(2\pi nh_n I_{r,\alpha}''(\tau))^{1/2}}(1 + o(1)) \qquad (33)$$

where $\tau \in]0, a_2[$ *and* $H_{r,\alpha}(\tau) = -(I_{r,\alpha}^2(\tau)/2 + J_{r,\alpha}(\tau))$.
Similarly, if K *also satisfies condition* $(A_r.5)$, *for any* $\alpha > 0$ *and* n *large enough, we get*

$$\mathbb{P}(\hat{r}_n(z_0) - r(z_0) < -\alpha) = \frac{\exp[nh_n I_{r,-\alpha}(-\tau) + cH_{r,-\alpha}(-\tau)]}{\tau(2\pi nh_n I_{r,-\alpha}''(-\tau))^{1/2}}(1 + o(1)) \qquad (34)$$

where $\tau \in]0, a_1[$.

Remark If z_0 is such that $f_X(z_0) = 0$ then we have the same kind of results as in Section 3.3.

Proof. The proof of (33) is very similar to the one of Theorem 3.1, so we omit the details. Let $\alpha > 0$.
Using $(B_r.1) - (B_r.3)$ and $\lim_{n \to \infty} nh_n^2 = c$, we have, for $t \in]-a_1, a_2[$,

$$\begin{aligned}
\Lambda_{n,\alpha}(t) &= \frac{1}{nh_n} \log \mathbb{E} \exp \left(t \sum_{j=1}^n W_{jn,\alpha} \right) \\
&= I_{r,\alpha}(t) + \frac{c}{nh_n} H_{r,\alpha}(t) + o\left(\frac{1}{nh_n}\right)
\end{aligned}$$

where $H_{r,\alpha}(t) = -(I_{r,\alpha}^2(t)/2 + J_{r,\alpha}(t))$.
Then there exists a parameter $\tau \in]0, a_2[$ such that $I_{r,\alpha}'(\tau) = 0$ and the rate function associated with the LDP is $\inf_{t \in]0,a_2[} -I_{r,\alpha}(t) = -I_{r,\alpha}(\tau)$.
As in the proof of Theorem 3.1, we consider the new distribution $\mathbb{P}_{n,\alpha}$ of $W_{1n,\alpha}$

obtained after an exponential change of probability using τ. We express the tail probability,

$$
\begin{aligned}
\mathbb{P}(\hat{r}_n(z_0) - r(z_0) > \alpha) &= \exp[nh_n I_{r,\alpha}(\tau) + cH_{r,\alpha}(\tau) + o(1)] \\
&\quad \times \int_0^\infty \exp(-\tau nh_n u) dQ_{n,\alpha}(u)
\end{aligned}
$$

where $Q_{n,\alpha}$ is the distribution of $V_{n,\alpha} = (nh_n)^{-1} \sum_{j=1}^n W_{jn,\alpha}$ over $\mathbb{P}_{n,\alpha}$.

We set $U_{n,\alpha} = (nh_n)^{1/2} V_{n,\alpha}/(I''_{r,\alpha}(\tau))^{1/2}$ and as in lemma 3.1, the distribution of $U_{n,\alpha}$ under $\mathbb{P}_{n,\alpha}$ converges to the standard Gaussian distribution. For the rest of the proof, we follow the same approach as in the case of the kernel density estimator. Using $(B_r.1) - (B_r.5)$, Hall's proof (in [8]) is applied to $W_{1n,\alpha}$ in place of $h_n K_{h_n}(x - X)$. Then we obtain an Edgeworth expansion of the same kind of (27) and we get the result (33).

$\square$

Acknowledgement

References

[1] R. Bahadur and R. Rao (1960) On deviations of the sample mean. *Ann. Math. Statist.*, **31**, 1015-1027.

[2] O.E. Barndoff-Nielsen and D.R. Cox (1989) *Asymptotic techniques for uses in statistics*. Chapman and Hall (Londres).

[3] B. Bercu, F. Gamboa and M. Lavielle (2000) Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM:P&S*, **3**, 1-29.

[4] P. Billingsley (1968) *Convergence of probability measures*. New York :Wiley.

[5] H. Cramer (1970) *Random variables and probability distributions*. Cambridge University Press.

[6] A. Dembo and O. Zeitouni (1998) *Large deviations techniques and applications*. Springer, New York. Second edition.

[7] J.D. Deuschel and D.W. Stroock (1989) *Large deviations*. Boston: Academic Press.

[8] P. Hall (1992) *The bootstrap and Edgeworth expansion*. Springer-Verlag, New York.

[9] J.L. Jensen (1995) *Saddlepoint approximations*. Oxford University Press.

[10] D. Louani (1998) Large deviations limit theorems for the kernel density estimator. *Scand. J. of Statist.*, **25**, 243-253.

[11] D. Louani (1999) Some large deviations limit theorems in conditional non-parametric statistics. *Statistics*, **33**, 171-196.

[12] L. Menneteau (2000) *Grandes déviations et lois du logarithme itéré pour les processus de Kiefer et les processus empiriques locaux*. Ph.D. Thesis, Paris VI.

[13] E.A. Nadaraya (1964) On estimating regression. *Theor. Prob. Appl.*, **9**, 141-142.

[14] E. Parzen (1962) On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**, 1065-1076.

[15] G.S. Watson (1964) Smooth regression analysis. *Sankhya Ser.A*, **26**, 359-372.