



Uniform L1-large deviations for the delta-sequence method density estimator

Noureddine Berrahou, Djamal Louani

► To cite this version:

Noureddine Berrahou, Djamal Louani. Uniform L1-large deviations for the delta-sequence method density estimator. Annales de l'ISUP, 2005, L (3), pp.9-25. hal-03637625

HAL Id: hal-03637625

<https://hal.science/hal-03637625>

Submitted on 11 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.

UNIFORM L_1 -LARGE DEVIATIONS FOR THE DELTA-SEQUENCE METHOD DENSITY ESTIMATOR

BY NOUREDDINE BERRAHOU AND DJAMAL LOUANI

L.S.T.A., Université de Paris 6

In this paper we obtain uniform L_1 -distance large deviations results for the delta-sequence method density estimator. More precisely, we consider uniformity over classes of density functions fulfilling some regularity conditions. It results from our study that the rate function is distribution free and also does not depend upon the delta-sequence used to estimate the density. A general result pertaining to any regular delta-sequence is stated and a discussion of hypotheses for the most usual methods is given.

1. Introduction. Let \mathcal{X} be an open interval of the real line \mathbf{R} that may be the whole real line. A sequence $\{\delta_m(x, u)\}$ of bounded measurable functions on $\mathcal{X} \times \mathcal{X}$ is a delta-sequence on \mathcal{X} if, for each $x \in \mathcal{X}$ and each \mathcal{C}^∞ -function φ with support in \mathcal{X} , we have

$$\lim_{m \rightarrow \infty} \int_{\mathcal{X}} \delta_m(x, u) \varphi(u) du = \varphi(x).$$

Let X_1, X_2, \dots, X_n be a sequence of i.i.d real random variable defined on a probability space (Ω, \mathcal{F}, P) and taking values in a set \mathcal{X} . Denote by f its probability density function with respect to the Lebesgue measure μ over \mathcal{X} . We shall associate an estimator of f with the sequence $\{\delta_m(x, u)\}$ by letting

$$f_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_m(x, X_i),$$

where $m = m_n$ is a sequence of positive real numbers that tends to infinity with n .

Several estimation methods of an unknown density function f by mean of functions of i.i.d random variables X_1, X_2, \dots, X_n have been proposed in the last five decades. These methods include the kernel method studied by Rosenblatt [13] and Parzen [11], the orthogonal series methods introduced by Chencov [3] and studied by Schwartz [14], Kronmal and Tarter [8] and

AMS 2000 subject classifications: Primary 60F10, 62G07; secondary 62G10

Keywords and phrases: Bahadur slope, composite hypothesis test, delta-sequence estimate, L_1 -distance, large deviations

Walter [15] and the characteristic function approach studied by Blum and Susarla [2]. The delta-sequence method gathers several density estimation methods including among others the following ones

- (i) *Kernel estimator*. For $(x, u) \in \mathbf{R}^2$, set

$$\delta_m(x, u) = mK(m(x - u)),$$

where K is a positive bounded function such that $\int_{-\infty}^{\infty} K(x)dx = 1$ and $\lim_{x \rightarrow \infty} |x| K(x) = 0$. Parzen [11] showed that $\{\delta_m\}$ constitutes a delta-sequence.

- (ii) *Histogram estimator*. For $(x, u) \in [0, 1]^2$, set

$$\delta_m(x, u) = m \sum_{j=1}^m \chi_j(x) \chi_j(u),$$

where χ_1 is the indicator function of the set $(0, 1/m)$, and χ_j is the indicator function of the interval $[(j-1)/m, j/m[$ for $j = 2, \dots, m$. It is easily seen that this case corresponds to be the usual histogram estimator.

- (iii) *Orthogonal series estimator*. Let $\{\psi_m(x)\}$ be a complete orthonormal system on an interval (a, b) consisting of eigenfunctions of a compact operator on $L^2(a, b)$. For $(x, u) \in (a, b)^2$, set

$$\delta_m(x, u) = \sum_{j=1}^m \psi_j(x) \psi_j(u).$$

Walter [15] stated that the sequence $(\delta_m(x, u))$ is a delta-sequence. The sequence $(\psi_m(x))$ may represent the trigonometric functions, the Legendre polynomials, the Haar orthogonal system, the Hermite functions.

- (vi) *Estimator of density type*. Let $\{Y_m\}$ be a sequence of i.i.d. random variables with mean 0 and finite variance, having a bounded density. Let $g_m(x)$ be the density of the sample mean. Then, the sequence of functions

$$\delta_m(x, u) = g_m(x - u) \quad x, u \in \mathbf{R},$$

is a delta-sequence. This may be shown by using the Chebychev's inequality.

- (v) *Fejér kernel estimator*. Consider the Fejér kernel defined, for any $u \in [-\pi, \pi]$, by

$$F_m(u) = \frac{\sin^2((m+1)u/2)}{2\pi(m+1)\sin^2(u/2)}.$$

Winter [16] showed that the sequence of function $\delta_m(x, u) = F_m(x - u)$ constitutes a delta-sequence.

In this paper, we are concerned with uniform L_1 -distance large deviations results of Chernoff-type for the delta-sequence method density estimator. More precisely, we consider uniformity over a class of density functions fulfilling some regularity conditions that will be given later on. A general result pertaining to any regular delta-sequence is stated and a discussion of hypotheses for the most usual methods is given. One of the most important application of large deviations results in statistical analysis is the comparison of tests via their asymptotic efficiency from the Bahadur point of view. This question has been previously addressed by Louani [9] where it is shown that the kernel density L_1 -error based goodness-of-fit test achieves better performances than the well-known Kolmogorov-Smirnov test when testing a simple hypothesis. The results we obtain here allow to consider composite hypotheses.

Let Θ_0 be a class of density functions. Consider the problem of testing the composite hypothesis $H_0 : "f \in \Theta_0"$ against the alternative $A : "f \notin \Theta_0"$ on the basis of observations X_1, X_2, \dots, X_n . For this purpose, use the test statistic

$$V_n = \inf_{h \in \Theta_0} \|f_n - h\|_{L_1}$$

which is naturally significant if larger than some positive threshold. The results we state allow to obtain the Bahadur exact slope associated to the statistic V_n and then to compare the test based on V_n with any other test provided to have its Bahadur slope.

There exists an extensive large deviations literature involving various areas of probability and statistics. For a large corpus of results and applications, we refer to the book of Dembo and Zeitouni [5].

In the sequel, we introduce some notations and tools that allow to perform our study. Define, for any $0 \leq a \leq 1$ the following functions,

$$(1.1) \mathbb{I}_a^+(\lambda) = \begin{cases} (a + \frac{\lambda}{2}) \log(1 + \frac{\lambda}{2a}) + (1 - a - \frac{\lambda}{2}) \log(1 - \frac{\lambda}{2(1-a)}) & \text{if } 0 < \lambda < 2 - 2a, \\ \infty & \text{elsewhere.} \end{cases}$$

$$(1.2) \mathbb{I}_a^-(\lambda) = \begin{cases} (a - \frac{\lambda}{2}) \log(1 - \frac{\lambda}{2a}) + (1 - a + \frac{\lambda}{2}) \log(1 + \frac{\lambda}{2(1-a)}) & \text{if } 0 < \lambda < 2a, \\ \infty & \text{elsewhere.} \end{cases}$$

Notice that $\Gamma_a^-(\lambda)$ is a reflection of $\Gamma_a^+(\lambda)$ about $a = 1/2$ and $\Gamma_a^+(\lambda) \leq \Gamma_a^-(\lambda)$ (respect. $\Gamma_a^+(\lambda) \geq \Gamma_a^-(\lambda)$) whenever $a \leq 1/2$ (respect. $a \geq 1/2$). $\Gamma_{1-a}^+(\lambda) =$

$\Gamma_a^-(\lambda)$, (see Louani [9]). Define now

$$\Gamma_a(\lambda) = \min\{\Gamma_a^+(\lambda), \Gamma_a^-(\lambda)\},$$

and

$$g(\lambda) = \inf\{\Gamma_a(\lambda) : 0 \leq a \leq 1\}.$$

2. Results. From now on, denotes by \mathcal{B} the σ -field of Borel subsets of \mathcal{X} . Let Λ_n be the empirical probability measure associated to the sample X_1, X_2, \dots, X_n and denote by Λ the probability measure associated to the density f . Obviously, one has

$$f_n(x) = \int_{\mathcal{X}} \delta_{m_n}(x, u) d\Lambda_n(u).$$

and

$$E(f_n(x)) = \int_{\mathcal{X}} \delta_{m_n}(x, u) f(u) du.$$

Consider a class Θ of density functions and set some hypotheses upon this class and the delta-sequence (δ_{m_n}) necessary to establish our results.

(H₁) For any $u \in \mathcal{X}$ and any $n \in \mathbb{N}$,

$$\int_{\mathcal{X}} \delta_{m_n}(x, u) du = 1.$$

(H₂) For any $u \in \mathcal{X}$ and any $B \in \mathcal{B}$, we have

$$\int_B |\delta_{m_n}(x, u)| dx \leq C_B,$$

and

$$\lim_{n \rightarrow \infty} \int_B \delta_{m_n}(x, u) dx = \mathbb{1}_{\overset{\circ}{B}}(u) \quad \text{almost everywhere,}$$

where $\overset{\circ}{B}$ is the interior of the set B and C_B is a constant depending only the set B .

(H₃) For any $\epsilon > 0$, there exists n_0 such that $\forall n > n_0$

$$\sup_{f \in \Theta} \|E(f_n) - f\|_{L_1} < \epsilon.$$

(H₄) For any $\epsilon > 0$ there exists a finite collection \mathcal{P}_n of subsets of \mathcal{X} such that

$$\int_{\mathcal{X}} |f_n(x) - E(f_n(x))| dx \leq \sum_{B \in \mathcal{P}_n} |\Lambda_n(B) - \Lambda(B)| + \epsilon,$$

From now on, denotes by $N(\mathcal{P}_n)$ the cardinality of \mathcal{P}_n . The main result of this paper is given in the following theorem.

Theorem 2.1 *Assume that the conditions (H_1) – (H_4) are satisfied. If $\lim_{n \rightarrow \infty} \frac{N(\mathcal{P}_n)}{n} = 0$, then, for any $\lambda > 0$, we have*

$$(2.1) \quad \lim_{n \rightarrow \infty} \sup_{f \in \Theta} \frac{1}{n} \log P(\|f_n - f\|_{L_1} > \lambda) = -g(\lambda).$$

Remark 2.1 The left hand side of the statement (2.1) is bounded from below under only the condition (H_2) .

Discussion of the conditions

When a function $h \in L_1([0, 1])$, it is well-known (see, for instance, Devroye and Györfi [6], page 291) that the Haar orthogonal system expansion of h converges to h almost everywhere. Since for any $B \subset [0, 1]$, $\mathbb{I}_B \in L_1([0, 1])$, it follows then that the condition (H_2) is satisfied.

In Lemma 2.1, Lemma 2.2 and Proposition below, sufficient conditions that allow hypotheses (H_2) and (H_3) to be satisfied are given. In the first place, we enumerate these conditions.

- (B_1) For n large enough, there exists a positive function γ_n defined on \mathcal{X} such that for any $(x, u) \in \mathcal{X}^2$, $|\delta_{m_n}(x, u)| \leq \gamma_n(x - u)$.
- (B_2) For n large enough, there exists a positive $C > 0$ such that $\int \gamma_n(y) dy < C$.
- (B_3) For any $\alpha > 0$ $\lim_{n \rightarrow \infty} \int_{|y| \geq \alpha} |\gamma_n(y)| dy = 0$.
- (B_4) Let $s \geq 1$ such that

$$\lim_{n \rightarrow \infty} \int |y|^j \gamma_{m_n}(y) dy = 0 \quad j = 1, \dots, s.$$

The following result, which is close to Bochner's Theorem, gives sufficient conditions for which (H_2) is satisfied.

Lemma 2.1 *Suppose that the conditions (B_1) – (B_3) are satisfied. If h is a continuous at a point u and $\sup_{x \in \mathcal{X}} |h(x)| < \infty$, then*

$$\int_{\mathcal{X}} h(x) \delta_{m_n}(x, u) dx \rightarrow h(u) \quad \text{as } n \rightarrow \infty.$$

Remark 2.2 The conditions (B_1) – (B_3) are satisfied for several usual delta-sequences, including, the Fejér kernel, kernel and density type methods. Since for any $B \in \mathcal{B}$, the function $\mathbb{I}_B(x)$ satisfies the conditions of Lemma 2.1, it follows then that the condition (H_2) holds true.

We begin the discussion of the condition (H_3) by an example relative to the Haar orthogonal system estimate.

For α and β universal positive constants, let

$$Lip(\alpha, 1, \beta) = \{f \in L_1([0, 1]) : w_1(f, \eta) \leq \beta \eta^\alpha\},$$

where

$$w_1(f, \eta) = \sup_{0 < h \leq \eta} \left\{ \int_0^1 |f(s+h) - f(s)| ds \right\}$$

is the modulus of continuity of the function f . It follows from Ciesielski [4], whenever $f \in L_1([0, 1])$, that

$$\|Ef_n - f\|_{L_1} \leq 6w_1(f, \frac{1}{m_n}),$$

where f_n is the Haar orthogonal system estimate of f . Therefore, when $\Theta = Lip(\alpha, 1, \beta)$, we have

$$\sup_{f \in \Theta} \|Ef_n - f\|_{L_1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Consider now the class Θ of density functions defined as follows. For any $\epsilon > 0$, there exists $\eta > 0$, such that

$$(2.2) \quad \sup_{f \in \Theta} \sup_{|y| < \eta} \int |f(x-y) - f(x)| dx < \epsilon.$$

The following result gives sufficient conditions that allow the condition (H_3) to hold true.

Lemma 2.2 Assume that the conditions (H_1) , (B_1) – (B_3) are satisfied. If Θ is the class defined by the statement (2.2), then

$$\sup_{f \in \Theta} \|Ef_n - f\|_{L_1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Suppose now that Θ is a class of densities with s absolutely continuous derivatives and that $\sup_{f \in \Theta} \|f^{(j)}(u)\|_{L_1} < \infty$, $j = 1, \dots, s$. The following

result gives sufficient conditions that allow the condition (\mathbf{H}_3) to hold true.

Proposition 2.1 *Assume that the conditions (\mathbf{H}_1) , (\mathbf{B}_1) , (\mathbf{B}_4) are satisfied. If Θ is the class of densities with s absolutely continuous derivatives and $\sup_{f \in \Theta} \|f^{(j)}(u)\|_{L_1} < \infty$, $j = 1, \dots, s$, then*

$$\sup_{f \in \Theta} \|Ef_n - f\|_{L_1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

As for the condition (\mathbf{H}_3) , we begin the discussion of the condition (\mathbf{H}_4) by the case pertaining to the Haar orthogonal system.

Following Haar [7], whenever $m = 2^l + j$ with $1 \leq j \leq 2^l$, $l = 0, 1, \dots$, the explicit form of the Haar kernel is given by

$$\delta_m(x, u) = \begin{cases} \frac{1}{t_k - t_{k-1}} & \text{if } (x, u) \in (t_{k-1}, t_k)^2, \text{ for some } k = 1, \dots, m, \\ 0 & \text{elsewhere,} \end{cases}$$

where

$$t_i = \begin{cases} \frac{i}{2^{l+1}} & \text{for } i = 0, \dots, 2j, \\ \frac{i-j}{2^l} & \text{for } i = 2j+1, \dots, m. \end{cases}$$

It is then clear that

$$\begin{aligned} \int_0^1 |f_n(x) - E(f_n(x))| dx &= \int_0^1 \left| \int_0^1 \delta_{m_n}(x, u) (d\Lambda_n(u) - d\Lambda(u)) \right| dx, \\ &\leq \int_0^1 \int_0^1 \left| \sum_{k=1}^{m_n} \frac{1}{t_k - t_{k-1}} \mathbb{I}_{(t_{k-1}, t_k)}(x) \mathbb{I}_{(t_{k-1}, t_k)}(u) (d\Lambda_n(u) - d\Lambda(u)) \right| dx, \\ &\leq \sum_{k=1}^{m_n} \frac{1}{t_k - t_{k-1}} \int_0^1 \mathbb{I}_{(t_{k-1}, t_k)}(x) \left| \int_0^1 \mathbb{I}_{(t_{k-1}, t_k)}(u) (d\Lambda_n(u) - d\Lambda(u)) \right| dx, \\ &\leq \sum_{k=1}^{m_n} |\Lambda_n((t_{k-1}, t_k)) - \Lambda((t_{k-1}, t_k))|. \end{aligned}$$

Therefore, the condition (\mathbf{H}_4) holds.

The following Lemma gives a sufficient condition for which hypothesis (\mathbf{H}_4) is satisfied.

Lemma 2.3 *Suppose that, for n large enough, there exists a positive function γ_n defined on \mathcal{X} such that for any $\epsilon > 0$ and any $u \in \mathcal{X}$, $\int_{\mathcal{X}} |\delta_{m_n}(x, u) -$*

$\gamma_n(x-u)|dx \leq \epsilon$ and $\int_{\mathcal{X}} \gamma_n(y)dy \leq 1 + \epsilon$. Then the condition (H_4) holds true.

Application

Large deviations results are useful and efficient tools to study the asymptotic efficiency of tests. This question has been widely investigated; we refer to Bahadur [1] and the book of Nikitin [10] for an account of results on this subject. In testing the composite H_0 , the rejection region associated to the test statistic V_n is given by

$$R_n = \{V_n > c\}$$

for some positive real number c . For any density f and any real number λ , set $D_n(\lambda, f) = P_f(V_n \leq \lambda)$, where P_f denotes the distribution of observations when f is the underlying density. It is obvious that, under the null hypothesis H_0 , we have

$$D_n(\lambda, f) = P_f(\|f_n - f\|_{L_1} \leq \lambda).$$

For any $\lambda \in \mathbf{R}$, define

$$\Delta_n(\lambda) = \inf\{D_n(\lambda, f) : f \in \Theta_0\}.$$

The P -value relative to the test statistic V_n is $L_n = 1 - \Delta_n(V_n)$. The following corollary gives the asymptotic behavior of the P -value relative to the statistic (V_n) .

Corollary 2.1 *Under conditions of theorem 2.1, for any $f \notin \Theta_0$, we have with P_f -probability one,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log L_n = -g\left(\inf_{h \in \Theta_0} \|f - h\|_{L_1}\right).$$

Remark 2.3 From Corollary above, we deduce that the Bahadur exact slope relative to the statistic (V_n) is $2g(\inf_{h \in \Theta_0} \|f - h\|_{L_1})$.

3. Proofs. Proof of Lemma 2.1 Using the conditions the (B_1) , it follows that

$$\begin{aligned} \left| \int_{\mathcal{X}} h(x) \delta_{m_n}(x, u) dx - h(u) \int_{\mathcal{X}} \delta_{m_n}(x, u) dx \right| &\leq \int_{\mathcal{X}} |(h(x) - h(u)) \delta_{m_n}(x, u)| dx, \\ &\leq \int_{\mathcal{X}} |h(u+y) - h(u)| \gamma_n(y) dy \end{aligned}$$

Since the function h is continuous at a point x , for any $\epsilon > 0$, there exists $\alpha > 0$ such that $|h(u+y) - h(u)| \leq \epsilon$, provided that $|y| \leq \alpha$.

Thus,

$$\begin{aligned} \int_{\mathcal{X}} |(h(u+y) - h(u)) \gamma_n(y)| dy &\leq \int_{|y| \leq \alpha} |h(u+y) - h(u)| \gamma_n(y) dy \\ &\quad + \int_{|y| \geq \alpha} |h(u)| \gamma_n(y) dy + \int_{|y| \geq \alpha} |h(u+y)| \gamma_n(y) dy, \\ &\leq \int_{|y| \leq \alpha} |h(u+y) - h(u)| \gamma_n(y) dy \\ &\quad + 2 \sup_{x \in \mathcal{X}} |h(x)| \int_{|y| \geq \alpha} \gamma_n(y) dy. \end{aligned}$$

Making use of the conditions (B_2) - (B_3) , it follows that the first term is not greater than ϵ and the second term tends to zero. ■

Proof of Lemma 2.2 Using the conditions the (H_1) and (B_1) , it follows that

$$\begin{aligned} \int_{\mathcal{X}} \left| \int_{\mathcal{X}} f(u) \delta_{m_n}(x, u) du - f(x) \right| dx &\leq \int_{\mathcal{X}} \left| \int_{\mathcal{X}} (f(u) - f(x)) \delta_{m_n}(x, u) \right| du dx, \\ &\leq \int_{\mathcal{X}} \int_{\mathcal{X}} |f(x-y) - f(x)| \gamma_n(y) dx dy. \end{aligned}$$

Using the statement (2.2), it follows that for any $\epsilon > 0$ and any $f \in \Theta$, there exists $\eta > 0$ such that

$$\sup_{|y| < \eta} \int |f(x-y) - f(x)| dx < \epsilon.$$

Therefore,

$$\begin{aligned} \int_{\mathcal{X}} \int_{\mathcal{X}} |f(x-y) - f(x)| \gamma_n(y) dx dy &\leq \int \int_{|y| < \eta} |f(x-y) - f(x)| \gamma_n(y) dy dx, \\ &\quad + \int_{|y| \geq \eta} \int f(x-y) \gamma_n(y) dx dy \\ &\quad + \int_{|y| \geq \eta} \int f(x) |\gamma_n(y)| dx dy, \\ &\leq \sup_{|y| < \eta} \int |f(x-y) - f(x)| dx \int_{|y| < \eta} |\gamma_n(y)| dy \\ &\quad + 2 \int_{|y| \geq \eta} |\gamma_n(y)| dy. \end{aligned}$$

the first term is not greater than ϵ for any $f \in \Theta$ and the second term tends to zero as $n \rightarrow \infty$. ■

Proof of Proposition 2.1 Using the Taylor's formula, we have

$$f(u) - f(x) = \sum_{j=1}^{s-1} \frac{f^{(j)}(x)}{j!} (u-x)^j + \int_0^1 (u-x)^s \frac{(1-t)^{s-1}}{(s-1)!} f^{(s)}(x+t(u-x)) dt.$$

Thus,

$$\begin{aligned} Ef_n(x) - f(x) &= \int_{\mathcal{X}} (f(u) - f(x)) \delta_{m_n}(x, u) du, \\ &= \sum_{j=1}^{s-1} \frac{f^{(j)}(x)}{j!} \int_{\mathcal{X}} (u-x)^j \delta_{m_n}(x, u) du \\ &\quad + \int_0^1 \int_{\mathcal{X}} \delta_{m_n}(x, u) (u-x)^s \frac{(1-t)^{s-1}}{(s-1)!} f^{(s)}(x+t(u-x)) du dt, \end{aligned}$$

and hence

$$\begin{aligned} |Ef_n(x) - f(x)| &\leq \sum_{j=1}^{s-1} \frac{f^{(j)}(x)}{j!} \int_{\mathcal{X}} |x-u|^j \gamma_{m_n}(x-u) du \\ &\quad + \int_0^1 dt \int_{\mathcal{X}} \gamma_{m_n}(y) |y|^s \frac{(1-t)^{s-1}}{(s-1)!} |f^{(s)}(x-ty)| du, \end{aligned}$$

we used here the inequality $|\delta_{m_n}(x, u)| \leq \gamma_{m_n}(x-u)$ and set $x-u = y$.

Therefore, by Young's and Minkowsky's inequalities,

$$\begin{aligned} \|Ef_n(x) - f(x)\|_{L_1} &\leq \sum_{j=1}^{s-1} \|f^{(j)}\|_{L_1} \int_{\mathcal{X}} \frac{|y|^j}{j!} \gamma_{m_n}(y) dy \\ &\quad + \int_0^1 \frac{(1-t)^{s-1}}{(s-1)!} dt \int_{\mathcal{X}} |y|^s \gamma_{m_n}(y) \|f^{(s)}\|_{L_1} dy. \end{aligned}$$

Making use of the condition (B_4) and the fact that $\sup_{f \in \Theta} \|f^{(j)}(u)\|_{L_1} < \infty$, $j = 1, \dots, s$, it follows that the condition (H_3) is satisfied. ■

Proof of Lemma 2.3 Using the classical result about the approximation of an integrable function γ_n by a step function, for any n and any $\epsilon > 0$, there exists a function

$$(3.1) \quad L_n(y) = \sum_{j=1}^{d_n} \alpha_j^n \mathbb{I}_{A_j^n}(y),$$

where d_n is a finite integer, (α_j^n) are non negative finite real numbers and A_j^n is a partition of \mathcal{X} such that

$$(3.2) \quad \int_{\mathcal{X}} |\gamma_n(y) - L_n(y)| dy < \epsilon.$$

Hence, for any $u \in \mathcal{X}$

$$\int_{\mathcal{X}} |\delta_{m_n}(x, u) - L_n(x - u)| dx < 2\epsilon.$$

For $\epsilon > 0$, choose $d_n, \alpha_1^n, \alpha_2^n, \dots, \alpha_{d_n}^n$ et $A_1^n, \dots, A_{d_n}^n$ such that the statements (3.1)-(3.2) are satisfied. Define \bar{f}_n as f_n with the function L_n instead of γ_n . It follows easily that

$$\begin{aligned} \int_{\mathcal{X}} |f_n(x) - E(f_n(x))| dx &\leq \int_{\mathcal{X}} |f_n(x) - \bar{f}_n(x)| dx \\ &+ \int_{\mathcal{X}} |\bar{f}_n(x) - E(\bar{f}_n(x))| dx + \int_{\mathcal{X}} |E(f_n(x)) - E(\bar{f}_n(x))| dx, \\ &\leq \int_{\mathcal{X}} |\bar{f}_n(x) - E(\bar{f}_n(x))| dx + 4\epsilon. \end{aligned}$$

It is clear that

$$\int_{\mathcal{X}} |\bar{f}_n(x) - E(\bar{f}_n(x))| dx \leq \sum_{j=1}^{d_n} |\alpha_j^n| \int_{\mathcal{X}} |\Lambda_n(x + A_j^n) - \Lambda(x + A_j^n)| dx,$$

where $x + A = \{z = x + y : y \in A\}$.

Consider the partition \prod_l of the real line into intervals of length $1/lm_n$, where l is a positive real number that will be chosen later on. For any $x \in \mathbf{R}$ and any $1 \leq j \leq d_n$, define $R_{j,n}^x$ as the union of all sets $B \in \prod_l$ such that $B \subset x + A_j^n$ and consider the set given by $S_{j,n}^x = (x + A_j^n) - R_{j,n}^x$. It is easily seen that $S_{j,n}^x$ is a subset of $x + (A_j^n - R_{j,n})$ for some set $R_{j,n}$ included in A_j^n . Thus

$$|\Lambda_n(x + A_j^n) - \Lambda(x + A_j^n)| \leq \sum_{B \in \prod_l, B \subset x + A_j^n} |\Lambda_n(B) - \Lambda(B)| + (\Lambda_n + \Lambda)(S_{j,n}^x),$$

and then,

$$\begin{aligned} \int_{\mathcal{X}} |\bar{f}_n(x) - E(\bar{f}_n(x))| &\leq \sum_{j=1}^{d_n} |\alpha_j^n| \int_{\mathcal{X}} \left(\sum_{B \in \prod_l, B \subset x + A_j^n} |\Lambda_n(B) - \Lambda(B)| \right. \\ &\quad \left. + (\Lambda_n + \Lambda)(S_{j,n}^x) \right) dx. \end{aligned}$$

Using the Young's inequality, we obtain

$$\begin{aligned}
 \int |\bar{f}_n(x) - E(\bar{f}_n(x))| dx &\leq \sum_{j=1}^{d_n} |\alpha_j^n| \sum_{B \in \prod_l} |\Lambda_n(B) - \Lambda(B)| \int_{B \subset x + A_j^n} dx \\
 &+ 2 \sum_{j=1}^{d_n} |\alpha_j^n| \mu(A_j^n - R_{j,n}), \\
 &\leq \sum_{j=1}^{d_n} |\alpha_j^n| \mu(A_j^n) \sum_{B \in \prod_l} |\Lambda_n(B) - \Lambda(B)| \\
 &+ 2 \sum_{j=1}^{d_n} |\alpha_j^n| \mu(A_j^n - R_{j,n}).
 \end{aligned}$$

The second term in the last inequality may be made smaller than ϵ by choosing l large enough such that for any $1 \leq j \leq d_n$, $\mu(A_j^n - R_{j,n}) \leq \epsilon \mu(A_j^n)$. Thus, since $\int |L_n(x)| dx = \sum_{j=1}^{d_n} |\alpha_j^n| \mu(A_j^n) \leq 1 + \epsilon$, it follows that

$$\begin{aligned}
 \int |\bar{f}_n(x) - E(\bar{f}_n(x))| dx &\leq \sum_{j=1}^{d_n} |\alpha_j^n| \mu(A_j^n) \sum_{B \in \prod_l} |\Lambda_n(B) - \Lambda(B)| + 6\epsilon \\
 &\leq \sum_{B \in \mathcal{P}_l} |\Lambda_n(B) - \Lambda(B)| + 8\epsilon,
 \end{aligned}$$

where \mathcal{P}_l is the collection of the sets of \prod_l having nonempty intersection with \mathcal{X} . Remark that if $\mu(\mathcal{X}) < \infty$, then $N(\mathcal{P}_l) \leq \mu(\mathcal{X}) l m_n$.

If $\mu(\mathcal{X}) = \infty$, the collection of sets \mathcal{P}_l is infinite and therefore is not usable as it is. It is necessary to cut off the tails of the distribution to reduce the cardinality of \mathcal{P}_l . To this end, consider a finite collection of sets $Q_{r,l}$ consisting of those sets of \mathcal{P}_l having nonempty intersection with $[-r, r]$, where $r > 0$ is a real number to be chosen later on. Define now the collection $\mathcal{P}_{r,l}$ to be $Q_{r,l} \cup \{(-\infty, -r) \cup (r, \infty)\}$. It is clear that

$$\begin{aligned}
 \sum_{B \in \mathcal{P}_l} |\Lambda_n(B) - \Lambda(B)| &\leq \sum_{B \in Q_{r,l}} |\Lambda_n(B) - \Lambda(B)| + (\Lambda_n + \Lambda)((-\infty, -r) \cup (r, \infty)) \\
 &\leq \sum_{B \in \mathcal{P}_{r,l}} |\Lambda_n(B) - \Lambda(B)| + 2\Lambda((-\infty, -r) \cup (r, \infty)) \\
 &\leq \sum_{B \in \mathcal{P}_{r,l}} |\Lambda_n(B) - \Lambda(B)| + 2\epsilon,
 \end{aligned}$$

since for r large enough, $\sup_{f \in \Theta} \Lambda((-\infty, -r) \cup (r, \infty)) \leq \epsilon$. Note in this case that $N(\mathcal{P}_{r,l}) \leq 2r l m_n + 2$. ■

Proof of Theorem 2.1 Observe first that, using Scheffé's Theorem, one may write

$$\|f_n - f\|_{L_1} = 2 \sup_{B \in \mathcal{B}} \left| \int_B f_n(x) dx - \int_B f(x) dx \right|.$$

The lower bound. For any set $B \in \mathcal{B}$ and any $\lambda > 0$, we have

$$P\left(2 \int_B (f_n(x) - f(x)) dx > \lambda\right) = P\left(\sum_{i=1}^n \int_B \delta_{m_n}(x, X_i) dx > n \left(\frac{\lambda}{2} + \int_B f(x) dx\right)\right).$$

Denote by $\phi_n^B(t)$ the moment generating function of $\sum_{i=1}^n \int_B \delta_{m_n}(x, X_i) dx$ which is given by

$$\begin{aligned} \phi_n^B(t) &= E\left[\exp\left\{t \sum_{i=1}^n \int_B \delta_{m_n}(x, X_i) dx\right\}\right] = \left(E\left[\exp\left\{t \int_B \delta_{m_n}(x, X_i) dx\right\}\right]\right)^n \\ &:= (\varphi_n^B(t))^n. \end{aligned}$$

Thus,

$$\varphi_n^B(t) = \int_{\mathcal{X}} \exp\left\{t \int_B \delta_{m_n}(x, u) dx\right\} f(u) du.$$

Note that, by condition (H_2) , we have

$$\int_B \delta_{m_n}(x, u) dx \rightarrow \mathbb{I}_B(u) \text{ almost everywhere as } n \rightarrow \infty,$$

Therefore, using the dominated convergence theorem, this yields

$$\lim_{n \rightarrow \infty} \varphi_n^B(t) = 1 + (e^t - 1) \int_B f(u) du.$$

Hence

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \phi_n^B(t) = \log \left[1 + (e^t - 1) \int_B f(u) du \right] := \phi^B(t).$$

The remainder of the proof of the result of large deviations when dealing with the probability $P(2 \int_B (f_n(x) - f(x)) dx > \lambda)$ uses essentially Chebychev's inequality for deriving the upper bound and an exponential change of measure to obtain the lower bound (see for instance, Plachky and Steinebach, [12]). The rate function is given in (1.1) where we have set $\int_B f(u) du = a$.

The negative version result corresponding to the probability

$$P\left(2\int_B(f(x) - f_n(x))dx > \lambda\right)$$

is obtained similarly. The rate function is given in (1.2). The result corresponding to the absolute deviation between $\int_B f_n(u)du$ and $\int_B f(u)du$ follows from the fact that

$$\begin{aligned} & \max \left\{ P\left(2\int_B(f(x) - f_n(x))dx > \lambda\right); P\left(2\int_B(f_n(x) - f(x))dx > \lambda\right) \right\} \\ & \leq P\left(2\left|\int_B f(x) - f_n(x)dx\right| > \lambda\right) \leq \\ & 2 \max \left\{ P\left(2\int_B(f(x) - f_n(x))dx > \lambda\right); P\left(2\int_B(f_n(x) - f(x))dx > \lambda\right) \right\}. \end{aligned}$$

Using the Scheffé's Theorem, it follows that for any $B \in \mathcal{B}$,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P\left(\|f_n - f\|_{L_1} > \lambda\right) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log P\left(2 \sup_{B \in \mathcal{B}} \left|\int_B f_n(x)dx - \int_B f(x)dx\right| > \lambda\right), \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P\left(2 \left|\int_B f_n(x)dx - \int_B f(x)dx\right| > \lambda\right), \\ &\geq -\Gamma_a(\lambda). \end{aligned}$$

Thus,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log P(\|f_n - f\|_{L_1} > \lambda) \geq -g(\lambda)$$

The upper bound. From conditions (H_3) -(H_4), it follows that there exists $\epsilon > 0$ independent of f , such that for any $f \in \Theta$,

$$P(\|f_n - f\|_{L_1} > \lambda) \leq P\left(\sum_{B \in \mathcal{P}_n} |\Lambda_n(B) - \Lambda(B)| > \lambda - 2\epsilon\right).$$

Again by Scheffé's Theorem, we obtain from the previous that

$$\begin{aligned} P(\|f_n - f\|_{L_1} > \lambda) &\leq P\left(\sup_{B \in \mathcal{P}_n} |\Lambda_n(B) - \Lambda(B)| > \frac{\lambda - 2\epsilon}{2}\right), \\ &\leq \sum_{B \in \mathcal{P}_n} P\left(|\Lambda_n(B) - \Lambda(B)| > \frac{\lambda - 2\epsilon}{2}\right), \end{aligned}$$

where \mathcal{P}_n is the set of all possible sets given by unions of elements of the collection \mathcal{P}_n .

The probability of large deviation version corresponding to absolute deviation between measures Λ_n and Λ over borel set B follows by the same arguments as for the absolute deviation between integrals, over some interval B , of the functions f_n and f . Hence,

$$P\left(|\Lambda_n(B) - \Lambda(B)| > \frac{\lambda - 2\epsilon}{2}\right) \leq \exp\left\{-n\Gamma_{\Lambda(B)}(\lambda - 2\epsilon)(1 + o(1))\right\}.$$

Therefore,

$$\begin{aligned} P(\|f_n - f\|_{L_1} > \lambda) &\leq \sum_{B \in \mathcal{P}_n} \exp\left\{-n\Gamma_{\Lambda(B)}(\lambda - 2\epsilon)(1 + o(1))\right\} \\ &\leq N(\mathcal{P}_n) \sup_{B \in \mathcal{P}_n} \exp\left\{-n\Gamma_{\Lambda(B)}(\lambda - 2\epsilon)(1 + o(1))\right\} \\ &\leq 2^{N(\mathcal{P}_n)} \exp\{-ng(\lambda - 2\epsilon)(1 + o(1))\}. \end{aligned}$$

Since $(N(\mathcal{P}_n)/n) \rightarrow 0$ as $n \rightarrow \infty$, it suffices then to tend n to infinity and to make use the fact that g is an increasing and continuous function to achieve the proof. ■

Proof of Corollary 2.1 Making use of the L_1 -consistency of the estimator f_n , it follows that

$$V_n \rightarrow \inf_{h \in \Theta_0} \|h - f\| \quad n \rightarrow \infty.$$

Therefore, for f fixed arbitrarily in the complementary of Θ_0 and $\epsilon > 0$, we have, for n large enough

$$P_f\left(\inf_{h \in \Theta_0} \|h - f\| - \epsilon \leq V_n \leq \inf_{h \in \Theta_0} \|h - f\| + \epsilon\right) > 1 - \delta$$

with $\delta > 0$ arbitrarily small. Since D_n is monotone, we have for any $f \in \Theta_0$

$$1 - D_n\left(\inf_{h \in \Theta_0} \|h - f\| + \epsilon, f\right) \leq 1 - D_n(V_n, f) \leq 1 - D_n\left(\inf_{h \in \Theta_0} \|h - f\| - \epsilon, f\right)$$

with P_f -probability $1 - \delta$. Therefore, with P_f -probability $1 - \delta$, we have

$$\begin{aligned} \sup_{f \in \Theta_0} [1 - D_n(\inf_{h \in \Theta_0} \|h - f\| + \epsilon, f)] &\leq \sup_{f \in \Theta_0} [1 - D_n(V_n, f)], \\ &\leq \sup_{f \in \Theta_0} [1 - D_n(\inf_{h \in \Theta_0} \|h - f\| - \epsilon, f)], \end{aligned}$$

and

$$1 - \Delta_n(\inf_{h \in \Theta_0} \|h - f\| - \epsilon, f) \leq L_n \leq 1 - \Delta_n(\inf_{h \in \Theta_0} \|h - f\| + \epsilon, f).$$

Since

$$1 - \Delta_n(\lambda) = \sup_{f \in \Theta_0} P_f(V_n > \lambda) = \sup_{f \in \Theta_0} P_f(\|f_n - f\|_{L_1} > \lambda),$$

making use of Theorem, it follows that

$$\begin{aligned} -g(\inf_{h \in \Theta_0} \|h - f\| - \epsilon) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log(L_n) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log(L_n) \leq -g(\inf_{h \in \Theta_0} \|h - f\| + \epsilon). \end{aligned}$$

Since g is continuous, we achieve the proof by making ϵ tend to zero. ■

References.

- [1] BAHADUR, R.R. (1971). *Some Limit Theorems in Statistic*. SIAM, Philadelphia, Pennsylvania.
- [2] BLUM, J. and SUSARLA, V. (1977). A Fourier inversion method for the estimation of a density and its derivatives. *J. Austral. Math. Soc.* **22**, 166–171.
- [3] CHENCOV, N.N. (1962). Evaluation of an unknown distribution density from observations. *Soviet. Math.* **3**, 1559–1562.
- [4] CIESIELSKI, Z. (1966). Properties of the orthonormal Franklin system. II. *Studia. Math.* **27**, 289–323.
- [5] DEMBO, A. and ZEITOUNI, O. (1993). *Large deviations techniques and applications*. Jones and Bartlett Publishers.
- [6] DEVROYE, L. and GYÖRFI, L. (1985). *Nonparametric Density Estimation. The L_1 -View*. John Wiley, New York.
- [7] HAAR, A. (1910). Zur Theorie der orthogonales Funcktionen-systeme. *Math. Ann.* **69**, 331–371.
- [8] KRONMAL, R. and TARTER, M. (1968). the estimation of probability densities and cumulatives by Fourier series methods. *J. Amer. Statist. assoc.* **63**, 925–952.
- [9] LOUANI, D. (2000). Large deviations for the L_1 -distance in kernel density estimation. *Journal of Statistical Planning and Inference.* **90**, 177–182.

- [10] NIKITIN, Y. (1995). *Asymptotic efficiency of nonparametric tests*. Cambridge University Press.
- [11] PARZEN, E. (1962). On the estimation of a probability density function and mode. *Ann. Math. Statist.* **33**, 1065–1076.
- [12] PLACHKY, D. and STEINEBACH, J. (1975). A Theorem about probabilities of large deviations with application to queuing theory. *Period. Math. Hungaria*. Vol. **6**, 343–345.
- [13] ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27**, 832–837.
- [14] SCHWARTZ, S.C. (1967). Estimation of probability density by orthogonal series. *Ann. Math. Statist.* **38**, 1261–1265.
- [15] WALTER, G. (1965). Expansions of distributions. *Trans. Amer. Math. Soc.* **116**, 492–510.
- [16] WINTER, B.B. (1975). Rate of strong consistency of two nonparametric density estimators. *Ann. Statist.* **3**, 759–766.

175, RUE DU CHEVALERET, 8ÈME ÉTAGE,
BATIMENT A, 75013 PARIS FRANCE
E-MAIL: louani@ccr.jussieu.fr
berrahou@ccr.jussieu.fr

1. Estimation récurrente en analyse des données

1.1. Cadre d'application

Dans [14], [15], [16], l'auteur présente que de nos jours, des volumes énormes de données posent sur des variables de probabilité, comme par exemple les transactions des produits des clients dans les magasins informatiques, sont collectées en ligne par des ordinateurs chaque jour, chaque semaine, chaque année et souvent non exploités par manque de méthodes adaptées. Ce sur de données abondantes est une collection de données qui ont des particularités de type de données [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100].

En outre, on sait que, par exemple dans l'industrie chimique, le nombre de variables de probabilité posées peut être de plusieurs milliers. On se trouve donc confronté à un double problème : des données abondantes, et souvent en ligne, et de grande dimension [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100].