Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 118 (2008) 452–473
stochastic processes and their applications
www.elsevier.com/locate/spa

Moderate deviations and law of the iterated logarithm in $L_1(\mathbb{R}^d)$ for kernel density estimators

Fuqing Gao

School of Mathematics and Statistics, Wuhan University, Wuhan 430072, PR China

Received 27 January 2005; received in revised form 25 April 2007; accepted 26 April 2007
Available online 29 April 2007

Abstract

Let $f_n(x)$ be the non-parametric kernel density estimator of a density function $f(x)$ based on a kernel function K. In this paper, we first prove two moderate deviation theorems in $L_1(\mathbb{R}^d)$ for $\{f_n(x), n \geq 1\}$. Then, as an application of the moderate deviations, we obtain a law of the iterated logarithm for $\{\|f_n - Ef_n\|_1, n \geq 1\}$.

MSC: primary 60F10; 62G07; secondary 62F12

Keywords: Kernel density estimator; Moderate deviations; Law of the iterated logarithm

1. Introduction and main results

Let $\{X_i; i \geq 1\}$ be a sequence of independent and identically distributed ($i.i.d.$) random variables taking values in $\mathbb{R}^d$, defined on a probability space $(\Omega, \mathcal{F}, P)$ with unknown density function $f(x)$. Let K be a measurable function such that

$$K \geq 0, \quad \int_{\mathbb{R}^d} K(x)\mathrm{d}x = 1. \tag{1.1}$$

The kernel density estimator of f, based on the kernel function K, is defined by

$$f_n(x) = \frac{1}{na_n^d} \sum_{i=1}^{n} K\left(\frac{x - X_i}{a_n}\right), \quad x \in \mathbb{R}^d \tag{1.2}$$

E-mail address: fqgao@whu.edu.cn.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2007.04.010

where $\{a_n, n \geq 1\}$ is a bandsequence (width of windows), that is, a sequence of positive numbers tending to 0 and satisfying

$$na_n^d \to +\infty \quad \text{as } n \to \infty. \tag{1.3}$$

As usual, we denote by $\|g\|_p = (\int_{\mathbb{R}^d} |g(x)|^p \mathrm{d}x)^{1/p}$, $p \geq 1$.

In [4] (see also [5]), Devroye proved that all types of L_1-consistency are equivalent to (1.3). The asymptotic normality of $\{\|f_n - Ef_n\|_1, n \geq 1\}$ was studied by Csörgö and Horváth [1] and Horváth [7]. More recently, Giné, Mason and Zaitsev [13] considered the asymptotic normality of the L_1-norm density estimator process, Louani [17] and Lei, Wu and Xie [15] (see Lei and Wu [16] for density estimator in a Markov process) studied the large deviations in $L_1(\mathbb{R}^d)$ for $\{f_n(x), n \geq 1\}$. For the uniform consistency, and uniform large deviations and uniform moderate deviations for $\{f_n(x), n \geq 1\}$, we refer to Einmahl and Mason [6], Giné and Guillou [10], Giné, Koltchinskii and Zinn [11], Louani [18], Gao [8] and the references therein. Giné and Mason [12] considered the law of the iterated logarithm for $\{\|f_n - Ef_n\|_2^2 - E\|f_n - Ef_n\|_2^2, n \geq 1\}$ by the KMT approximation, and indicated that their methods do not extend to the cases $\|\cdot\|_p$, $p \neq 2$. The purpose of this paper is to study the moderate deviations and the law of the iterated logarithm in $L_1(\mathbb{R}^d)$ for $\{f_n, n \geq 1\}$. We find the best condition on the bandsequence such that $\{\|f_n - Ef_n\|_1, n \geq 1\}$ satisfies the moderate deviation principle. A law of the iterated logarithm for $\{\|f_n - Ef_n\|_1, n \geq 1\}$ is also obtained.

Let $b_n, n \geq 1$ be a sequence of positive real numbers satisfying

$$\frac{n}{b_n} \to +\infty \quad \text{and} \quad \frac{n}{b_n^2} \to 0 \quad \text{as } n \to +\infty. \tag{1.4}$$

We introduce the following condition:

$$(\mathrm{H}_1) \int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right) K^2(x)\mathrm{d}x < \infty, \qquad \int_{\mathbb{R}^d} |x|^{pd} f(x)\mathrm{d}x < \infty, \quad \text{for some } p > 1.$$

Remark 1.1. If (H_1) holds, then

$$\int_{\mathbb{R}^d} \sqrt{f(x)}\mathrm{d}x \leq \left(\int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right)^{-1} \mathrm{d}x\right)^{1/2} \left(\int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right) f(x)\mathrm{d}x\right)^{1/2} < \infty,$$

and

$$\begin{aligned}
\int_{\mathbb{R}^d} \sqrt{\int_{\mathbb{R}^d} \frac{1}{a_n^d} K^2\left(\frac{x-y}{a_n}\right) f(y)\mathrm{d}y\mathrm{d}x} &= \int_{\mathbb{R}^d} \sqrt{\int_{\mathbb{R}^d} K^2(z) f(x - a_n z)\mathrm{d}z\mathrm{d}x} \\
&\leq \left(\int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right)^{-1} \mathrm{d}x\right)^{1/2} \left(\int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right) \int_{\mathbb{R}^d} K^2(z) f(x - a_n z)\mathrm{d}z\mathrm{d}x\right)^{1/2} \\
&\leq \left(\int_{\mathbb{R}^d} \left(1 + |x|^{pd}\right)^{-1} \mathrm{d}x\right)^{1/2} \\
&\quad \times \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(1 + 2^{pd-1}(|u|^{pd} + |a_n v|^{pd})\right) K^2(u) f(v)\mathrm{d}u\mathrm{d}v\right)^{1/2},
\end{aligned}$$

and so

$$\sup_{n \geq 1} \int_{\mathbb{R}^d} \sqrt{\int_{\mathbb{R}^d} \frac{1}{a_n^d} K^2\left(\frac{x-y}{a_n}\right) f(y)\mathrm{d}y\mathrm{d}x} < \infty.$$