Statistics and Probability Letters 91 (2014) 32–40

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On wavelet projection kernels and the integrated squared error in density estimation

Evarist Giné *, W.R. Madych

Department of Mathematics, University of Connecticut, Storrs, CT 06269, United States

ARTICLE INFO

Article history:
Received 3 March 2014
Received in revised form 27 March 2014
Accepted 27 March 2014
Available online 8 April 2014

Keywords:
Wavelet projections
Integrated squared error
Linear density estimators

ABSTRACT

It is shown that the integrated squared errors of wavelet projection estimators of a density f satisfy both the central limit theorem and the law of the iterated logarithm under the essentially minimal assumption $f \in L^p$ for some $p > 2$ and very mild conditions on the scaling function.

1. Introduction

Arguably the two main types of linear density estimators are the convolution kernel estimators and the more recent wavelet projection estimators. The former has been extensively studied since its introduction in the 50's by Akaike, Parzen and Rosenblatt. In particular, for f bounded, the central limit theorem (CLT) and the law of the iterated logarithm (LIL) for the integrated squared deviation of these estimators from the true density were respectively obtained by Hall (1984) and by Giné and Mason (2004). The analogous results for wavelet projection estimators have been recently obtained respectively by Zhang and Zheng (1999) and by Lu (2013) under the assumption that the density f is bounded, Riemann integrable and eventually monotone. These hypotheses are not satisfactory for projection estimators as we now elaborate: given a closed subspace $V \in L^2(\mathbb{R})$ and letting π_V be the orthogonal projection onto V, the V-projection estimator of f is an unbiased estimator $f_n \in V$ of $\pi_V f$. Hence, since $\pi_V f - f$ is orthogonal to V, by the Pythagorean theorem,

$$\begin{aligned} I_n &:= \|f_n - f\|_2^2 - E\|f_n - f\|_2^2 \\ &= \|f_n - \pi_V f\|_2^2 + \|\pi_V f - f\|_2^2 - (E\|f_n - \pi_V f\|_2^2 + \|\pi_V f - f\|_2^2) \\ &= \|f_n - Ef_n\|_2^2 - E\|f_n - Ef_n\|_2^2 =: J_n, \end{aligned} \quad (1)$$

and it is typical of density estimation that smoothness conditions on f are only required in the estimation of the bias $Ef_n - f$ so that, since J_n does not include the bias, the CLT and the LIL for $I_n = J_n$ should not require any 'smoothness' conditions, in particular, Riemann integrability or eventual monotonicity.

It turns out that these extraneous conditions on f occur only at very precise places in the proofs of the CLT and the LIL: not in the many estimates that these proofs require, but only in the computation of the limiting variance. Zhang and Zheng (1999) proved a proposition to handle this limit computation to the effect that if $K(x, y)$ is the kernel of the orthogonal

* Corresponding author. Tel.: +1 860 486 1290.
E-mail addresses: gine@math.uconn.edu (E. Giné), madych@math.uconn.edu (W.R. Madych).

http://dx.doi.org/10.1016/j.spl.2014.03.029

projection of $L^2(\mathbb{R})$ onto the closed subspace $V_0 \subset L^2(\mathbb{R})$ spanned by the translates of a scaling function of a multiresolution analysis of L^2 satisfying some natural properties, and if f is bounded, Riemann integrable and eventually monotone, then

$$\lim_{N\to\infty} N \int\int K^2(Nx, Ny)f(x)f(y)dxdy = \int f^2(x)dx. \tag{2}$$

Since the union of the nested spaces V_j that define the multiresolution analysis is dense in L^2, clearly one has the limit (2) with K replacing K^2, and it is an interesting observation that K^2 satisfies this limit as well. Again, the problem with this result is that the hypothesis on f is unnatural. Lu's (2013) proof used similar arguments to compute the limiting variance in a more complex situation.

Also, Giné and Mason (2004) proved the LIL for the integrated squared deviation of convolution kernel estimators from their expected value (not from their true density) assuming only that $f \in L^p(\mathbb{R})$ for some $p > 2$. Thus, again because of (1), the assumption of boundedness for f in the wavelet projection case seems to be too strong both for the CLT and for the LIL.

The object of this note is to obtain the CLT and the LIL for the integrated squared error of wavelet projection density estimators for all f such that $\int f^p < \infty$ for some $p > 2$. Here is an example for each $p > 2$ of a density f that is neither bounded nor Riemann integrable nor eventually monotone: for $0 < \alpha p < 1$, and $\lambda_k > 0$ such that $\sum_{k=1}^{\infty}\lambda_k = (1-\alpha)/2$, take f to be $f(x) = \sum_{k=1}^{\infty} \frac{\lambda_k}{|x-2k|^{\alpha}} I_{0<|x-2k|<1}$.

A portion of our development consists of showing that both (2) and a slightly more complicated limit needed for the LIL are consequences of a general variant of the Riemann–Lebesgue lemma applied to the kernel $K(x,x)-1$. This allows us to remove the regularity assumptions on f.

In order to replace the boundedness assumption on f by $f \in L^p$ for some $p > 2$, we use both Young's inequality for convolutions and a substitute of it for $\int (|K| * f)^q(x)(1+|x|)^t dx$, $0 < q < 1$, $t \geq 0$.

2. A property of the diagonals of wavelet projection kernels

Let ϕ be a scaling function of a multiresolution analysis of $L^2(\mathbb{R})$ (*MRA* for short) satisfying the following properties:

(a) $\int_{\mathbb{R}} \phi(x)dx = 1$ and $\sum_{k\in\mathbb{Z}} \phi(x-k) = 1$, and

(b) ϕ is rapidly decaying at $\pm\infty$, that is, $|x|^m|\phi(x)| \to 0$ as $|x| \to \infty$ for all $m > 0$.

All the usual scaling functions satisfy these two properties, for instance $\phi = I_{(0,1]}$ (generating the Haar system), the Daubechies scaling functions (which are bounded and compactly supported), the scaling functions ϕ of band limited wavelets if e.g. ϕ is in Schwartz space, and the Battle–Lemarié or spline wavelets (for which ϕ has exponential decay, see e.g. Daubechies (1992), Corollary 5.4.2). So, we say, for short, that a multiresolution analysis is standard if it admits a scaling function ϕ satisfying properties (a) and (b), and in this case we also call such a function ϕ standard.

Let $K(x,y) = \sum_{k\in\mathbb{Z}} \phi(x-k)\phi(y-k)$ denote the (orthogonal) projection kernel onto the closed subspace V_0 of $L^2(\mathbb{R})$ generated by the translates $\phi(x-k)$ of ϕ, $k \in \mathbb{Z}$. For convenience and further reference, here we recall that if ϕ is standard, then the projection kernel K has the following two properties:

(c) $\sum_{k\in\mathbb{Z}} |\phi(x-k)| \in L^{\infty}(\mathbb{R})$, and $\int_{\mathbb{R}} K(x, x-y)dy = 1$ for all x, and

(d) for every $R > 0$ there exists $C_R < \infty$ such that

$$|K(x,y)| \leq \frac{C_R}{(1+|x-y|)^R} =: \Phi^{\{R\}}(x-y), \quad x, y \in \mathbb{R}. \tag{3}$$

See Härdle et al. (1998), Section 8.6 or Hernandez and Weiss (1996), Lemma 3.12, p. 220, where they show that there exists $c < \infty$ such that if $|\phi(x)| \leq \tilde{\Phi}(|x|)$, $x \in \mathbb{R}$, for $\tilde{\Phi}$ non-increasing on $[0,\infty)$, then $|K(x,y)| \leq c\tilde{\Phi}(|x-y|/2)$, $x, y \in \mathbb{R}$.

Usually, we will drop the superindex R from Φ, and will take R large enough so that Φ enjoys enough integrability for proofs to work (in most situations, $R = 2$ suffices).

As it will become apparent below, the limit in (2) reduces to showing the existence of the limit

$$\lim_{N\to\infty} \int_{\mathbb{R}} K(Nx, Nx)F(x)dx, \quad F \in L^1(\mathbb{R}), \tag{4}$$

and computing it. This is done in the following general lemma about periodic functions and in the subsequent theorem.

Lemma 1. *Suppose $P(x)$ is a measurable bounded periodic function whose integral over a period is 0. If $F(x)$ is in $L^1(\mathbb{R})$ then*

$$\lim_{\lambda\to\infty} \int_{-\infty}^{\infty} P(\lambda x)F(x)dx = 0.$$

Proof. Assume without loss of generality that P has period 1. The proof consists in showing that the lemma holds for a convenient subspace of functions dense in L^1 and then extending it to L^1 by continuity. Two typical sets of such functions