Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 65(2) 2011 169-179

On Linear Wavelet Density Estimation: Some Recent Developments

Yogendra P. Chaubey[1], Christophe Chesneau[2] and Hassan Doosti[3*]
[1]*Concordia University, Montréal, Québec, Canada*
[2]*LMNO, Université de Caen Basse-Normandie, France*
[3]*Tarbiat Moallem University, Tehran, Iran*

Received 14 Setember 2010; Revised 25 October 2010; Accepted 27 October 2010

SUMMARY

The theory of wavelets has found wide applications in nonparametric estimation, especially for density and related functionals. It has been adapted to many other situations in addition to density estimation for iid data. Such procedures may potentially be useful for nonparametric density estimation in agricultural setting such as in modeling yield of crops and crop insurance claims distribution. This article presents some recent developments in this area in a comprehensive way dealing with different data types in addition to the iid setup.

Keywords : Biased data, Censored data, Components of a mixture, Decon-volution model, Density estimation, Multiplicative censoring, Wavelets.

1. INTRODUCTION

In many statistical applications it is assumed that the underlying random variable X of interest is absolutely continuous with a distribution function $F(x)$ that admits a probability density function (pdf) $f(x)$, $x \in \mathbb{R}$. As is well known that various population characteristics can be obtained from the pdf, it is therefore of interest to estimate the underlying pdf.

The standard approach of density estimation, known as the parametric approach assumes that this pdf belongs to some family characterized by a set of parameters. Estimates of these parameters naturally provide a plug-in estimator of the pdf. However, if the family of the underlying distributions can not be very well established, one requires non-parametric estimation of the underlying density. Such inferences are useful in many applied fields including that of agricultural economics and crop insurance. Chaubey and Dewan (2010) cite examples of density estimation in the context of cotton yield and maize production.

Kernel density estimators developed around 1960's (Rosenblatt 1956 and Parzen 1962) have been very popular in the statistical literature and a plethora of articles dealing with various aspects of the technique and issues are now available in the textbooks, (see Prakasa Rao 1983, Silverman 1986, Wand and Jones 1995) with along with developed soft wares (see Härdle 1995). In the last 20 years, the subject of nonparametric density estimation is enriched by considerable mathematical advance in theory of wavelets that provide an orthogonal expansion of functions that may be continuous or with jump discontinuities. This representation is especially useful for density estimation due to the fact that the coefficients in the expansion are expectations of certain known functions given in terms of the wavelets and hence they can be estimated easily by the corresponding sample averages. Furthermore, with the availability of the new data these coefficients can be easily updated without knowledge of the complete data set. This

*Corresponding author : Hassan Doosti
E-mail address : hassandoosti1353@yahoo.com

feature is very useful for large data sets. An additional advantage of the wavelet methodology is its applicability for the data sets that are not necessarily generated as a sequence of independent and identically distributed (*iid*) random variables. It has been demonstrated by Walter and Blum (1979) and rigorously proved by Terrel and Scott (1992) that virtually all non-parametric density estimation algorithms are asymptotically kernel methods and it can be shown that it holds exactly true for a special class of wavelet estimators called linear wavelet density estimators. This feature makes the linear wavelet density estimators easily amenable to the methods of kernel estimators for studying its properties. Walter and Ghorai (1992) discuss the advantages and disadvantages of the method of wavelet density estimation. The reader may refer to Härdle *et al*. (1998) and Vidakovic (1999) for a detailed coverage of wavelet theory in statistics and to Prakasa Rao (1999b) for a recent comprehensive review and application of these and other methods of nonparametric functional estimation.

The objective of this article is to describe some recent developments in the theory of linear wavelet density estimators. We present some technical details in Section 2 for making the article self contained, but this section could be quickly browsed before the next sections. Section 3 gives the basic form of the linear density estimator for the iid data and Section 4 discusses its adaptation for the censored data. We discuss the applicability of this new technology for data under multiplicative censoring in Section 5 and for biased (or weighted) data in Section 6. Section 7 is devoted to linear wavelet density estimation for a component density from a mixture of densities. An unattractive feature of the wavelet method is in producing possibly negative estimate of the density in some regions. This problem has been addressed by a few authors that has been elaborated in Section 8. In Section 9 we introduce linear wavelet estimator in a density convolution model.

2. PRELIMINARIES ON WAVELETS AND BESOV SPACES

Here we provide a brief introduction to the wavelet system and Besov spaces that have become essential to the statistical literature. For the details of the theory and applications of wavelets, the reader may refer to the excellent text by Vidakovic (1999) or to the excellent survey by Antoniadis *et al*. (1994). For properties of the Besov spaces, the reader is referred to Meyer (1992) and Triebel (1992) (cf. Leblanc 1996 and Härdle *et al*. 1998).

2.1 Wavelet System

A wavelet system is composed of an infinite collection of functions that are obtained by dilation and translation of two basic functions ϕ and ψ called the *scaling function and mother wavelet*, respectively. The function ϕ is assumed to satisfy

$$\int_{-\infty}^{\infty} \phi(x)dx = 1$$

and is obtained as the solution from the equation

$$\phi(x) = \sum_{k\in\mathbb{Z}} C_k\, \phi(2x-k),$$

for a given sequence of constants $\{C_k\}$, and the function ψ is given by

$$\psi(x) = \sum_{k\in\mathbb{Z}} (-1)^k C_{-k+1}\, \phi(2x-k).$$

Define

$$\phi_{j,k}(x) = 2^{j/2}\, \phi(2^j x - k),\ (j, k) \in \mathbb{Z}^2$$

and

$$\psi_{j,k}(x) = 2^{j/2}\, \psi(2^j x - k),\ (j, k) \in \mathbb{Z}^2.$$

Suppose that the coefficients $\{C_k\}$ satisfy

$$\sum_{k\in\mathbb{Z}} C_k C_{k+2l} = \begin{cases} 2 & \text{if} \quad l = 0 \\ 0 & \text{if} \quad l \neq 0. \end{cases}$$

It is known (*cf*. Daubechies 1992) that under some additional conditions on ϕ, the collection $\{\psi_{j,k}; (j, k) \in \mathbb{Z}^2\}$ forms an orthonormal basis for $L^2(\mathbb{R})$, and $\{\phi_{j_0,k}; k \in \mathbb{Z}\}$ constitute an orthonormal basis for V_{j_0}, for every fixed $j_0 \in \mathbb{Z}$ as well.

Definition 2.1. The scaling function ϕ is said to be r-regular for an integer $r \geq 1$, if for every nonnegative integer $l \leq r$, the l-th derivative of ϕ, denoted $\phi^{(l)}$, is such that, for any integer $k \geq 1$,

$$|\phi^{(l)}(x)| \leq c_k(1 + |x|)^{-k},$$

for some $c_k \geq 0$ depending only on k.

Definition 2.2. A multiresolution analysis of $L^2(\mathbb{R})$ consists of an increasing sequence of closed spaces $\{V_j\}$ of $L^2(\mathbb{R})$ such that

(i) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$;

(iii) there is a scaling function $\phi \in V_0$ such that $\{\phi(x-k);\ k \in \mathbb{Z}\}$ forms an orthonormal basis for V_0;

(iv) for all $k \in \mathbb{Z}$ and all $h \in L^2(\mathbb{R})$, $h(x) \in V_0 \Rightarrow h(x-k) \in V_0$;

(v) $h(x) \in V_j \Rightarrow h(2x) \in V_{j+1}$.

Mallat (1989) has connected the multiresolution analysis to wavelet theory by showing that given any multiresolution analysis, it is possible to construct a function ψ, (called the mother wavelet), such that for any fixed $j \in \mathbb{Z}$ the family $\{\psi_{j,k};\ k \in \mathbb{Z}\}$ constitutes an orthonormal basis of the orthogonal complement W_j of V_j in V_{j+1} so that $\{\psi_{j,k};\ (j,k) \in \mathbb{Z}^2\}$ is an orthonormal basis of $L^2(\mathbb{Z})$ (*cf.* Daubechies 1992).

The corresponding multiresolution analysis is said to be r-regular if the scaling function ϕ is so. Suppose that both the functions ϕ and ψ belong the space of functions with r continuous derivatives denoted by C^r, for some $r \geq 1$, and have compact supports included in $[-N, N]$, for some $N > 0$. It follows, from Corollary 5.5.2 in Daubechies (1988), that the mother wavelet ψ is orthogonal to polynomials of degree $\leq r$, *i.e.*, for any $l \in \{0,..., r\}$

$$\int_{-\infty}^{\infty} x^l \psi(x)\, dx = 0.$$

Using such wavelets, any function $f \in L^2(\mathbb{R})$ can be expanded in the form

$$f(x) = \sum_{k \in \mathbb{Z}} \alpha_{j0,k}\phi_{j0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} \beta_{j,k}\psi_{j,k}(x)$$

for any $j_0 \in \mathbb{Z}$. The so called wavelet coefficients $\alpha_{j,k}$ and $\beta_{j,k}$ are given by

$$\alpha_{j,k} = \int_{-\infty}^{\infty} f(x)\phi_{j,k}(x)dx$$

and

$$\beta_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx$$

respectively (cf. Daubechies 1992).

Note that any $f \in L^2(\mathbb{R})$ satisfying $\mathrm{supp}(f) \subset [-L, L]$ with $L > 0$ can be expanded in the form

$$f(x) = \sum_{k \in K_{j0}} \alpha_{j0,k}\phi_{j0,k}(x) + \sum_{j \geq j_0} \sum_{k \in K_j} \beta_{j,k}\psi_{j,k}(x),$$

where $K_j = K_j(N, L)$ is a set of consecutive integers with a length proportional to 2^j. This wavelet decomposition (and the associated notations) will be consider in our statistical results.

2.2 Besov Spaces

Besov spaces are normed spaces defined for weakly-differentiable functions belonging to $L^2(\mathbb{R})$. We present the following definition of a weakly differentiable function f from Härdle *et al*. (1998).

Definition 2.3. Let $f \in L^2(\mathbb{R})$ be an integrable function on every bounded interval. It is said to be weakly differentiable if there exists a function g defined on the real line which is integrable on every bounded interval such that

$$\int_x^y g(u)du = f(y) - f(x).$$

The function g is defined almost everywhere and is called the weak derivative of f.

Definition 2.4. Let $p \geq 1$ and $m \geq 0$ be an integer. A function $f \in L^p(\mathbb{R})$ belongs to the Sobolev space $W_p^m(\mathbb{R})$, if it is m-times weakly-differentiable and the m-th weak derivative $f^{(m)} \in L^p(\mathbb{R})$. The space $W_p^m(\mathbb{R})$ is equipped with the norm $\|f\|_{W_p^m}$ defined by

$$\|f\|_{W_p^m} = \left(\sum_{u=0}^{m} \|f^{(u)}\|_p^p \right)^{1/p}$$

where $\|f\|_p$ denotes the norm for $L^p(\mathbb{R})$.

Let $f \in L^p(\mathbb{R})$ for some $p \geq 1$. Let $\Delta_h f(x) = f(x+h) - f(x)$ and define $\Delta_h^2 f = \Delta_h(\Delta_h f)$. For $t \geq 0$, let

$$w_p^1(f, t) = \sup_{|h| \leq t} \|\Delta_h f\|_p$$

and

$$w_p^2(f, t) = \sup_{|h| \leq t} \|\Delta_h^2 f\|_p.$$

Let $q \geq 1$ and ε be a function on $[0, \infty)$ and define $\|\varepsilon\|_q^*$ by

$$\|\varepsilon\|_q^* = \begin{cases} \left(\int_0^\infty t^{-1} |\varepsilon(t)|^q\, dt \right)^{1/q} & \text{if } 1 \leq q < \infty \\ \mathrm{essup}_t |\varepsilon(t)| & \text{if } q = \infty \end{cases}$$

whenever these quantities exist.

Definition 2.5. Let $p \geq 1$, $q \geq 1$ and $s = m + \alpha$ where $m \geq 0$ is an integer and $0 < \alpha \leq 1$. The Besov space $B_{p,q}^s$ is the space of all functions f such that $f \in W_p^m(\mathbb{R})$ and $w_p^2(f^{(m)}, t) = \varepsilon(t)t^\alpha$ where $\|\varepsilon\|_q^* < \infty$. The space $B_{p,q}^s$ is equipped with the norm $\|f\|_{B_{p,q}^s}$, called the Besov norm, defined by

$$\|f\|_{B_{p,q}^s} = \|f\|_{W_p^m} + \|\varepsilon\|_q^*$$

Suppose that f belongs to the Besov class

$$F^s_{p,q}(M, L) = \left\{f \,\varepsilon\, B^s_{p,q}, \|f\|_{B^s_{p,q}} \le M, \right.$$
$$\left. \text{supp}(f) \subset [-L, L]\right\}.$$

Let us consider the wavelet basis and notations in (2.1). Then, taking $j_0 = 0$, for any $0 \le s \le r + 1$, $p \ge 1$ and $q \ge 1$, the Besov norm of f can be written in terms of the wavelet coefficients

$$\|f\|_{B^s_{p,q}} = \|\alpha_{0,\cdot}\|_{\ell_p} + \left(\sum_{j\ge 0}(2^{j\sigma}\|\beta_{j,\cdot}\|_{\ell_p})^q\right)^{1/q}$$

where $\sigma = s + 1/2 - 1/p$ and $\|\gamma_{j,\cdot}\|_{\ell_p}$ represents the following norm for a double sequence $\{\gamma_{j,k}\}$

$$\|\gamma_{j,\cdot}\|_{\ell_p} = \left(\sum_{k\in K_j}|\gamma_{j,k}|^p\right)^{1/p}$$

(cf. Härdle *et al*. (1998), p. 123).

3. LINEAR WAVELET DENSITY ESTIMATION FOR i.i.d. DATA

We observe n iid random variables $X_1, ..., X_n$ with a common unknown pdf f. We want to estimate f from $X_1,..., X_n$. Doukhan and Leon (1990) and Kerkyacharian and Picard (1992) introduced the linear density estimator in terms of the projection of f on V_{j0} with j_0 depending on n, $j_0 \nearrow \infty$ when $n \to \infty$,

$$\hat{f}(x) = \sum_{k\in K_{j0}} \hat{\alpha}_{j0,k}\phi_{j0,k}(x), \qquad (3.1)$$

where

$$\hat{\alpha}_{j0,k} = \frac{1}{n}\sum_{i=1}^{n}\phi_{j0,k}(X_i).$$

Note that $\hat{\alpha}_{j0,k}$ is an unbiased estimator of $\alpha_{j0,k} = \int_{-\infty}^{\infty} f(x)\phi_{j0,k}(x)dx$

$$\mathbf{E}(\hat{\alpha}_{j0,k}) = \mathbf{E}(\phi_{j0,k}(X_1))$$
$$= \int_{-\infty}^{\infty} f(x)\phi_{j0,k}(x)dx = \alpha_{j0,k}.$$

The properties of the above linear estimator have been studied for a variety of error measures and density classes, see Leblanc (1996), Tribouley (1995), Varron (2008) and Giné and Nickl (2009). One of the main results about the linear estimator is following theorem which shows that the rate of convergence is optimal.

Theorem 3.1. Suppose that $f \in F^s_{p,q}(M, L)$ with $s > 1/p$, $p \ge 2$ and $q \ge 1$. Let $\hat{f}$ be (3.1) with j_0 satisfying $2^{j0} \approx n^{1/(1+2s)}$. Then there exists a constant $C > 0$ such that

$$\mathbf{E}\|\hat{f} - f\|_2^2 \le Cn^{-2s/(1+2s)}$$

Donoho and Johnstone (1995) showed that linear smoothing methods are incapable of achieving the optimal mean-square rate of convergence for curves whose smoothness is distributed inhomogeneously (*i.e*., when $1 \le p < 2$). In order to obtain a result like Theorem 3.1, one has to consider a non-linear wavelet estimator (hard thresholding, soft thresholding, . . .). There are some contributions in the case of dependency of observations. Leblanc (1996) obtained L^p-losses where $\{X_i\}$ are sequence of mixing random variables. Prakasa Rao (2003) and Doosti *et al*. (2006) derived the rate of convergence in the case of positively associated and negatively associate sequences, respectively. Doosti and Nezakati (2008) extended the results for m-dependent random variables.

4. LINEAR WAVELET DENSITY ESTIMATION UNDER RANDOM CENSORSHIP

We consider the random censorship model from the right, where two sequences of random variables, $\{X_i\}$ and $\{Y_i\}$, are considered. We regard $\{X_i\}$ as survival times (or failure times), having a common unknown distribution function F and pdf f. Let the survival times X_i be censored from the right by the censoring times Y_i, with a common distribution function G. We only observe the n pairs $(Z_1, \delta_1), . . ., (Z_n, . . .,\delta_n)$ where, for any $i \in \{1, . . ., n\}$,

$$Z_i = \min(Y_i, X_i),$$
$$\delta_i = I(X_i \le Y_i),$$

I denotes the indicator function. In this random censorship model, we assume that the survival times $\{X_i\}$ are independent of the censoring times $\{Y_i\}$. Following the convention in the survival analysis literature, we assume that, for any $i \in \{1, . . ., n\}$, both X_i and Y_i are nonnegative random variables.

The distribution functions F may be estimated by using the Kaplan-Meier estimator

$$\hat{F}_n(x) = 1 - \prod_{i=1}^{n}\left[1 - \frac{\delta_{(i)}}{n-i+1}\right]^{I(Z_{(i)}\le x)}$$

where $Z_{(1)} \leq Z_{(2)} \leq \ldots \leq Z_{(n)}$ denote the order statistics of $Z_1, \ldots, Z_n$, and $\delta_{(i)}$ is the concomitant of $Z_{(i)}$, i.e., $\delta_{(m)} = \delta_k$ if $Z_{(m)} = Z_k$. The Kaplan-Meier estimator of the censoring distribution may similarly be given by

$$\hat{G}_n(x) = 1 - \prod_{i=1}^{n} \left[1 - \frac{1-\delta_{(i)}}{n-i+1}\right]^{I(Z_{(i)} \leq x)}.$$

Note that $\delta_{(m)}/n(1 - \hat{G}_n(Z_{(m)}^-))$ is the jump of the Kaplan-Meier estimator $\hat{F}_n$ at Z_m. Here our interest is in estimating f based on $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$.

There is an extensive literature on the right censorship model with independent failure and censoring times. Density estimation was studied by Antoniadis *et al*. (1999), Li (2003) and Li *et al*. (2008). Let $T < \tau_H$ be a fixed constant, where $\tau_H = \inf\{x : H_{(x)} = 1\} \leq \infty$ is the least upper bound for the support of H, the distribution function of Z_1 and $f_1(x) = f(x)I(x \leq T)$. Here we estimate $f_1(x)$, for $x \in (-\infty, T)$, that in turn provides the estimate of $f(x)$ over the interval $(-\infty, T)$.

A wavelet based density estimator may be motivated from Li (2003) as given by

$$\hat{f}_1(x) = \sum_{k \in K_{j0}} \hat{\alpha}_{j0,k} \phi_{j0,k}(x) \tag{4.1}$$

where

$$\hat{\alpha}_{j0,k} = \int_{-\infty}^{\infty} \phi_{j0,k}(x) I(x \leq T)\, d\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i I(Z_i \leq T) \phi_{j0,k}(Z_i)}{1 - \hat{G}_n(Z_i^-)}$$

The following theorem is a result of Chaubey *et al*. (2010), when we consider $d = 0$.

Theorem 4.1. Suppose that $f_1 \in F_{p,q}^{s}(M, L)$ with $s > 1/p$, $p \geq 2$ and $q \geq 1$. Let $\hat{f}_1$ be (4.1) with j_0 satisfying $2^{j0} \approx n^{1/(1+2s)}$. Then there exists a constant $C > 0$ such that

$$\mathbf{E} \| \hat{f}_1 - f_1 \|_2^2 \leq C n^{-2s/(1+2s)}.$$

5. LINEAR WAVELET DENSITY ESTIMATION UNDER MULTIPLICATIVE CENSORING

We observe n *iid* random variables $Y_1, \ldots, Y_n$ with a common unknown pdf g supported on $(0, \infty)$. For any $i \in \{1, \ldots, n\}$, we know that

$$Y_i = U_i X_i,$$

where $U_1, \ldots, U_n$ are n *iid* unobserved random variables with uniform distribution on [0, 1] and $X_1, \ldots, X_n$ are n *iid* unobserved random variables with a common unknown pdf f supported on $(0, \infty)$. We want to estimate f from $Y_1, \ldots, Y_n$.

It is straightforward to derive that density g can be expressed in terms of f in the following manner

$$g(y) = \int_y^{\infty} \frac{f(x)}{x} dx,\ y \in (0, \infty).$$

Hence, the problem is seen to be a statistical inverse problem in the following sense: we observe $Y_1, \ldots, Y_n$ with pdf g which is related to another density f by $g = Kf$, where K is a linear operator. Vardi (1989) termed the above model as a multiplicative censoring model and showed how it unifies several well-studied statistical problems, including non-parametric inference for renewal processes, certain non-parametric deconvolution problems, and estimation of decreasing densities.

A natural attempt for estimating f would be to make an estimate $\hat{g}$ of g by standard methods and then if K^{-1} exists use $\hat{f} = K^{-1}\hat{g}$ as an estimate of f. For a review of these ideas see Andersen and Hansen (2001).

Different estimators may be obtained based on a Kernel method or a series expansion of the desired pdf f. For a review of these ideas see O'Sullivan (1986) and references therein. Another direction is to expand f based on a singular-value decomposition of K. In the statistics literature this was popularized by Johnstone and Silverman (1990, 1991). For recent contributions drawing on more general spectral theory for bounded operators, see Dey *et al*. (1996), Mair and Ruymgaart (1996) and Van Roij and Ruymgaart (1996). The most recent direction is that of applying wavelets as ba-sis functions in the reconstruction, see Abramovich and Silverman (1998) and Donoho *et al*. (1995).

Abbaszadeh *et al*. (2010) introduced the two estimators described below. We define the linear wavelet estimator $\hat{f}_1$ by

$$\hat{f}_1(x) = \sum_{k \in K_{j0}} \hat{\alpha}_{j0,k} \phi_{j0,k}(x), \tag{5.1}$$

where

$$\hat{\alpha}_{j0,k} = \frac{1}{n} \sum_{i=1}^{n} (\phi_{j0,k}(Y_i) + Y_i (\phi_{j0,k})'(Y_i)).$$

Note that $\hat{\alpha}_{j0,k}$ is an unbiased estimator of $\alpha_{j0,k} = \int_0^\infty f(x)\phi_{j0,k}(x)dx$. It follows from an integration by parts that

$$\begin{aligned}\mathbf{E}(\hat{\alpha}_{j0,k}) &= \mathbf{E}(\phi_{j0,k}(Y_1) + Y_1(\phi_{j0,k})'(Y_1)) \\ &= \int_0^\infty (\phi_{j0,k}(x) + x(\phi_{j0,k})'(x)g(x)dx \\ &= -\int_0^\infty xg'(x)\phi_{j0,k}(x)dx \\ &= \int_0^\infty f(x)\phi_{j0,k}(x)dx = \alpha_{j0,k}.\end{aligned}$$

We define the linear wavelet estimator $\hat{f}_2$ by

$$\hat{f}_2(x) = \frac{x}{n}\sum_{k\in K_{j0}}\sum_{i=1}^{n}(\phi_{j0,k})'(Y_i)\phi_{j0,k}(x). \quad (5.2)$$

They derived upper bound on the rate of convergence which provides a pseudo-consistency result for their estimators given in the following theorems.

Theorem 5.1. Suppose that $f \in F^s_{p,q}(M, L)$ with $s \geq 1/p$, $p \geq 1$ and $q \geq 1$. Let $\hat{f}_1$ be (5.1) with j_0 satisfying $2^{j0} \approx n^{1/(2(1+2s'))}$ where, for any $p' \geq \max(2, p)$, $s' = s + 1/p' + 1/p$. Then there exists a constant $C > 0$ such that

$$\mathbf{E}\|\hat{f}_1 - f\|^2_{p'} \leq Cn^{-\frac{2s'}{1+2s'}}.$$

Theorem 5.2. Suppose that $f' \in F^s_{p,q}$ (M, L) with $s \geq 1/p$, $p \geq 1$ and $q \geq 1$. Let $\hat{f}_2$ be (5.2) with j_0 satisfying $2^{j0} \approx n^{1/(1(1+2s'))}$ where, for any $p' \geq \max(2, p)$, $s' = s + 1/p' - 1/p$. Then there exists a constant $C > 0$ such that

$$\mathbf{E}\|\hat{f}_2 - f\|^2_{p'} \leq Cn^{-\frac{2(s'-1)}{1+2s'}}.$$

6. DENSITY ESTIMATION FOR A BIASED SAMPLE

We observe *n iid* random variables $Y_1, \ldots, Y_n$ with an unknown pdf g of the form

$$g(x) = \frac{w(x)\,f(x)}{\mu}, \; x \in R,$$

where w is a known positive function, f is an unknown pdf of a random variable X and $\mu = \int_{-\infty}^{\infty} w(x)f(x)dx$ is an unknown normalization parameter. We want to estimate f from $Y_1, \ldots, Y_n$.

This model has several applications in various domains such as biology, see Buckland *et al*. (1993), industry, see Cox (1969), and economics, see Heckman (1985). We may equally refer to the survey by Patil and Rao (1977) on several practical examples of biased distributions.

The density estimation problem for biased data has been considered in several papers. See Efromovich (2004), Brunel *et al*. (2009), Chesneau (2010c) and Ramirez and Vidakovic (2010).

Ramirez and Vidakovic (2010) proposed the linear wavelet estimator $\hat{f}$ defined by

$$\hat{f}(x) = \sum_{k\in K_{j0}} \hat{\alpha}_{j0,k}\phi_{j0,k}(x), \quad (6.1)$$

where

$$\hat{\alpha}_{j0,k} = \frac{\mu}{n}\sum_{i=1}^{n}\frac{\phi_{j0,k}(Y_i)}{w(Y_i)}.$$

Note that $\hat{\alpha}_{j0,k}$ is an unbiased estimator of $\alpha_{j0,k} = \int_{-\infty}^{\infty} f(x)\phi_{j0,k}(x)dx$

$$\begin{aligned}\mathbf{E}(\hat{\alpha}_{j0,k}) &= \mathbf{E}\left(\mu\frac{\phi_{j0,k}(Y_1)}{w(Y_1)}\right) \\ &= \int_{-\infty}^{\infty}\mu\frac{\phi_{j0,k}(x)}{w(x)}\frac{w(x)\,f(x)}{\mu}dx \\ &= \int_{-\infty}^{\infty} f(x)\phi_{j0,k}(x)dx = \alpha_{j0,k}.\end{aligned}$$

As the parameter µis not known, it is estimated by $\hat{\mu} = n\Big/\sum_{i=1}^{n}(1/w(Y_i))$. Ramirez and Vidakovic (2010) showed that the inverse of this estimator is unbiased for $1/\mu$. Doosti and Dewan (2010) investigated an upper bound on L_p-loss for the estimator given by Ramirez and Vidakovic (2010) which extends such a result for the L^2-consistency given in Ramirez and Vidakovic (2010). The following theorem is a version of the main result in Doosti and Dewan (2010).

Theorem 6.1. Suppose that $f \in F^s_{p,q}$ (M, L) with $s \geq 1/p$, $p \geq 1$ and $q \geq 1$, and there exists a constant $B > 0$ such that $w(x) \geq B$, $x \in \mathbb{R}$. Let $\hat{f}$ be (6.1) with j_0 satisfying $2^{j0} \approx n^{1/(1(1+2s'))}$ where, for any $p' \geq \max(2, p)$, $s' = s+1/p' - 1/p$. Then there exists a constant $C > 0$ such that

$$\mathbf{E}\|\hat{f} - f\|^2_{p'} \leq Cn^{-\frac{2s'}{1+2s'}}$$

Chaubey and Doosti (2010) proposed positive wavelets density estimator for biased sample and derived the formula for its IMSE.

7. DENSITY ESTIMATION OF A COMPONENT FROM MIXTURES

We observe n independent random variables $X_1, \ldots, X_n$ such that, for any $i \in \{1, \ldots, n\}$, X_i depends on a random indicator Y_i taking its values in $\{1, \ldots, m\}$. Applying the Bayes theorem, the pdf of X_i is

$$h_i(x) = \sum_{d=1}^{m} w_d(i) f_d(x),\ x \in \mathbb{R}, \tag{7.1}$$

where $w_d(i) = \mathbb{P}(Y_i = d)$ and, for any $d \in \{1, \ldots, m\}$, f_d is the conditional density of X_i given $\{Y_i = d\}$. We suppose that all these densities are unknown and weight $w_d(i)$ is known. For a fixed $\nu \in \{1, \ldots, m\}$, we aim to estimate f_ν from $X_1, \ldots, X_n$.

Such an estimation problem has several applications in various domains such as biology, industry and economics. It was firstly studied by Maiboroda (1996). The constructions of wavelet estimators (linear and non-linear) has been considered by Pokhyl'ko (2005). Let us now present the construction of the linear one.

First of all, we remark that there exist m real numbers $a_\nu(1), \ldots, a_\nu(m)$ such that, for any $d \in \{1, \ldots, m\}$,

$$\frac{1}{n}\sum_{d=1}^{m} a_\nu(i) w_d(i) = \begin{cases} 1 & \text{if } d = \nu, \\ 0 & \text{otherwise} \end{cases}$$

and

$$(a_\nu(1), \ldots, a_\nu(n)) = argmin_{(b_1, \ldots, b_n) \in \mathbb{R}^n} \frac{1}{n}\sum_{i=1}^{n} b_i^2 .$$

Pokhyl'ko (2005) developed the linear estimator $\hat{f}$ defined by

$$\hat{f}(x) = \sum_{k \in K_{j0}} \hat{\alpha}_{j0,k} \phi_{j0,k}(x), \tag{7.2}$$

where

$$\hat{\alpha}_{j0,k} = \frac{1}{n}\sum_{i=1}^{n} a_\nu(i) f_{j0,k}(X_i).$$

Note that $\hat{\alpha}_{j0,k}$ is an unbiased estimator of $\alpha_{j0,k} = \int_{-\infty}^{\infty} f_\nu(x)\phi_{j0,k}(x)dx$

$$\begin{aligned}
\mathbf{E}(\hat{\alpha}_{j0,k}) &= \frac{1}{n}\sum_{i=1}^{n} a_\nu(i)\mathbf{E}(\phi_{j0,k}(X_i)) \\
&= \frac{1}{n}\sum_{i=1}^{n} a_\nu \text{ (i)} \ldots \\
&\qquad \left(\sum_{d=1}^{m} w_d(i)\int_{-\infty}^{\infty} f_d(x)\phi_{j0,k}(x)dx\right) \\
&= \sum_{d=1}^{m}\int_{-\infty}^{\infty} f_d(x)\phi_{j0,k}(x)dx \ldots \\
&\qquad \left(\frac{1}{n}\sum_{i=1}^{n} a_\nu(i) w_d(i)\right) \\
&= \int_{-\infty}^{\infty} f_\nu(x)\phi_{j0,k}(x)dx = \alpha_{j0,k}.
\end{aligned}$$

Theorem 9.1 below is proved by Pokhyl'ko (2005).

Theorem 7.1. Suppose that $f_\nu \in F^s_{p,q}$(M, L) with $s > 1/p$, $p \geq 2$ and $q \geq 1$. Let $\hat{f}$ be (7.2) with j_0 satisfying $2^{j0} \approx (n/z_n)^{1/(1+2s)}$ and $z_n = (1/n)\sum_{i=1}^{n} a_\nu^2(i)$. Then there exists a constant $C > 0$ such that

$$\mathbf{E}\|\hat{f} - f_\nu\|_2^2 \leq C\left(\frac{z_n}{n}\right)^{2s/(1+2s)}$$

This result has been extended to the estimation of the derivatives of f_ν by Prakasa Rao (2010). Chesneau (2010) has investigated the estimation of f_ν by a linear wavelet estimator from pairwise positively quadrant dependent (PPQD) $X_1, \ldots, X_n$.

8. POSITIVE WAVELET DENSITY ESTIMATOR

Janssen (1994) showed that there are no continuous non-negative orthogonal scaling functions. Hence, the pdf may be negative in the tails. This is not an attractive feature of the usual wavelet method, hence modifications may be necessary.

There are two approaches. First we estimate a transformation of f and then the density function estimator is obtained by taking the inverse transformation. For instance, The log-transformation was introduced by Leonard (1973) and Clutton-Brock (1990). Penev and Dechevsky (1997) and Pinheiro and Vidakovic (1997) discuss estimation of the square root of a density. For further details see Chapter 7 of Vidakovic (1999). Walter and Shen (1999) introduced the second approach by proposing direct smooth non-negative wavelet estimator that is defined through the function ρ_α given by

$$\rho_\alpha(x) = \sum_{j\in\mathbb{Z}} \alpha^{|j|} f(x-j).$$

Walter and Shen (1999) showed that there exists $\alpha_0 > 0$, such that for $\alpha \in [\alpha_0, 1]$, $\rho_\alpha(x) > 0$ for all $x \in \mathbb{R}$. Non-negative wavelet kernels are defined as

$$K_\alpha(x, y) = \left(\frac{1-\alpha}{1+\alpha}\right)^2 \sum_{k\in\mathbb{Z}} \rho_\alpha(x-k)\rho_\alpha(y-k).$$

The dilation of $K_\alpha(x, y)$ is defined as

$$K_{\alpha,j}(x, y) = 2^j K_\alpha(2^j x, 2^j y).$$

We observe n *iid* random variables $X_1,...,X_n$ with a common unknown pdf f. Then a positive wavelet density estimate of f is defined as

$$\hat{f}_{j0}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{\alpha,j0}(x, X_i).$$

Ghorai (2003) derived the asymptotic distribution of integrated squared error of positive wavelet density estimator by using Martingale central limit theorem. Ghorai and Yu (2004) derived asymptotic formula for the IMSE of positive wavelet density estimator and they proved the consistency of the density estimator. Theorem 8.1 shows the formula for IMSE of the estimator.

Theorem 8.1. Assume f, f', f'' are continuous and f'' exists piecewise with well-defined left and right limits. Also assume f is monotone in the tails. Then

$$\begin{aligned} IMSE(\hat{f}_{j0}) &= \int_{-\infty}^{\infty} \mathbf{E}((\hat{f}_{j0}(x) - f(x))^2)dx \\ &= 2^{j0} B(\alpha, \rho) n^{-1} \\ &\quad + A(\alpha, \|f''\|_2^2)\, 2^{-4j_0} + o_{j_0}(2^{j_0} n^{-1}) \\ &\quad + O_{j_0}(2^{-5j0}), \end{aligned}$$

where $A(\alpha, \|f''\|_2^2) = 4\alpha^2(1-\alpha)^{-4} \|f''\|_2^2$ and $B(\alpha, \rho) = \frac{1-\alpha}{1+\alpha}\sum_l [|l| - 1 + 2\alpha(1-\alpha^2)^{-1}]^2 \alpha^{2|l|}$.

9. DENSITY ESTIMATION IN A DENSITY CONVOLUTION MODEL

We observe n *iid* random variables $Y_1, \ldots, Y_n$ such that, for any i $\{1, \ldots, n\}$,

$$Y_i = X_i + \epsilon_i,$$

where $X_1, \ldots, X_n$ are n unobserved *iid* random variables with a common un-known pdf f and $\varepsilon_1, \ldots, \varepsilon_n$ are n unobserved *iid* random variables with a common known pdf g. For any $i \in \{1, \ldots, n\}$, X_i and ε_i are independent. We want to estimate f from $Y_1, \ldots, Y_n$.

Methods and results can be found in Caroll and Hall (1988), Devroye (1989), Fan (1991), Pensky and Vidakovic (1999), Fan and Koo (2002), Butucea and Matias (2005), Comte *et al*. (2006), Delaigle and Gijbels (2006) and Lacour (2006). Wavelet estimators have been developed by Pensky and Vidakovic (1999) and Fan and Koo (2002).

Let us present the construction of the linear one developed by Fan and Koo (2002). We define the Fourier transform of a function h by

$$\mathcal{F}(h)(x) = \int_{-\infty}^{\infty} h(y)e^{-ixy}\, dy,$$

whenever this integral exists. The notation $\overline{f}$ will be used for the complex conjugate f.

We consider the ordinary smooth case on g: there exist two constants, $c_* > 0$ and $\delta > 1$, such that

$$|\mathcal{F}(g)(x)| \geq \frac{c_*}{(1+x^2)^{\delta/2}}.$$

This assumption controls the decay of the Fourier coefficients of g, and thus the smoothness of g.

Fan and Koo (2002) studied the linear estimator $\hat{f}$ defined by

$$\hat{f}(x) = \sum_{k\in K_{j0}} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \quad (9.1)$$

with

$$\hat{\alpha}_{j_0,k} = \frac{1}{2\pi n}\sum_{i=1}^{n}\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})(x)}}{\mathcal{F}(g)(x)} e^{-ixY_i} dx..$$

Note that $\hat{\alpha}_{j_0,k}$ is an unbiased estimator of $\alpha_{j0,k} = \int_{-\infty}^{\infty} f(x)\phi_{j0,k}(x)dx$. It follows using the equality $\mathbf{E}(e^{-ixY_1}) = \mathcal{F}(f)(x)\mathcal{F}(g)(x)$ and the Parseval-Plancherel theorem, we obtain

$$\mathbf{E}(\hat{\alpha}_{j_0,k}) = \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})(x)}}{\mathcal{F}(g)(x)} \mathbf{E}(e^{-ixY_i})\, dx$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j_0,k})\,(x)}}{\mathcal{F}(g)(x)} \mathcal{F}(f)\,(x)\,\mathcal{F}(g)(x)\,dx$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\mathcal{F}(\phi_{j_0,k})}\,(x)\mathcal{F}(f)\,(x)\,dx$$

$$= \int_{-\infty}^{\infty} f(x)\phi_{j0,\,k}(x)dx = \alpha_{j0,\,k}.$$

Theorem 9.1 below is proved in Fan and Koo (2002).

Theorem 9.1. Suppose that $f \in F^{s}_{p,q}(M, L)$ with $s > 1/p$, $p \geq 2$ and $q \geq 1$. Let $\hat{f}$ be (9.1) with j_0 satisfying $2^{j0} \approx n^{1/(1+2\delta+2s)}$. Then there exists a constant C > 0 such that

$$\mathbf{E}\|\hat{f} - f\|_2^2 \leq Cn^{-2s/(1+2\delta+2s)}.$$

This result has been extended to the deconvolution of a component from mixtures by Chesneau (2010b).

ACKNOWLEDGEMENTS

The main part of the paper has been done while the third author was visiting Concordia University, Montreal, Canada on an invitation from the Department of Mathematics and Statistics. He is very grateful to the Department for the financial support and thankful to the members of the Department for their hospitality.

REFERENCES

Abbaszadeh, M., Doosti, H. and Gachpazan, M. (2010). Density estimation by wavelets under multiplicative censoring, *Preprint*.

Abramovich, F. and Silverman, B.W. (1998). Wavelet decomposition approaches to statistical inverse problems. *Biometrika*, **85**, 115-129.

Andersen, K. and Hansen, M. (2001). Multiuplicative censoring: Density estimation by a series expansion approach. *J. Stat. Plann. Inf.*, **98**, 137-155.

Antoniadis, A., Grégoire, G. and McKeague, I. (1994). Wavelet methods for curve estimation. *J. Amer. Statist. Assoc.*, **89**, 1340-1353.

Antoniadis, A., Gregoire, G. and Nason, G. (1999). Density and hazard rate estimation for right-censored data by using wavelet methods. *J. Roy. Statist. Soc.*, **61**, 63-84.

Brunel, E., Comte, F. and Guilloux, A. (2009). Nonparametric density estimation in presence of bias and censoring. *Test*, **18(1)**, 166-194.

Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. (1993). *Distance Sampling: Estimating Abundance of Biological Populations.* Chapman and Hall, London.

Butucea, C. and Matias, C. (2005). Minimax estimation of the noise level and of the signal density in a semiparametric convolution model. *Bernoulli*, **11(2)**, 309-340.

Caroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, **83**, 1184-1186.

Chaubey, Y.P. and Dewan, I. (2010). Smooth estimation of survival and density functions for stationary associated sequences: Some recent developments. *J. Ind. Soc. Agril. Statist.*, **64**, 261-272.

Chaubey, Y.P. and Doosti, H. (2010). Positive wavelet linear density function estimation. *Preprint*, Department of Mathematics and Statistics, Concordia University, Montréal, Québec, Canada.

Chaubey, Y.P., Doosti, H., Shirazi, E. and Prakasa Rao, B.L.S (2010). Linear wavelet-based estimation for derivative of a density under random censorship. *J. Iran. Statist. Soc.*, **9**, *To appear*.

Chesneau, C. (2010a). Wavelet linear estimation of a density from observations of mixtures under quadrant dependence, *Preprint*, LMNO.

Chesneau, C. (2010b). Wavelet density estimators for the deconvolution of a component from a mixture, *Preprint*, LMNO.

Chesneau, C. (2010c). Wavelet block thresholding for density estimation in the presence of bias. *J. Korean Statist. Soc.*, **39**, 43-53.

Clutton-Brock, M. (1990). Density estimation using exponentials of orthogonal series. *J. Amer. Statist. Assoc.*, **85**, 760-764.

Comte, F., Rozenholc, Y. and Taupin, M.L. (2006). Penalized contrast estimator for density deconvolution. *The Can. J. Stat.*, **34**, 431-452.

Cox, D. (1969). Some sampling problems in technology. *In New Developments in Survey Sampling* (N. L. Johnson and H. Smith, Jr., eds.). Wiley, New York, 506-527.

Daubechies, I. (1988). Orthogonal bases of compactly supported wavelets. *Statistica Sinica*, **41**, 909- 996.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS-NSF regional conferences series in applied mathematics. SIAM, Philadelphia.

Delaigle, A. and Gijbels, I. (2006). Estimation of boundary and discontinuity points in deconvolution problems. *Statistica Sinica*, **16**, 773-788.

Devroye, L. (1989). Consistent deconvolution in density estimation. *Can. J. Stat*., **17**, 235-239.

Dey, A.K., Mair, B.A. and Ruymgaart, F.H. (1996). Cross-validation for parameter selection in inverse estimation. *Scand. J. Statist*., **23**, 609-620.

Donoho, D.L. and Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinking. *J. Amer. Statist. Assoc*., **90**, 1200-1224.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: asymptopia? (with discussion). *J. Roy. Statist. Soc*., **B57**, 301-369.

Doosti, H. and Dewan, I. (2010). Wavelet linear density estimation for associated stratified size-biased sample, *Preprint*, Indian Statistical Institute, Delhi Centre, New Delhi, India.

Doosti, H. and Nezakati, A. (2008). Wavelet linear density estimation for M-dependent random variables. JDUBS, **1(2)**, 51-55.

Doosti, H., Fakoor, V. and Chaubey, Y.P. (2006). Wavelet linear density estimation for negative associated sequences. *J. Ind. Statist. Assoc*., **44**, 127-135.

Doukhan, P. and Leon, J.R. (1990). Une note sur la dé viation quadratique déstimateurs de densités par projections orthogonales.*C.R. Acad. Sci. Paris t310*, série, 1, 425-430.

Efromovich, S. (2004). Density estimation for biased data. *J. Statist. Plann. Inf.*, **32(3)**, 1137-1161.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist*., **19**, 1257-1272.

Fan, J. and Koo, J.Y. (2002). Wavelet deconvolution. *IEEE Trans. Inf. Theo*., **48**, 734-747.

Giné, E. and Nickl, R. (2009). Uniform limit theorems for wavelet density estimators. *Ann. Prob*., **37**, 1607-1646.

Ghorai, J.K. (2003). A central limit theorem for the l_2 error of positive wavelet density estimator. *Ann. Inst. Statist. Math*., **55**, 619-637.

Ghorai, J.K. and Yu, D. (2004). Data-based resolution selection in positive wavelet density estimation. *Comm. Statist.-Theory. Methods*., **33(10)**, 2393-2408.

Härdle, W. (1990). *Smoothing Techniques: with Implementation in S*. Springer, New York.

Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, T. (1998). *Wavelets, Approximations, and Statistical Applications*. Lecture Notes in Statistics, **129**, Springer, New York.

Heckman, J. (1985). Selection bias and self-selection. In: *The new Palgrave : A Dictionary of Economics*, 287-296. MacMillan Press, Stockton, New York.

Janssen, A.J.E.M. (1994). The Smith -Barnwell condition and non-negative scaling functions. *IEEE Trans. Info. Theo*., **38**, 884-885.

Johnstone, I.M. and Silverman, B.W. (1990). Speed of estimation in positron emission tomograghy and related inverse problems. *Ann. Statist*., **18**, 251-280.

Johnstone, I.M. and Silverman, B.W. (1991). Discretization effects in statistical inverse problems. *J. Complexity*, **7**, 1-34.

Kerkyacharian, G. and Picard, D. (1992). Density estimation in Besove spaces. *Statist. Prob. Lett*., **13**, 15-24.

Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Acad. Sci. Paris Ser. I Math*., **342(11)**, 877-882.

Leblanc, F. (1996). Wavelet linear density estimator for a discrete-time stochastic process: L_p-losses. *Statist. Prob. Lett*., **27**, 71-84.

Leonard, T. (1973). A Bayesian method for histograms. *Biometrika*, **60**, 297-308.

Li, L. (2003). Non-linear wavelet-based density estimators under random censorship. *J. Statist. Plann. Inf*., **117**, 35-58.

Li, L., MacGibbon, B. and Valenta, C. (2008). On the optimality of wavelet-based nonparametric regression with censored data. *J. Appl. Prob. Statist.,* **3**, 243-261.

Maiboroda, R.E. (1996). Estimators of components of a mixture with varying concentrations. *Ukrain. Mat. Zh*., **48(4)**, 562-566.

Mair, B.A. and Ruymgaart, E.H. (1996). Statistical inverse estimation in Hilbert scales. *SIAM J. Appl. Math*., **56**, 1424-1444.

Mallat, S.G. (1989). A theory for multiresoution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intel*., **11**, 674-693.

Meyer, Y. (1992). *Wavelets and Operators*. Cambridge University Press, Cambridge.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems (with discussion). *Statist. Sci*., bf **1** 502-527.

Parzen, E. (1962). On estimation of probability density and mode. *Ann. Math. Stat*., **33**, 1065-1070.

Patil, G.P. and Rao, C.R. (1977). The weighted distributions: A survey of their applications. In: *Applications of Statistics (P.R. Krishnaiah, ed.)* 383-405. North-Holland, Amsterdam.

Penev, S. and Dechevsky, L. (1997). On non-negative wavelet-based density estimators. *J. Nonpar. Stat*., **7**, 365-394.

Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for non-parametric density deconvolution. *Ann. Stat*., **27**, 2033-2053.

Pinheiro, A. and Vidakovic, B. (1997). Estimating the square root of a density via compactly supported wavelets. *Comput. Statist. Data Anal.*, **25**, 399-415.

Pokhyl'ko, D. (2005). Wavelet estimators of a density constructed from observations of a mixture. *Theor. Prob. Math. Statist*., **70**, 135-145.

Prakasa Rao, B.L.S. (1983). *Non-parametric Functional Estimation*. Academic Press, New York.

Prakasa Rao, B.L.S. (1999a). Estimation of the integrated squared density derivative by wavelets. *Bull. Inform. Cyb*., **31**, 47-65.

Prakasa Rao, B.L.S. (1999b). Nonparametric functional estimation: An overview. In: *Asymptotics, Nonparametrics and Time Series*, Ed. Subir Ghosh, Marcel Dekker Inc., New York, 461-509.

Prakasa Rao, B.L.S. (2003). Wavelet linear density estimation for associated sequences. *J. Ind. Stat. Assoc*., **41**, 369-379.

Prakasa Rao, B.L.S. (2010). Wavelet linear estimation for derivatives of a density from observations of mixtures with varying mixing proportions. *Ind. J. Pure Appl. Math*., **41**, 275-291.

Ramirez, P. and Vidakovic, B. (2010). Wavelet density estimation for stratified size-biased sample. *J. Statist. Plann. Inf*., **140**, 419-432.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of density functions. *Ann. Math. Stat*., **27**, 832-837.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Terrel, G.R. and Scott, D.W. (1992). Variable kernel density estimation. *Ann. Stat*., **20**, 1236-1265.

Tribel, H. (1992). *Theory of Function Space II*, Birkhaüser Verlag, Berlin.

Tribouley, K. (1995). Density estimation by cross-validation wavelet method. *Statistica Neerlandica*, **45**, 41-62.

Van Roij, A., Ruymgaart, F. (1996). Asymptotic minimax rates for abstract linear estimators. *J. Statist. Plann. Inf*., **53**, 389-402.

Vardi, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: Non-parametric estimation. *Biometrika*, **76**, 751-761.

Varron, D. (2008). Some asymptotic results on density estimators by wavelet projections. *Stat. Prob. Lett*., **78**, 2517-2521.

Vidakovic, A. (1999). *Statistical Modelling by Wavelets, Springer*, New York.

Walter, G. and Blum, J. (1979). Probability density estimation using delta sequences. *Ann. Stat*., **7**, 328-340.

Walter, G. and Ghorai, J. (1992). Advantages and disadvantages of density estimation with wavelets. In: *Proceedings of the 24th Symp. on the Interface*, Ed. H. Joseph Newton, In-terface FNA, VA, **24**, 234-343.

Walter, G.G. and Shen, X. (1999). Continuous non-negative wavelets and their use in density estimation. *Commu. Statist.-Theory Methods*, **28(1)**, 1-17.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. Chapman and Hall, New York.