

A Note on Central Limit Theorem for Deconvolution Wavelet Density Estimator¹⁾

Sungho Lee²⁾

Abstract

The problem of wavelet density estimation based on Shannon's wavelets is studied when the sample observations are contaminated with random noise. In this paper we will discuss the asymptotic normality for deconvolving wavelet density estimator of the unknown density $f(x)$ when Fourier transform of random noise has polynomial descent.

Keywords : Central limit theorem, deconvolution, Shannon's wavelets, polynomial descent.

1. Introduction

Let X and Z be independent random variables with density functions $f(x)$ and $q(z)$, respectively, where $f(x)$ is unknown and $q(z)$ is known. One observes a sample of random variables $Y_i = X_i + Z_i$, $i=1, 2, \dots, n$. The objective is to estimate the density function $f(x)$ where $g(y)$ is the convolution of $f(x)$ and $q(z)$, $g(y) = \int_{-\infty}^{\infty} f(y-z) q(z) dz$.

The problem of measurements being contaminated with noise exists in many different fields(see, for example, Louis(1991), Zhang(1992)). The most popular approach to the problem was to estimate $f(x)$ by a kernel estimator and Fourier transform (see, for example, Carroll and Hall (1988), Taylor and Zhang(1990), Fan(1991)). Fan(1991) proved that the estimators of $f(x)$ are asymptotically optimal pointwise and globally if the Fourier transform of the kernel has bounded support.

The present paper deals with estimation of a deconvolution density using a wavelet decomposition. The underlying idea is to present $f(x)$ via a wavelet expansion and then to estimate the coefficients using a deconvolution algorithm. Wavelet methods, introduced to statistics by the work of Donoho and Johnstone in early 90's, show remarkable potential in nonparametric function estimation(see, for example, Donoho, Johnstone, Kerkyacharian and

1) This research was supported(in part) by the Taegu University Research Grant, 2001.

2) Professor, Department of Statistics, Taegu University, Kyungsan, 712-714, Korea
E-mail : shlee1@biho.taegu.ac.kr

Picard(1995,1996)). There are several important families of wavelets(for example, Haar's wavelets, Shannon's wavelets, Meyer's wavelets, Daubechies' compactly supported wavelets). In this work we consider a wavelet decomposition based on Shannon's wavelets rather than on wavelets with bounded support. Shannon's wavelets and Meyer's wavelets allow immediate deconvolution and form a subset of the set of band-limited wavelets, that is, the Fourier transform of the wavelet has bounded support. Pensky and Vidakovic(1999) proposed the estimators based on Meyer-type wavelets to estimate $f(x)$ for two different cases in the well-known Sobolev space H^α : the case when the distribution of the error Z is supersmooth, that is, the Fourier transform \tilde{q} of q has exponential descent, and the case when the distribution of the error Z is ordinary smooth, that is, \tilde{q} has polynomial descent. They showed that, in the case of exponential descent, the linear wavelet estimator (2.7) in Section 2 is asymptotically optimal in the sense that the rate of convergence of the mean integrated squared error can't be improved. Lee(2001) showed that the linear wavelet estimator (2.7) is L_1 strongly consistent when \tilde{q} has polynomial descent or exponential descent. Lee and Hong(2002) showed that the linear wavelet estimator (2.7) is a uniformly, strongly consistent estimator of $f(x) \in H^\alpha, \alpha > 0$ when \tilde{q} has polynomial descent or exponential descent.

In deconvolution density estimation Fan(1991) discussed the asymptotic normality of the estimator, constructed by kernel and Fourier transform, by assuming either the distribution of the error Z is ordinary smooth or supersmooth. In this paper we will discuss the asymptotic normality for deconvolving the linear wavelet density estimator (2.7) of the unknown density $f(x)$ when Fourier transform $\tilde{q}(\xi)$ of $q(z)$ has polynomial descent. This result can be considered as the wavelet counterpart to the Fan's ordinary smooth case(1991b) for the kernel density estimator. Gamma or double exponential distribution functions satisfy polynomial descent

2. Preliminaries

Throughout this paper we use the notation $\hat{f}(\omega)$ for the Fourier transform $\int_{-\infty}^{\infty} e^{-i\omega x} f(x) dx$ of a function $f(x)$. We assumed that the reader is familiar with the elements of wavelet theory(see, for example, Vidakovic(1999)). Assume that $f(x)$ is square integrable and that $\tilde{q}(\omega)$ does not vanish for real ξ . If $\varphi(x)$ and $\psi(x)$, respectively, are a scaling function and a wavelet generated by an orthonormal multiresolution decomposition of $L^2(-\infty, \infty)$, then for any integer m the density function $f(x)$ allows the following representation:

$$f(x) = \sum_{k \in \mathbb{Z}} a_{m,k} \varphi_{m,k}(x) + \sum_{k \in \mathbb{Z}} \sum_{j=m}^{\infty} b_{j,k} \psi_{j,k}(x) , \tag{2.1}$$

where $\varphi_{m,k}(x) = 2^{m/2} \varphi(2^m x - k)$ and $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$, and the coefficients $a_{m,k}$ and $b_{j,k}$ have the forms

$$a_{m,k} = \int_{-\infty}^{\infty} \varphi_{m,k}(x) f(x) dx , \quad b_{j,k} = \int_{-\infty}^{\infty} \psi_{j,k}(x) f(x) dx$$

respectively.

A special class of wavelets are band-limited wavelets, the Fourier transform of which have bounded support. In this paper, we shall use a particular type of band-limited wavelet, Shannon's wavelets (see Walter(1994)). The Shannon scaling function is

$$\varphi(x) = \frac{\sin \pi x}{\pi x} \quad \text{and} \quad \tilde{\varphi}(\omega) = I_{[-\pi, \pi]}(\omega). \tag{2.2}$$

A possible wavelet is given by

$$\psi(x) = \frac{\sin \pi(x-0.5) - \sin 2\pi(x-0.5)}{\pi(x-0.5)} .$$

The coefficients $a_{m,k}$ and $b_{j,k}$ can be viewed as mathematical expectations of the functions $u_{m,k}$ and $v_{j,k}$

$$a_{m,k} = \int_{-\infty}^{\infty} u_{m,k}(y) g(y) dy , \quad b_{j,k} = \int_{-\infty}^{\infty} v_{j,k}(y) g(y) dy , \tag{2.3}$$

provided that $u_{m,k}(y)$ and $v_{j,k}(y)$ are solutions of the following equations:

$$\int_{-\infty}^{\infty} q(y-x) u_{m,k}(y) dy = \varphi_{m,k}(x) , \quad \int_{-\infty}^{\infty} q(y-x) v_{j,k}(y) dy = \psi_{j,k}(x). \tag{2.4}$$

Taking the Fourier transform of both sides in (2.4), we obtain $u_{m,k}(x) = 2^{m/2} U_m(2^m x - k)$, $v_{j,k}(x) = 2^{j/2} V_j(2^j x - k)$, where $U_m(\cdot)$ and $V_j(\cdot)$ are the inverse Fourier transforms of the functions

$$\tilde{U}_m(\omega) = \tilde{\varphi}(\omega) / \tilde{q}(-2^m \omega) , \quad \tilde{V}_j(\omega) = \tilde{\psi}(\omega) / \tilde{q}(-2^j \omega) , \tag{2.5}$$

respectively. Therefore, estimating $a_{m,k}$ and $b_{j,k}$ by

$$\hat{a}_{m,k} = n^{-1} \sum_{i=1}^n 2^{m/2} U_m(2^m Y_i - k) , \quad \hat{b}_{j,k} = n^{-1} \sum_{i=1}^n 2^{j/2} V_j(2^j Y_i - k) \tag{2.6}$$

and truncating the series (2.1), we obtain a linear wavelet estimator

$$\hat{f}_n(x) = \sum_{k \in \mathbb{Z}} \hat{a}_{m,k} \varphi_{m,k}(x) . \tag{2.7}$$

3. Asymptotic Normality

The main result of this paper is Theorem 3.1 which establishes central limit theorem of the linear wavelet estimator $\hat{f}_n(x)$. To discuss asymptotic normality of the estimator, note

that $\hat{f}_n(x)$ is the sum of an i.i.d. sequence. Thus, it is sufficient to show that the usual triangular CLT conditions hold. A sufficient condition for asymptotic normality

$$\frac{\hat{f}_n(x) - E \hat{f}_n(x)}{\sqrt{\text{var}(\hat{f}_n(x))}} \rightarrow^L N(0, 1) \tag{3.1}$$

is that Lyapounov's condition holds, i.e. for some $\delta > 0$,

$$\frac{E |Z_{n1} - E Z_{n1}|^{2+\delta}}{n^{\delta/2} (\text{var}(Z_{n1}))^{1+\delta/2}} \rightarrow 0 \tag{3.2}$$

where $Z_{nj} = \sum_{k \in Z} 2^{m/2} U_m(2^m Y_j - k) \varphi_{m,k}(x)$.

The following lemma is used to find a lower bound for $\text{var}(Z_{n1})$.

Lemma 3.1. (Fan(1991)) Suppose that K_n is a sequence of Borel functions satisfying

$$K_n(y) \rightarrow K(y) \quad \text{and} \quad \sup_n |K_n(y)| \leq K^*(y),$$

where $K^*(y)$ satisfies

$$\int_{-\infty}^{\infty} K^*(y) dy < \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} |y K^*(y)| = 0.$$

If x is a continuity point of a density f , then for any sequence $h_n \rightarrow 0$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} \int_{-\infty}^{\infty} K_n\left(\frac{x-y}{h_n}\right) f(y) dy = f(x) \int_{-\infty}^{\infty} K(y) dy.$$

Theorem 3.1. Let X and Z be independent random variables and $Y = X + Z$. Assume that (i) $\tilde{q}(\omega) \omega^\beta \rightarrow c (\neq 0)$ as $\omega \rightarrow \infty, \beta \geq 0$ and $\tilde{q}(\omega)$ does not vanish for all real $\omega \in R$, (ii) $2^m \rightarrow \infty$ and $2^m/n \rightarrow 0$ as $n \rightarrow \infty$. (iii) g is continuous at x . Then,

$$\frac{\hat{f}_n(x) - E \hat{f}_n(x)}{\sqrt{\text{var}(\hat{f}_n(x))}} \rightarrow^L N(0, 1).$$

Proof. From the Fourier inversion formula and (2.2),

$$\begin{aligned} E(Z_{n1}^2) &= E \left(\sum_{k \in Z} 2^{m/2} \varphi_{m,k}(x) 2\pi^{-1} \int_{-\infty}^{\infty} \frac{\tilde{\varphi}(\omega) e^{i\omega(2^m Y_1 - k)}}{\tilde{q}(-2^m \omega)} d\omega \right)^2 \\ &= 2^{2m} \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\omega 2^m y}}{\tilde{q}(-2^m \omega)} \left(\sum_{k \in Z} e^{-ik\omega} \varphi(2^m x - k) \right) d\omega \right)^2 g(y) dy. \end{aligned} \tag{3.3}$$

Note that $\{e^{ik\omega}, k \in Z\}$ is a complete orthogonal system on $L_2[-\pi, \pi]$. Hence,

$$\tilde{\varphi}(-\omega) e^{-i\omega 2^m x} = \sum_{k \in Z} e^{-ik\omega} \varphi(2^m x - k).$$

$$\text{and } E(Z_{n1}^2) = 2^{2m} \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\omega 2^m(y-x)}}{\tilde{q}(-2^m \omega)} d\omega \right)^2 g(y) dy.$$

Then, by Lemma 3.1 and the same procedure in Fan(1991), we can obtain that for sufficiently large n

$$\begin{aligned} E(Z_{n1}^2) &\simeq 2^{m(1+2\beta)} g(x) \int_{-\infty}^{\infty} \left(\frac{1}{2\pi c} \int_{-\pi}^{\pi} \omega^\beta e^{i\omega y} d\omega \right)^2 dy \\ &= \frac{2^{m(1+2\beta)} g(x)}{2\pi c^2} \int_{-\pi}^{\pi} |\omega|^{2\beta} d\omega \end{aligned} \tag{3.4}$$

by Parseval's identity. We can also show that, by Fubini's theorem and Parseval's identity,

$$\begin{aligned} E(Z_{n1}) &= \sum_{k \in \mathbb{Z}} 2^{m/2} \varphi_{m,k}(x) \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{\varphi}(\omega) e^{i\omega(2^m y - k)}}{\tilde{q}(-2^m \omega)} d\omega \right) g(y) dy \\ &= \sum_{k \in \mathbb{Z}} 2^{-m/2} \varphi_{m,k}(x) \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{\tilde{\varphi}(-\omega/2^m)}{\tilde{q}(\omega)} \tilde{g}(\omega) e^{i\omega k/2^m} d\omega \\ &= \sum_{k \in \mathbb{Z}} a_{m,k} \varphi_{m,k}(x) \rightarrow f(x) \text{ as } n \rightarrow \infty. \end{aligned}$$

Similarly, for sufficiently large n, we can get

$$\begin{aligned} E |Z_{n1} - EZ_{n1}|^{2+\delta} &\simeq C_1 2^{m(1+\delta+2\beta+\delta\beta)} g(x) \int_{-\pi}^{\pi} |\omega|^{2\beta} d\omega. \end{aligned}$$

Therefore, $\frac{E |Z_{n1} - EZ_{n1}|^{2+\delta}}{n^{\delta/2} (var(Z_{n1}))^{1+\delta/2}} \rightarrow 0$ if $2^m/n \rightarrow 0$.

Even though we obtain asymptotic normality for a wavelet density estimator, it still involves the unknown quantity $var(\hat{f}_n(x))$. Thus we will replace it by sample variance. In the following corollaries $f(x)$ belongs to a Sobolev space $H^\alpha, \alpha > 0$. Such spaces are characterized by their Fourier transforms to be all tempered distributions whose Fourier transform is a locally bounded and integrable function satisfying $\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 (1 + \omega^2)^\alpha d\omega < \infty$.

Lemma 3.2. Under the assumptions of Theorem 3.1,

$$\sum_{j=1}^n Z_{nj}^2 / (nEZ_{n1}^2) \rightarrow^P 1.$$

Proof. By the weak law of large numbers(Chow and Teicher, p.340), it is sufficient to show that for all $\epsilon > 0$,

$$\frac{1}{EZ_{n1}^2} E(Z_{n1}^2 I_{\{Z_{n1}^2 \geq \epsilon nEZ_{n1}^2\}}) \rightarrow 0. \tag{3.5}$$

By repeating similar calculations, we obtain for sufficiently large n

$$|Z_{n1}|^2 = 2^m \left| \sum_{k \in Z} \varphi_{mk}(x) U_m(2^m Y_1 - k) \right|^2 \leq C_2 2^{2m(1+\beta)} \int_{-\pi}^{\pi} |\omega|^{2\beta} d\omega.$$

Then, by(3.4), $Z_{n1}^2/(nEZ_{n1}^2) \rightarrow 0$ if $2^m/n \rightarrow 0$ as $n \rightarrow \infty$, and hence (3.5) holds.

Corollary 3.1. Under the assumptions of Theorem 3.1, if $2^{-m} = o(n^{-1/(1+2\beta+2\alpha)})$ and $f(x) \in H^\alpha, \alpha > 0$, then

$$\frac{\sqrt{n} (\hat{f}_n(x) - f(x))}{S_n} \rightarrow^L N(0, 1),$$

where $S_n^2 = n^{-1} \sum_{j=1}^n Z_{nj}^2$.

proof. For sufficiently large n,

$$\text{var}(\hat{f}_n(x)) = \frac{1}{n} \text{var}(Z_{n1}) = \frac{C}{n} 2^{m(1+2\beta)} g(x) \int_{-\pi}^{\pi} |\omega|^{2\beta} d\omega.$$

Let $f(x) \in H^\alpha$. Then by Lemma 3.1 of Walter(1999), $|\sum_{k \in Z} a_{mk} \varphi_{mk}(x) - f(x)| \leq A 2^{-am}$.

Thus by Lemma 3.2,

$$\begin{aligned} \frac{\sqrt{n} (\hat{f}_n(x) - f(x))}{S_n} &= \left(\frac{\hat{f}_n(x) - E\hat{f}_n(x)}{\sqrt{\text{var}(\hat{f}_n(x))}} + \frac{E\hat{f}_n(x) - f(x)}{\sqrt{\text{var}(\hat{f}_n(x))}} \right) \times \sqrt{\frac{\text{var}(Z_{n1})}{EZ_{n1}^2}} \times \sqrt{\frac{EZ_{n1}^2}{S_n^2}} \\ &\rightarrow^L N(0, 1), \end{aligned}$$

if $2^{m(1+2\beta+2\alpha)}/n \rightarrow \infty$ as $n \rightarrow \infty$.

Lemma 3.3. Under the assumptions of Theorem 3.1, if $2^m = o(n^{1/(1+2\beta)})$, then

$$\frac{1}{n} \sum_{j=1}^n Z_{nj} - EZ_{n1} \xrightarrow{p} 0.$$

Proof. When n is sufficiently large,

$$\begin{aligned} &P\left[\left| \frac{1}{n} \sum_{j=1}^n Z_{nj} - EZ_{n1} \right| > \varepsilon \right] \\ &\leq \frac{C}{\varepsilon^2} \frac{2^{m(1+2\beta)}}{n} g(x) \int_{-\pi}^{\pi} |\omega|^{2\beta} d\omega \\ &\rightarrow 0, \end{aligned}$$

if $2^{m(1+2\beta)}/n \rightarrow 0$ as $n \rightarrow \infty$.

Corollary 3.2. Under the assumptions of Theorem 3.1, if $2^{-m} = o(n^{-1/(1+2\beta+2\alpha)})$, $2^m = o(n^{1/(1+2\beta)})$, and $f(x) \in H^\alpha, \alpha > 0$, then

$$\frac{\sqrt{n} (\hat{f}_n(x) - f(x))}{S_n} \rightarrow^L N(0, 1),$$

where $S_n^2 = n^{-1} \sum_{j=1}^n Z_{nj}^2 - (n^{-1} \sum_{j=1}^n Z_{nj})^2$.

proof. By Lemma 3.3 and Corollary 3.1. it follows.

Remark. In the wavelet framework 2^{-m} plays the role of usual window h_n and hence the wavelet scale m is very important. In fact, it is not true unless 2^{-m} tends to 0 sufficiently slow as the above theorems indicate.

Example. Let $f(x) \in H^\alpha$ and $q(x) = 0.5ae^{-a|x|}$, the probability density function of a double exponential distribution. For example, if $f(x)$ is the normal p.d.f. with mean μ and variance σ^2 , $f(x) \in H^\alpha, \alpha > 0$. Since $\tilde{q}(\omega) = (\alpha^2\omega^2 + 1)^{-1}$, \tilde{q} has polynomial descent. Thus under the assumptions of Corollary 3.2 with $\beta=2$, $\frac{\sqrt{n}(\hat{f}_n(x) - f(x))}{S_n} \rightarrow^L N(0,1)$, where $S_n^2 = n^{-1} \sum_{j=1}^n Z_{nj}^2 - (n^{-1} \sum_{j=1}^n Z_{nj})^2$.

References

[1] Carroll, R.J. and Hall, P. (1998). Optimal rates of convergence for deconvoluting a density, *Journal of American Statistics Association*, 83, 1184-1886.

[2] Chow, Y.S. and Teicher, H. (1988). *Probability Theory*, Springer-Verlag, New York Inc.

[3] Daubechies. I. (1992). Ten Lectures on Wavelets, CBMS-MSF, SIAM, philadelphia.

[4] Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D.(1995). Wavelet shrinkage: asymptopia? (with discussion), *Journal of Royal Statisticl Society*, Ser. B. 57, 301-369.

[5] Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding, *Annals of Statistics*, 24, 508-539.

[6] Fan, J. (1991a). On the Optimal Rates of Convergence for Nonparametric deconvolution problem, *Annals of Statistics*, 19, 1257-1272.

[7] Fan, J. (1991b). Asymptotic normality for deconvolution kernel density estimators, *Sankhya* 53, Series A, 97-110.

[8] Lee, S. (2001). A Note on L_1 strongly consistent wavelet density estimator for the deconvolution problem, *The Korean Communications in Statistics*, 8.

[9] Lee, S. and Hong, D. H. (2002). On a strongly consistent wavelet density estimator for the deconvolution problem, to appear in *Communications in Statistics: Theory and Methods*, 31(8).

[10] Louis, T.A. (1991). Using empirical Bayes methods in biopharmaceutical research, *Statistics in Medicine*, 10, 811-827.

[11] Pensky, M. and Vidakovic, B. (1999). Adaptive Wavelet estimator for nonparametric

- density deconvolutoin, *Annals of Statistics*, 27(6), 2033-2053 .
- [12] Taylor, R. L. and Zhang, H. M. (1990). On a strongly consistent nonparametric density estimator for deconvolution problem, *Communications in Statistics: Theory and Methods*, 19, 3325-3342.
- [13] Vidakovic, B. (1999). *Statistical Modeling by Wavelets*, John Wiley & Sons, Inc., New York, NY.
- [14] Walter, G. G. (1994). *Wavelets and Other Orthogonal Systems with Applications*, CRC Press, Boca Raton, FL.
- [15] Walter, G. G. (1999). Density estimation in the presence of noise, *Statistics & Probability Letters*, 41, 237-246.
- [16] Zhang, H. P. (1992). On deconvolution using time of flight information in positron emission tomography, *Statistica Sinica*, 2, 553-575.

[2002년 1월 접수, 2002년 3월 채택]