Scientific
Research

Wavelet Density Estimation of Censoring Data and Evaluate of Mean Integral Square Error with Convergence Ratio and Empirical Distribution of Given Estimator

Mahmoud Afshari

Department of Statistics, College of Science, Persian Gulf University, Bushehr, Iran
Email: afshar@pgu.ac.ir

Abstract

Wavelet has rapid development in the current mathematics new areas. It also has a double meaning of theory and application. In signal and image compression, signal analysis, engineering technology has a wide range of applications. In this paper, we use wavelet method, for estimating the density function for censoring data. We evaluate the mean integrated squared error, convergence ratio of given estimator. Also, we obtain empirical distribution of given estimator and verify the conclusion by two simulation examples.

Keywords

1. Introduction

One of data types, which researchers are extremely interested in, is caring to the time interval till the occurrence of certain events such as death etc. Any process waiting for a specific event produces survival data. Survival function, which is shown by $S(t)$, indicates the ratio of people who survived since the base time which is the point they enter the experiment. Failure in survival analysis means the occurrence of the event we were waiting for. The time, where survival is measured after that point, is called the start time. The failure time is the time that failure occurs for each individual which is denoted by T_i for $i = 1,2,3,\cdots$. The failure time is occurred from the base time up to when the failure occurs and it's known as T_i. It's not always possible to observe the failure time for each individual. In such cases, censorship occurs. The rate of occurrences of an event (failure) in a spe-

cific short period of time providing that no failure occurred before that time is the concept which is discussed by the name hazard function in survival analysis. Hazard function for the failure time line is as follows:

$$h_t(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t + \Delta t | T \ge t)}{\Delta t} = \frac{F'(t)}{1 - F(t)} = -\frac{S'(t)}{S(t)}.$$

Wavelets can be used for transient phenomena analysis or functions analysis which sometimes changes rapidly, and they are symmetrical and have limited period unlike rugged Sine waves, thus the signals with radical changes are analyzed better. The close relationship between wavelet coefficients and some spaces, wavelet bases being orthogonal and also useful properties of them in wavelet issues simplify the computational algorithms.

Wavelets theory was proposed by Alfred Harr [1] for the first time in 1910. He showed that a continuous function can be approximated as follows:

$$f_n(x) = \langle f, \phi_0 \rangle \phi_0(x) + \langle f, \phi_1 \rangle \phi_1(x) + \cdots + \langle f, \phi_n \rangle \phi_n(x). \tag{1}$$

Such that $\langle f, \phi_i \rangle = \int f(x) \phi_i(x) dx$,

$$\phi_n(x) = I_{\frac{k}{2^j} \le x < \frac{k+\frac{1}{2}}{2^j}} - I_{\frac{k+\frac{1}{2}}{2^j} \le x < \frac{k+1}{2^j}} \quad , \quad j \ge 0 \ \ k = 0, 1, \cdots, 2^j - 1 \ \ , \ \ n = 2^j + k$$

Also for mother wavelet and father wavelets the following:

$$\phi_{j,k}(x) = \phi(2^j x - k), \quad \psi_{j,k}(x) = \psi(2^j x - k).$$

Definition 1-1: Assume that $V_j = \overline{\text{span}}\{\varphi_{j,k} : k \in z\}$; $\{\varphi_{j,k}, k \in z\}$ is an orthogonal unit base for V_j and V_j contains all sectionally constant functions and their exact length is twice the interval length of V_{j+1}.

Spaces $\{V_j, j \in Z\}$ are called multiresolatio analysis or scale function φ, if it satisfies the following conditions:

1- $V_j \subset V_{j+1} \forall j \in Z$, 2- $\overline{\bigcup_{j \in z} V_j} = L^2(R)^2$, 3- $\bigcap_{j \in z} V_j = \{0\}$,

4- $j \in z$. $f(x) \in V_j$, $f(x) \in V_j$, 5- ' $-k \in z. \Leftrightarrow f(x) \in V_0$, $f(x-h) \in V_0$.

6- $\exists \varphi(x) \in V_0$ in condition that $\{\varphi(x-k) : k \in z\}$ is an orthogonal base for V_0.

If we consider the scale function in the interval $[0,1]$, then the image of f on the space V_j is defined as $P_{V_j}^f = \sum_{k \in z} \alpha_{j,k} \varphi_{j,k}(x)$ which is a function with the resolution, 2^j and because of the fact that $\overline{\bigcup_{j \in z} V_j} = L^2(R)$

thus $P_{V_j}^f$ is a good approximation of function f for large amounts of j.

Let the nested sequence of closed subspaces; ... $\mathbf{V}_{j-1} \subset \mathbf{V}_j \subset \mathbf{V}_{j+1} \subset \cdots, j \in \mathbf{Z}$, be a multiresolutuon approximation to $\mathbf{L}^2(R)$. Define $\mathbf{W}_j$, $j \in \mathbf{Z}$ to be orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j+1}$.

The term wavelets are used to refer to a set of basis functions with very special structure. The special of wavelets basis for function $f \in \mathbf{L}^2(R)$ as scaling function φ and mother wavelet ψ such that $\{\varphi(x-k)\}_{k \in Z}$ forms an orthogonal basis for $\mathbf{V}_0$ and $\{\psi(x-k)\}_{k \in Z}$ forms an orthonormal basis for $\mathbf{W}_0$. Other wavelets in the basis are then generated by translation of the scaling function and dilations of the mother wavelet by using the relationships:

$$\varphi_{m_0,k}(x) = 2^{m_0/2} \varphi(2^{m_0} x - k), \quad \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k) \tag{2}$$

Given above Wavelet basis, a function $f \in \mathbf{L}^2(R)$ can be written a formal expansion:

$$f = \sum_{k \in Z} \alpha_{m_0,k} \varphi_{m_0,k} + \sum_{j=m_0}^{\infty} \sum_{k \in Z} \beta_{j,k} \psi_{j,k} \tag{3}$$

where $\alpha_{j_0,k} = \int f(x) \varphi_{j_0,k}(x) dx$, $\delta_{j,k} = \int f(x) \psi_{j,k} dx$

As for general orthogonal series estimator, Daubechies [2], density estimator can be written as:

$$\hat{f} = \sum_{k \in Z} \hat{\alpha}_{m_0,k} \phi_{m_0,k}(x) + \sum_{j \ge m_0} \sum_{k \in Z} \hat{\beta}_{j,k} \psi_{j,k}(x) = \mathbf{P}_{m_0} f + \sum_{j \ge m_0} \sum_{k \in Z} \hat{\beta}_{j,k} \psi_{j,k}. \tag{4}$$

where the obvious coefficient estimator can be written:

$$\hat{\alpha}_{m_0,k} = E\left[\varphi_{m_0,k}(X)\right] = \frac{1}{n}\sum_{i=1}^{n}\varphi_{m_0,k}(X_i), \quad \hat{\beta}_{j,k} = E\left[\psi_{j,k}(X)\right] = \frac{1}{n}\sum_{i=1}^{n}\psi_{j,k}(X_i) \tag{5}$$

We divide time axis into two parts, the intervals and the number of events in each interval. We determine number of events and hazard function according to the observations. Then we flatten them separately via linear wavelet density estimation on the whole time and then we calculate the function estimator and evaluate the asymptotic distribution.

In this paper we obtain estimator density for censoring data by using wavelet method and evaluate mean integral square error with convergence ratio and empirical distribution of given estimator.

2. Estimator of Density by Using Wavelet Method

Wavelets can be used for transient phenomena analysis or functions analysis which sometimes changes rapidly, and they are symmetrical and have limited period unlike rugged Sine waves, thus the signals with radical changes are analyzed better. The close relationship between wavelet coefficients and some spaces, wavelet bases being orthogonal and also useful properties of them in wavelet issues simplify the computational algorithms. As a result, numerous articles have been published about density function estimation. The mathematical theorem of wavelets and their application in statistics have been studied as a technique for nonparametric curve estimators by Antoniadys [3].

Afshari [4]-[6] have done some researches about density function estimator, the density functional derivative and the nonparametric regression function for the mixing random variables. Donohu [7], kyacharyan, Picard [8], Malat [9], Meyer [10], and some articles have been published in this field. Hall and Patil [11] have found a formula for the Mean Integrated Squared Error of Nonlinear Wavelet based on density estimators. Antoniadys *et al.* [12] achieved the density function estimator and the hazard function for right-censored data with the wavelets. In this section we obtain estimator of density function for censoring data by using wavelet method.

Suppose $X_1, X_2, X_3, \cdots, X_n$ are failure time of n tests that are studied. They are non-negative, independent, identically distributed, with the density function f and distribution function F and $C_1, C_2, C_3, \cdots, C_n$ are corresponding to censored times, non-negative, independent, identically distributed, with the density function g and distribution function G.

Assuming independency of failure times and censored time of the observed random variable, Z_i and the function δ_i and Hazard function are shown as below:

$$Z_i = \min(X_i, C_i), \quad \delta_i = I_{(X_i \leq C_i)}, \quad h(t) = \frac{f(t)}{1-F(t)}, \quad F(t) < 1.$$

Such that $I_{(A)}$ is indicator function of A. For data censoring, if $G(t) < 1$ then we have as the following:

$$h(t) = \frac{f(t)(1-G(t))}{(1-F(t))(1-G(t))}, \quad F(t) < 1.$$

Also we definite as follows:

$$L(t) = P(Z_i \leq t) = 1 - P(Z_i > t) = 1 - P(X_i > t, C_i > t) = 1 - (1-F(t))(1-G(t)).$$

$$f^*(t) = f(t)(1-G(t)). \quad h(t) = \frac{f^*(t)}{1-L(t)}, \quad L(t) < 1.$$

To estimate $f^*(t)$, we divide the time axis into two parts of small intervals and the amounts of events (0 or 1) in each interval, and then we divide these values to the length of intervals.

Estimation procedures of $f^*(t)$ can be summarized as the following:

Select $\Delta > 0$ and collect the observed failures in $k+1$ intervals with the length Δ and using wavelet estimation on the collected data. We find an estimate of sub density. This means that we calculate the collected wavelet coefficients data on the scale of $j(n)$ by choosing the decomposition level $j(n)$ and then we estimate $f^*(t)$. It is necessary to state the following symbols to show the details:

$$T_F = \sup\{t : F(t) < 1\} \cdot \quad T_G = \sup\{t : G(t) < 1\} \cdot \quad T_L = \sup\{t : L(t) < 1\} = \min\{T_F, T_G\}.$$

We figure estimators on the finite interval $[0, \tau]$ in which $\tau < T_L$. Note that if $Z_{(i)}$ is the ordinal order statistic i of the sequence Z_i then $T_{L_n} = Z_{(n)} \xrightarrow{n \to \infty} T_L$. In fact we suppose $\tau = Z_{(n)}$.

Suppose that N is an integer that could be dependent to n and the estimated points are as follows:

$$t_k = \frac{k\tau}{2^N}, \qquad k = 0, \cdots, K = 2^N - 1 \cdot$$

Suppose that $\Delta = \tau 2^{-N}$ and we divide the interval $[0, \tau]$ of time axis to $k+1$ intervals with Δ long

$$\tau_0 = -\frac{\Delta}{2}, \quad \tau_k = t_k - \frac{\Delta}{2}, k = 1, \cdots, K, \quad \tau_{K+1} = \tau.$$

The k-th interval is marked by J_k so: $J_k = [\tau_K, \tau_{k+1})$ for $k = 0, \cdots, K-1$, $J_K = [\tau_K, \tau]$.

Now we define the following indicator function that indicates the number of uncensored failures in the time interval $J_k : Y_{ik} = I_{J_k}(Z_i)\delta_i$, $i = 1, \cdots, n$ $k = 0, \cdots, K \cdot$ We assume that U_k the observed failures ratio in the interval J_k n other words: $U_k = \frac{1}{n}\sum_{i=1}^{n} Y_{ik}$, $k = 0, \cdots, K$.

Theorem 2-1: Suppose that the sub density f^* is a continuous function on $[0, \tau]$ and it's m times differentiable, then if $_V \Delta \to 0$ or $n \to \infty$, we have:

$$Var\left(\frac{U_k}{\Delta}\right) = \frac{f^*(t_k)}{n\Delta} + O\left(\frac{1}{n}\right), E\left(\frac{U_k}{\Delta}\right) = f^*(t_k) + O(\Delta)$$

$$Cov\left(\frac{U_k}{\Delta}, \frac{U_\ell}{\Delta}\right) = -\frac{1}{n} f^*(t_k) f^*(t_\ell) + O\left(\frac{\Delta}{n}\right), \quad k \neq \ell.$$

Proof: see [13].

We smooth the data $\frac{U_k}{\Delta}$ by an appropriate wavelet smoother to find the estimation of f^*.

We can write,

$$f^*(t) = \sum_{k=0}^{2^{j_0}-1} \left\langle f^*, \phi_{j_0,k} \right\rangle \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_{\ell=0}^{2^j-1} \left\langle f^*, \psi_{j,\ell} \right\rangle \psi_{j,\ell}(t). \tag{6}$$

where, $\left\langle f, g \right\rangle \equiv \int_0^\tau f(t) g(t) \mathrm{d}t$

The complex structural polymorphism analysis causes an efficient tree construction algorithm for analysis of functions in V_N with theoretic scale wavelet coefficients $\left\langle f^*, \varphi_{N,k} \right\rangle$. However, the integral scale $\left\langle f^*, \varphi_{N,k} \right\rangle$ is not well available and we need an initial value for a fast wavelet transform. Antonyadys [4] suggested the following initial amount:

$$\left\langle f^*, \phi_{N,k} \right\rangle = 2^{-\frac{N}{2}} f^* t_k + O\left(2^{-\frac{N}{2}} 2^{-Nm}\right), \quad 0 \leq k \leq 2^N - 1$$

As a result a reasonable estimate for image of f^* with clarity N is:

$$\tilde{f}_N^*(t) = 2^{-\frac{N}{2}} \sum_{k=0}^{K} \frac{U_k}{\Delta} \varphi_{N,k}(t). \tag{7}$$

If we assume that the collected values U_k which are equal to the estimators of $\tilde{f}_N^*(t)$, are in Sobolev space $W^m([0, \tau])$ and φ is regular of degree m. We estimate the unknown function f^* as follows to level the data with a better rate for the sample size n and the sequence $j(n) < N$:

$$\hat{f}_n = P_{V_{j(n)}}^{\tilde{f}_N^*} \tag{8}$$

That it is the orthogonal image of $\tilde{f}_N^*(t)$ on the leveler approximation space $V_{j(n)}$.

Theorem 2-2: Suppose that the sub density f^* is a continuous function on $[0, \tau]$ and it's m times diffe-ren- tiable, then if $\Delta \to 0$ for $n \to \infty$ we have:

$$E\left[\tilde{f}_N^*(t)\right] = P_{V_N}^{f^*(t)} + O(\Delta), Var\left[\tilde{f}_N^*(t)\right] = O\left\{(n\Delta)^{-1}\right\} + O\left(n^{-1}\right).$$

Proof: by using theorem (2-1) we can write:

$$\left[\tilde{f}_N^*(t)\right] = 2^{-\frac{N}{2}} \sum_{k=0}^{K} E\left[\frac{U_k}{\Delta}\right] \phi_{N,k}(t) = 2^{-\frac{N}{2}} \sum_{k=0}^{K} \left[f^*(t_k) + O(\Delta)\right] \phi_{N,k}(t)$$

$$= 2^{-\frac{N}{2}} \sum_{k=0}^{K} f^*(t_k) \phi_{N,k}(t) + O(\Delta) 2^{-\frac{N}{2}} \sum_{k=0}^{K} \phi_{N,k}(t) \tag{9}$$

Since, $\sup_t \sum_k |\phi(t-k)| = M$, then $\varphi_{N,k}(t) = 2^{\frac{N}{2}} \varphi(2^N t - k)$ and we can write as the following:

$$\sup_t 2^{-\frac{N}{2}} \sum_k |\phi_{N,k}(t)| = M.$$

So Equations (9) can be written as follows:

$$E\left[\tilde{f}_N^*(t)\right] \leq 2^{-\frac{N}{2}} \sum_{k=0}^{K} f^*(t_k) \phi_{N,k}(t) + O(\Delta) M \tag{10}$$

By using Equation (1) we have:

$$2^{-\frac{N}{2}} \sum_{k=0}^{K} f^*(t_k) \phi_{N,k}(t) + O(\Delta) M = \sum_{k=0}^{K} \left\{ \langle f^*, \phi_{N,k} \rangle + O\left(\Delta^{m+\frac{1}{2}}\right) \right\} \phi_{N,k}(t) + O(\Delta)$$

$$= \sum_{k=0}^{K} \langle f^*, \phi_{N,k} \rangle \phi_{N,k}(t) + O\left(\Delta^{m+\frac{1}{2}}\right) + O(\Delta). \tag{11}$$

By using Equations (10) and (11) we have:

$$E\left[\tilde{f}_N^*(t)\right] = P_{V_N}^{f^*(t)} + O(\Delta).$$

$$Var\left[\tilde{f}_N^*(t)\right] = 2^{-\frac{N}{2}} \sum_{k=0}^{K} Var\left(\frac{U_k}{\Delta}\right) \phi_{N,k}^2(t) + 2^{-N} \sum_{k=0}^{K} \sum_{l=0, l \neq k}^{K} Cov\left(\frac{U_k}{\Delta}, \frac{U_l}{\Delta}\right) \phi_{N,k}(t) \phi_{N,k}(t)$$

By using theorem (2-1) we can writhe as follows:

$$Var\left[\tilde{f}_N^*(t)\right] = 2^{-N} \sum_{k=0}^{K} \left[\frac{f^*(t_k)}{n\Delta} + O\left(n^{-1}\right)\right] \phi_{N,k}^2(t)$$

$$+ 2^{-N} \sum_{k=0}^{K} \sum_{l=0, l \neq k}^{K} \left[-\frac{1}{n} f^*(t_k) f^*(t_l) + O\left(\Delta n^{-1}\right)\right] \phi_{N,k}(t) \phi_{N,l}(t)$$

Using this fact that f^* is uniformly bounded on $[0, \tau]$ and $\Delta = O\left(2^{-N}\right)$, we have:

$$Var\left[\tilde{f}_N^*(t)\right] \leq \frac{1}{n} C_1 \sum_{k=0}^{K} \phi_{N,k}^2(t) + O\left(\frac{\Delta}{n}\right) \sum_{k=0}^{K} \phi_{N,k}^2(t) + C_2 \frac{\Delta}{n} \sum_k \sum_{k \neq l} |\phi_{N,k}(t) \phi_{N,l}(t)|$$

$$+ O\left(\frac{\Delta^2}{n}\right) \sum_k \sum_{k \neq l} |\phi_{N,k}(t) \phi_{N,l}(t)|. \tag{12}$$

Since φ is regular in order m we can write:

$$\sum_{k=0}^{K} \phi_{N,k}^2(t) = O\left(2^N\right), \quad \sum_k \sum_{k \neq l} |\phi_{N,k}(t)| |\phi_{N,l}(t) = O\left(2^{2N}\right)| \tag{13}$$

According Equation (13), we can write: $Var\left[\tilde{f}_N^*\left(t\right)\right]=O\left\{\left(n\Delta\right)^{-1}\right\}+O\left(n^{-1}\right)$, complete the proof.

3. Evaluate of Mean Integral Square Error with Convergence Ratio

In this section we evaluate mean integral square error and convergence ratio is investigated.

Definition 3-1: The mean integrated square error (MISE) of kernel estimator of a density function f is given $MISE \approx C_1\left(nh\right)^{-1}+C_2h^{2r}$. In this formula $\approx$ denotes the right and left convergence, when $n\to\infty$, n denotes the sample size, h denotes the estimator bandwidth core, r denotes core level and C_1, C_2 denote kernel dependent quantities with unknown density.

Theorem 3-1: Suppose that the sub density f^* is a continuous function on $[0,\tau]$ and it's m times differentiable, then if $\Delta\to 0$ for $n\to\infty$ and $j\left(n\right)\to\infty$, then $n2^{-j\left(n\right)}\to\infty$,

$$\text{MISE}\left(\hat{f}_n\right)=E\left[\int_0^\tau\left\{\hat{f}_n\left(t\right)-f^*\left(t\right)\right\}^2\mathrm{d}t\right]\le O\left(2^{-2j\left(n\right)m}\right)+O\left(\Delta^2\right)+O\left(\frac{2^{j\left(n\right)}}{n}\right) \tag{14}$$

Proof:

$$E\left(t\right)=\hat{f}_n\left(t\right)-f^*\left(t\right)=P_{V_{j(n)}}^{\tilde{f}_N^*\left(t\right)}-f^*\left(t\right)=P_{V_{j(n)}}^{\tilde{f}_N^*\left(t\right)}-P_{V_{j(n)}}^{f^*\left(t\right)}+P_{V_{j(n)}}^{f^*\left(t\right)}-f^*\left(t\right)=S\left(t\right)+A\left(t\right). \tag{15}$$

By using Equation (15) and theorem (2-2) for $m\ge 1$, we can write as the following:

$$E\left[E\left(t\right)\right]=E\left[S\left(t\right)\right]+A\left(t\right)=P_{V_{j(n)}}^{f^*\left(t\right)}+O\left(\Delta\right)-P_{V_{j(n)}}^{f^*\left(t\right)}+A\left(t\right)$$

Because $V_{j\left(n\right)}\subset V_N$ we can write as the following:

$$E\left[E\left(t\right)\right]=O\left(\Delta\right)+A\left(t\right) \tag{16}$$

$$\left\|P_{V_{j(n)}}^{f^*\left(t\right)}-f^*\left(t\right)\right\|_2=O\left(2^{-mj\left(n\right)}\right) \tag{17}$$

So by using Equations (16) and (17), we can write:

$$\text{MISE}\left(\hat{f}_n\right)=O\left(\Delta^2\right)+O\left(2^{-mj\left(n\right)}\right). \tag{18}$$

For evaluate $Var\left[\text{MISE}\left(\hat{f}_n\right)\right]$, we can write:

$$P_{V_{j(n)}}^{\tilde{f}_N^*\left(t\right)}=\sum_{k=0}^{2^{j(n)}-1}\left\langle\tilde{f}_N^*,\phi_{j(n),k}\right\rangle\varphi_{j(n),k}\left(t\right)=\sum_{k=0}^{2^{j(n)}-1}\left\{\sum_{l=0}^{2^N-1}2^{-\frac{N}{2}}\left\langle\frac{U_l}{\Delta}\varphi_{N,l},\varphi_{j(n),k}\right\rangle\right\}\varphi_{j(n),k}\left(t\right)$$

Also we can write:

$$P_{V_{j(n)}}^{\tilde{f}_N^*\left(t\right)}-P_{V_{j(n)}}^{E\left(\tilde{f}_N^*\left(t\right)\right)}=\sum_k 2^{-\frac{N}{2}}\sum_l\frac{U_l-E\left(U_l\right)}{\Delta}\left\langle\phi_{N,l},\phi_{j(n),k}\right\rangle\phi_{j(n),k}\left(t\right) \text{ then,}$$

$$\int\left\{P_{V_{j(n)}}^{\tilde{f}_N^*\left(t\right)}-P_{V_{j(n)}}^{E\left(\tilde{f}_N^*\left(t\right)\right)}\right\}^2\mathrm{d}t=2^{-N}\sum_l\left\{\sum_k\frac{U_l-E\left(U_l\right)}{\Delta}\left\langle\phi_{N,k},\phi_{j(n),l}\right\rangle\right\} \tag{19}$$

By using theorem (2-1) and expectation of Equation (19), we can write as the following:

$$Var=\sum_l 2^{-N}\sum_k E\left\{\frac{\left(U_k-E\left(U_k\right)\right)^2}{\Delta^2}\right\}\left\langle\varphi_{N,k},\varphi_{j(n),l}\right\rangle$$
$$+\sum_l 2^{-N}\sum_k\sum_{h\ne k}Cov\left(\frac{U_k}{\Delta},\frac{U_h}{\Delta}\right)\left\langle\varphi_{N,k},\varphi_{j(n),l}\right\rangle\left\langle\varphi_{N,h},\varphi_{j(n),l}\right\rangle^2 \tag{20}$$

By using theorem (2-1) we have:

$$\sum_l 2^{-N} \sum_k E\left\{\frac{\left(U_k - E(U_k)\right)^2}{\Delta^2}\right\} = \Delta \sum_{k,l}\left\{\frac{f^*(t_k)}{n\Delta} + O(n^{-1})\right\}\left\langle \phi_{N,k}, \phi_{j(n),l}\right\rangle^2$$

$$= \Delta \sum_{k,l}\left\{\frac{f^*(t_k)}{n\Delta} + O(n^{-1})\right\}\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle^2$$

(21)

$$\sum_k \left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle^2 = \left\|P_{V_N}^{\varphi_{j(n),l(t)}}\right\|_2^2 = \left\|\varphi_{j(n),l}\right\|_2^2 = \int \varphi_{j(n),l}(t)\,\mathrm{d}t = \int 2^{j(n)}\phi^2\left(2^{j(n)}t - l\right)\mathrm{d}t = 1 \qquad (22)$$

By using Equation (22) and this fact that f^* is uniformly bounded, we can write as the following:

$$\sum_l 2^{-N} \sum_k E\left\{\frac{\left(U_k - E(U_k)\right)^2}{\Delta^2}\right\}\left\langle \phi_{N,k}, \phi_{j(n),l}\right\rangle^2 = O\left(\frac{2^{j(n)}}{n}\right) + O\left(\frac{\Delta 2^{j(n)}}{n}\right) = O\left(\frac{2^{j(n)}}{n}\right).$$

The second part of Equation (20) can be written as the following:

$$\sum_l 2^{-N} \sum_k \sum_{h\neq k} Cov\left(\frac{U_k}{\Delta}, \frac{U_h}{\Delta}\right)\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle\left\langle \varphi_{N,h}, \varphi_{j(n),l}\right\rangle$$

$$= \sum_l 2^{-N} \sum_k \sum_{h\neq k}\left\{-\frac{1}{n}f^*(t_k)f^*(t_h) + O(\Delta n^{-1})\right\}\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle\left\langle \varphi_{N,h}, \varphi_{j(n),l}\right\rangle$$

$$= \sum_k \sum_{h\neq k}\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle\left\langle \varphi_{N,h}, \varphi_{j(n),l}\right\rangle$$

$$\leq \sum_k \sum_h \left|\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle\right|\left|\left\langle \varphi_{N,h}, \varphi_{j(n),l}\right\rangle\right| = \left(\sum_k \left|\left\langle \varphi_{N,k}, \varphi_{j(n),l}\right\rangle\right|\right)^2$$

$$\leq 2^N \sum_k \left\langle \phi_{N,k}, \phi_{j(n),l}\right\rangle^2 = 2^N \left\|\phi_{j(n),l}\right\|_2^2 = 2^N$$

By using $O\left(\dfrac{2^{j(n)}}{n}\right) + O\left(\dfrac{\Delta 2^{j(n)}}{n}\right) = O\left(\dfrac{2^{j(n)}}{n}\right)$, the proof is complete.

4. Empirical Distribution of Purpose Estimator

In this section we investigate empirical distribution of estimator under some condition.

Theorem 4-1 Suppose that the sub density f^* is a continuous function on $[0,\tau]$ and it's m times differentiable, for $n \to \infty$, $n\Delta \to \infty$, $n\Delta^3 \to 0$, $n\Delta 2^{-j(n)(2m-1)} \to 0$, then for interval $[0,\tau]$, we have:

$$\sqrt{n\Delta}\left(\hat{f}_n(t) - f^*(t)\right) \to N\left(0, f^*(t)\sum_{k=-\infty}^{\infty}\varphi^2(k)\right).$$

Proof:

$$\sqrt{n\Delta}\left(\hat{f}_n(t) - f^*(t)\right) = \sqrt{n\Delta}\left(\hat{f}_n(t) - E\left[\hat{f}_n(t)\right]\right) + \sqrt{n\Delta}\left(E\left[\hat{f}_n(t)\right] - f^*(t)\right)$$

By using theorems (2-1) and (2-2), we can write as the following:

$$E\left[\hat{f}_n(t)\right] - f^*(t) = O(\Delta) + \left(P_{V_{j(n)}}^{f_N^*(t)}\right) - f^*(t),$$

$$\sup_{f\in W^m}\left\|P_{V_{j(n)}}^{f(t)} - f(t)\right\|_\infty = O\left(2^{-j(n)\left(m-\frac{1}{2}\right)}\right), \left(P_{V_{j(n)}}^{f^*(t)} - f^*(t)\right) = O\left(2^{-j(n)\left(m-\frac{1}{2}\right)}\right). \qquad (23)$$

$$E\left[\hat{f}_n(t)\right] - f^*(t) = O(\Delta) + O\left(2^{-j(n)\left(m-\frac{1}{2}\right)}\right). \qquad (24)$$

So by using equation of (23) and (24) we can write as the following:

$$\sqrt{n\Delta}\left(\hat{f}_n(t) - E\left(\hat{f}_n(t)\right)\right) = \sqrt{n\Delta}\left(\hat{f}_n(t) - \tilde{f}_N^*(t)\right) + \sqrt{n\Delta}\left(\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right)$$

$$+ \sqrt{n\Delta}\left(E\left(\tilde{f}_N^*(t)\right) - E\left(\hat{f}_n(t)\right)\right) = \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

We prove that II has asymptotically normal distribution and also I, III tend to zero when $n \to \infty$
First, we show that I, III tend to zero when $n \to \infty$. According to Equation (24) we have:

$$\hat{f}_n(t) - \tilde{f}_N^*(t) = P_{V_{j(n)}}^{\tilde{f}^*(t)}\tilde{f}_N^*(t) = O\left(2^{-j(n)\left(m-\frac{1}{2}\right)}\right).$$

$$E\left(\tilde{f}_N^*(t)\right) - E\left(\hat{f}_n(t)\right) \le \left|E\left(\tilde{f}_N^*(t)\right) - E\left(\hat{f}_n(t)\right)\right| \le E\left|\tilde{f}_N^*(t) - \hat{f}_n(t)\right| = O\left(2^{-j(n)\left(m-\frac{1}{2}\right)}\right). \quad (25)$$

By using Equation (23) we have:

$$\left|P_{V_{j(n)}}^{f^*(t)} - f(t)\right| \le K2^{-j(n)\left(m-\frac{1}{2}\right)}.$$

So by using Equation (24) and (25), the phrase I, III tend to zero when $n \to \infty$, and finally we have:

$$\sqrt{n\Delta}\tilde{f}_N^*(t) = \sqrt{n\Delta}\sum_{k=0}^{K} 2^{-\frac{N}{2}}\frac{U_k}{\Delta}\cdot 2^{\frac{N}{2}}\phi\left(2^N t - k\right) = \frac{\sqrt{n}}{\sqrt{\Delta}}\sum_{k=0}^{K} U_k\phi\left(2^N t - k\right)$$

$$= \frac{1}{\sqrt{n\Delta}}\sum_{k=0}^{K} nU_k\phi\left(2^N t - k\right) = \sum_{i=1}^{n}\frac{1}{\sqrt{n\Delta}}\sum_{k=0}^{K} Y_{ik}\phi\left(2^N t - k\right).$$

So we have:

$$\sqrt{n\Delta}\left(\hat{f}_n(t) - E\left(\hat{f}_n(t)\right)\right) = \sum_{i=1}^{n}\frac{1}{\sqrt{n\Delta}}\sum_{k=0}^{K}\left(Y_{ik} - p_k\right)\phi\left(2^N t - k\right) = \sum_{i=1}^{n} Z_{ni} \quad (26)$$

Such that for each fixed k, while $i = 1, 2, \cdots, n$, Y_{ik} is defined as an independent and identically distributed random sample with the mean as follows:

$$p_k\left|\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right| \le \frac{1}{n\Delta}2^{-\frac{N}{2}}\left(\sum_{k=0}^{K}\left\{\left|\phi_{N,k}(t)\right|\left|\sum_{i=1}^{n}\left(Y_{ik} - p_k\right)\right|\right\}\right).$$

By using cushy Schwartz inequality:

$$\left|\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right|^2 \le \frac{1}{n^2\Delta^2}2^{-N}\left(\sum_{k=0}^{K}\left|\phi_{N,k}(t)\right|^2\right)\left(\sum_{k=0}^{K}\left|\sum_{i=1}^{n}\left(Y_{ik} - p_k\right)\right|^2\right) \le \frac{M^2}{n^2\Delta^2}\left(\sum_{k=0}^{K}\left|\sum_{i=1}^{n}\left(Y_{ik} - p_k\right)\right|^2\right). \quad (27)$$

So we can write as the following:

$$E\left[\sup_{0 \le t \le \tau}\left|\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right|^2\right] \le \frac{M^2}{n^2\Delta^2}\left(\sum_{k=0}^{K} Var\left(Y_{ik}\right)\right).$$

Using this fact that f^* is uniformly bounded and, $Var\left(Y_{ik}\right) = \Delta f^*(t) + O\left(\Delta^2\right)$, $k = 1, 2, \cdots, 2^N - 1$, we can write:

$$Var\left(Y_{ik}\right) = \Delta f^*(t) + O\left(\Delta^2\right),$$

$$E\left[\sup_{\circ \le t \le \tau}\left|\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right|^2\right] \le \frac{M^2}{n^2\Delta^2}\left(\sum_{k=0}^{K}\left(f^*(t_k) + c\right)\right) \le \frac{c}{n\Delta^2}$$

Thus, the Equation (26) state is convergent in L^2 and thus in the distribution.
Also by using Theorem (2-2), we have:

$$Var\left[Z_{ni}\right] = \frac{1}{n}\left(\sum_{k=0}^{K} f^*(t)\phi^2\left(2^N t - k\right) + O\left(\Delta\right)\right).$$

Thus we have:

$$Var\left[\sqrt{n\Delta}\ \left(\tilde{f}_N^*(t) - E\left(\tilde{f}_N^*(t)\right)\right)\right] \to f^*(t)\sum_{k=0}^{K}\phi^2(k)\cdot$$

We control the Lindberg condition in order to prove that II is asymptotically normal. For this purpose, we set: $U_{ni} = \dfrac{Z_{ni}}{\sqrt{VarZ_{ni}}}$, and we show that $E\left\{U_{ni}^2\ \mathrm{I}_{|U_{ni}|>\varepsilon\sqrt{n}}\right\} \to 0\cdot$

By using cushy Schwartz inequality:

$E\left(U_{ni}^4\right) = O\left(n^2\right)E\left(Z_{ni}^4\right), E\left(Z_{ni}^4\right) = O\left(\dfrac{1}{n^2\Delta}\right), E\left\{U_{ni}^2\ \mathrm{I}_{|U_{ni}|>\varepsilon\sqrt{n}}\right\} \leq \left\{E\left(U_{ni}\right)^4\right\}^{\frac{1}{2}}\left(\varepsilon\sqrt{n}\right)^{-1}$, So we can write as the following:

$E\left\{U_{ni}^2\ \mathrm{I}_{|U_{ni}|>\varepsilon\sqrt{n}}\right\} = O\left(\dfrac{1}{\sqrt{n\Delta}}\right)$, and complete the proof.

5. Simulation and Numerical Computation for Target Estimator

In this section we simulate, $\hat{f}_n(t_k)$ on the data of size n by using Semlayt's wavelet. We consider convergence ratio of given estimator by computing of average mean square error of given estimators. We use R software and wavelet package for simulation.

Example 1: We generate $X_1, X_2, X_3, \cdots, X_n \sim \Gamma(5,1)$ and $C_1, C_2, C_3, \cdots, C_n \sim E(6)$ from the Samples of size $n = 400$ and $n = 600$ with $K = 8$, $K = 16$, $K = 32$ and $\Delta = 0.05$ for optimal surface $j = 2$.

The results in **Table 1** displays the average mean square errors of subdensity function estimator for sample sizes $n = 400$ and $n = 600$.

The panel in **Figure 1** displays the wavelet estimator of subdensity $\hat{f}_n(t_k)$ of observed failures for a traditional censoring data. The solid line is the density estimator and the dotted line is the true density.

Example 2: Suppose that $X_1, X_2, X_3, \cdots, X_n \sim f = 0.6Y + 0.4W$, where $Y \sim LN(0,1)$ and $W \sim N(3,.04)$. We generate $C_1, C_2, C_3, \cdots, C_n \sim E(3)$ from sample size of $n = 400$ and $n = 600$ with $K = 8$, $K = 16$, $K = 32$ and $\Delta = 0.05$.

The results in **Table 2** displays the average mean square errors of subdensity function estimator for sample sizes $n = 400$ and $n = 600$.

The panel in **Figure 2** displays the wavelet estimator of subdensity of observed failures for a traditional censoring data. The solid line displays the subdensity estimates based actual data and the dotted line is the true density.

Table 1. The average mean square errors of subdensity function estimator by wavelet method.

$\mathrm{AMSE}(f^*) = \dfrac{1}{K}\sum_{k=0}^{K}\left(\hat{f}_n(t_k) - f^*(t_k)\right)^2$		
K	$n = 400$	$n = 600$
8	26.1	17.9
16	19.2	10.1
32	18.6	7.2

Table 2. The average mean square errors of subdensity function estimator by wavelet method.

$\mathrm{AMSE}(f^*) = \dfrac{1}{K}\sum_{k=0}^{K}\left(\hat{f}_n(t_k) - f^*(t_k)\right)^2$		
K	$n = 400$	$n = 600$
8	680	610
16	420	275
32	379	278
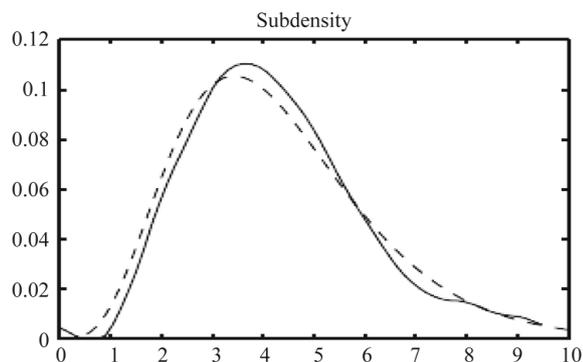
Figure 1. The wavelet subdensity and true density estimator.
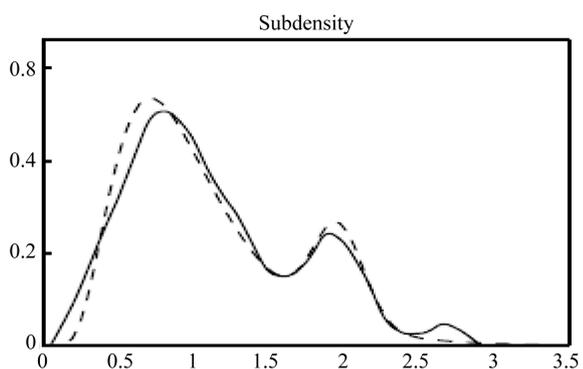
Figure 2. The wavelet subdensity and true density estimator.

6. Conclusion

In this paper we obtain density estimation for censoring data by using wavelet method and evaluate mean integral square error. We show that convergence ratio is acceptable and empirical distribution of given estimator under some condition is normal.

Acknowledgements

References

[1] Harr, A. (1910) Zur Theorie der Orthogonalen Funktionen. *Mathematische Annalen*, **69**, 331-371.

[2] Daubechies, I. (1988) Orthogonal Bases of Compactly Supported Wavelets. *Communication in Pure and Applied Mathematics*, **41**, 909-996.

[3] Antoniadis, A. (1996) Smoothing Noisy Data with Tapered Coiflets Series. *Scandinavian Journal of Statistics*, **23**, 313-330.

[4] Afshari, M. (2013) A Fast Wavelet Algorithm for Analyzing of Signal Processing and Empirical Distribution of Wavelet Coefficients with Numerical Example and Simulation. *Communication of Statistics-Theory and Methods*, **42**, 4156-4169.

[5] Afshari, M. (2014) Estimation of Hazard Function for Censoring Random Variable by Using Wavelet Decomposition and Evaluate of MISE, AMSE With Simulation. *Journal of Data Analysis and Information Processing*, **2**, 1-5. http://dx.doi.org/10.4236/jdaip.2014.21001

[6] Afshari, M. (2008) Wavelet-Kernel Estimation of Regression Function for Uniformly Mixing Process. *Word Applied Sciences Journal*, **4**, 605-609.

[7] Donoha, D.L. and Johnstone, I.M. (1994) Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika Journal*, **81**, 425-455. http://dx.doi.org/10.1093/biomet/81.3.425

[8] Kerkyacharian, G. and Picard, D. (1993) Density Estimation Bykernel and Probability. McGraw-Hill Science, New York, 327-336.

[9] Mallat, S.G. (1989) A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *Transformations on Pattern Analysis and Machine Intelligence*, **11**, 674-693.

[10] Meyer, Y. (1990) On de lettes et operateurs. Hermann, Paris.

[11] Hall, P. and Patil, P. (1995) Formula for Mean Integrated Squarederror of Non-Linear Wavelet Based Density Estimators. *Annals of Statistics*, **23**, 905-928. http://dx.doi.org/10.1214/aos/1176324628

[12] Antoniadis, A., Gregoire, G. and Nason, P. (1999) Density and Hazard Rate Estimation for Right Censored Data Using Wavelet Methods. *Journal of Royal Statistical Society Series B*, **23**, 313-330.

[13] Vidakovik, B. (1999) Statistical Modeling by Wavelets. Wiley, New York. http://dx.doi.org/10.1002/9780470317020

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.