Strong Convergence Rates of Wavelet Estimators in Semiparametric Regression Models with Censored Data*

Hongchang Hu*

School of Mathematics and Statistics, Hubei Normal University, Huangshi 435002, China

Abstract: The paper studies a semiparametric regression model

\[ Y_i = X_i \beta + g(T_i) + \epsilon_i, \quad i = 1, 2, \ldots, n. \]

where \( Y \) is censored on the right by another random variable \( C_i \) with known or unknown distribution \( G \). Firstly, the wavelet estimators of parameter \( \beta \) and nonparameter \( g(t) \) are given by wavelet smoothing and the synthetic data methods. Secondly, under general conditions, the laws of the iterated logarithm for the wavelet estimators of parameter and strong uniform convergence rates for the wavelet estimators of nonparameter are investigated. Lastly, the validity of method is illuminated by the simulation example.

Keywords: Semiparametric regression model, Wavelet estimate, Censored data, Law of the iterated logarithm, Strong uniform convergence rate.

1. INTRODUCTION

Consider the semiparametric regression model

\[ Y_i = X_i \beta + g(T_i) + \epsilon_i, \quad i = 1, 2, \ldots, n. \]  \hspace{1cm} (1)

where \( Y \)’s are scalar response variables, \( X \)’s and \( T \)’s are explanatory variables, \( \beta \) is a one-dimensional unknown parameter, \( g(t) \) is an unknown regression function on \([0,1]\), and \( \epsilon_i \)’s are independent and identically distributed random errors with zero mean and finite variance \( \sigma^2 \).

Following Speckman (See [1]), denote \( X_i = f(T_i) + \eta_i, \quad 1 \leq i \leq n. \) \hspace{1cm} (2)

where \( f(t) \) is some unknown smooth function on \([0,1]\). \( \eta_i \)’s are independent and identically distributed random errors with zero mean and finite variance \( \sigma^2_\eta \) and independent of \( Y \)’s and \( T \)’s.

Since the introductory work by Engle et al. (See [2]), the model (1) has been widely studied (See [1,3-6] and references therein) and put into use in many fields of applied statistics. A partial list of estimation methods for \( \beta \) and \( g(t) \) includes penalized least squares method (See [7]), smoothing splines method (See [8]), piecewise polynomial method (See [3]), near neighbor method (See [9]) and wavelet method (See [10-13]), etc.

In practice, particularly in medical studies, \( Y \) may be censored randomly on the right by some censoring variable \( C_i, i = 1, 2, \ldots, n. \) and hence cannot be observed completely. One only observes \( \{X_i, T_i, Z_i, \delta_i\}, i = 1, 2, \ldots, n. \) where \( Z_i = \min(\delta_i, C_i), \delta_i = I(Y_i \leq C_i), \) \( C_i \) are assumed to be independent of \( Y_i \)’s and \( \eta_i \)’s. For right censored data, the model (1) has been studied by Wang and Li [14] and Wang and Zheng [15], and the model (1)-(2) has been investigated by Qin and Cai [16], Qin and Jing [17], Pan and Fu [18] and Liang and Zhou [19].

With wavelet method, the paper discusses the model (1)-(2). The organization of the paper is as follows. In section 2 the wavelet estimators of parameter \( \beta \) and nonparameter \( g(t) \) are given by wavelet smoothing and the synthetic data methods. Under general conditions, the laws of the iterated logarithm for the wavelet estimators of parameter and strong uniform convergence rates for the wavelet estimators of nonparameter are investigated in section 3. The main proofs are presented in section 4, with the simulation example in section 5.

2. ESTIMATION METHOD

Let \( F(t) \) and \( G(t) \) denote the distributions for \( Y_i \) and \( C_i \), respectively. Assume that \( G(t) \) is continuous and known at the moment. Using the synthetic data method (See [16,17,20], and references therein), we transform the censored data \( \{(Z_i, \delta_i), i = 1, 2, \ldots, n\} \) into the following synthetic data

\[ Y_i' = \phi_1(Z_i; G)\delta_i + \phi_2(Z_i; G)(1-\delta_i) \hspace{1cm} (3) \]

where \( \phi_1, \phi_2 \) are continuous functions which satisfy

(i) \( (1-G(Y))\phi_1(Y; G) + \int_{0}^{Y} \phi_2(t; G) dG(t) = Y \),

(ii) \( \phi_1 \) and \( \phi_2 \) do not depend on \( F \).

The class of all pairs \( (\phi_1, \phi_2) \) of such functions will be denoted by \( K \). It can be easily showed that, if \( (\phi_1, \phi_2) \in K \), then

8-13 2008 Bentham Science Publishers Ltd.
\[ E(Y_i | X, T) = E(Y_i | X, T) \quad \text{and} \quad E(Y_i) = E(Y) , \]

where \( X = \{X_1, L, X_n\}, T = \{T_1, L, T_n\} \). The class \( \mathcal{K} \) includes many interesting cases (See [19–20]).

If \( G(\cdot) \) is unknown, we use a modified version \( \partial \beta(t) \) of the Kaplan-Meier estimator of \( G(t) \) defined by

\[
\partial \beta(t) = \begin{cases} \hat{G}\_n(t), & \text{if } t \leq Z_{(n)} \text{ or } \delta_{(n)} = 0 \\ \hat{G}\_n(Z_{(n)}), & \text{if } t > Z_{(n)} \text{ and } \delta_{(n)} = 1 \end{cases}
\]

where

\[
\hat{G}\_n(t) = 1 - \prod_{i \leq Z_{(n)}} \left(1 + N\_n(Z_i)\right)^{\delta_{(i)}},
\]

\( N\_n(t) = \sum_{i \leq Z_{(n)}} 1\{Z_i > t\} \), and \( \delta_{(n)} \) is the corresponding \( \delta_i \). It is easily showed that

\[
\partial \beta(t) \leq 1 - (n+1)^{-1} < 1
\]

for all \( t \). Substituting \( \partial \beta(t) \) for \( G(t) \) in (3), we obtain the following synthetic data:

\[
\hat{Y}_i = \phi_i(Z_i; \partial \beta) \delta_i + \phi_1(Z_i; \partial \beta)(1 - \delta_i)
\]

Suppose that there exists a scaling function \( \phi(e) \) in the Schwartz space \( S_t \) and a multiresolution analysis \( \{V_n\} \) in the concomitant Hilbert space \( L^2(\mathbb{R}) \), with its reproducing kernel \( E_n(t,s) \) given by

\[
E_n(t,s) = 2^n E_n(2^n t, 2^n s) = 2^n \sum_{k \in \mathbb{Z}} \phi(2^n t - k) \phi(2^n s - k)
\]

where \( Z \) denotes the collection of integers. Let \( A_i = [s_{i-1}, s_i] \) be intervals that partition \([0,1]\) with \( t_i \in A_i \) and \( 1 \leq i \leq n \).

The estimate method will be introduced as following:

Firstly, suppose that \( \beta \) is known, we define estimator of \( g(\cdot) \) by

\[
g_n(t, \beta) = \sum_{j=1}^{\infty} \left( Y_i - X, \beta \right) J_n(t, s) ds.
\]

In succession, we define wavelet estimator \( \beta_n \) by minimizing

\[
\beta_n = \arg \min \left\{ \sum_{j=1}^{\infty} \left| \sum_{i=1}^{\infty} (Y_i - X, \beta) \right|^2 \right\},
\]

where

\[
X_{\infty} = X_i - \sum_{j=1}^{\infty} J_n(t, s) ds X_j, \quad Y_{\infty} = Y_i - \sum_{j=1}^{\infty} J_n(t, s) ds Y_j.
\]

Finally, we define linear wavelet estimator of \( g(\cdot) \) by

\[
g_n(t, \beta) = \sum_{j=1}^{\infty} \left( Y_i - X, \beta \right) J_n(t, s) ds.
\]

Using the similar procedure, if \( G(\cdot) \) is unknown, then we define estimators of \( \beta \) and \( g(\cdot) \) by

\[
\hat{\beta}_n = \sum_{j=1}^{\infty} \hat{X}_{\infty} \hat{Y}_{\infty} J_n(t, s) ds, \quad \hat{g}_n(t) = \sum_{j=1}^{\infty} \hat{X}_{\infty} \hat{Y}_{\infty} J_n(t, s) ds.
\]

To obtain our results, the following conditions are sufficient:

(1) \( g(\cdot) \) and \( f(\cdot) \) belong to the Sobolev space with order \( \alpha > \gamma / 2 \).

(2) \( g(\cdot) \) and \( f(\cdot) \) satisfy the Lipschitz condition with order \( \gamma > 0 \).

(3) \( \phi(\cdot) \) is in the Schwartz space with order \( L < \alpha \), satisfies the Lipschitz condition with order 1 and has a compact support. Furthermore, \( \left| \phi(\xi) - 1 \right| = O(\xi) \) as \( \xi \to 0 \), where \( \phi \) is the Fourier transform of \( \phi \).

(4) \( \max_{i} |s_i - s_{i-1}| = O(\gamma^{-1}) \) and \( n = O(\gamma^{-3}) \).

(5) \( F(\cdot) \) and \( G(\cdot) \) are continuous functions. In the case \( G(t_f) < 1 \), assume that

\[
\int_{t}^{t_f}(1 - F(u)) \leq G(s) < \infty,
\]

where \( t_f = \inf \{ t; F(t) = 1 \} \). In the case \( G(t_f) = 1 \), assume that

\[
(1 - G(u))^\gamma = O\left(1 - F(u)\right) \quad \text{as} \quad u \uparrow t_f,
\]

for some \( \gamma \in (0,1) \).

In the following discussion, we shall confine ourselves to some more restricted classes of \( (\phi_1, \phi_2) \) than the class \( \mathcal{K} \). In particular, we define

\[
K^* = \{ (\phi_1, \phi_2) \in \mathcal{K} : \text{for each } s \leq \tau_g, \text{there exists a constant } 0 < C < \infty \text{ such that} \}
\[
\sup_{r \in \mathbb{R}} \left\| \phi_j (t; G) \right\| \leq C, \quad j = 1, 2 \}
\]

and

\[
\hat{\mathcal{K}} = \{ (\phi_1, \phi_2) \in \mathcal{K} : \text{there exist constants } 0 < L = L(s) < \infty \text{ and} \}
\[
\delta > 0 \text{ such that} \quad \sup_{r \in \mathbb{R}} \left\| \phi_j (t; \hat{\mathcal{K}}) - \phi_j (t; G) \right\| \leq L_{\sup} \left\| \mathcal{K} (t; G) - \mathcal{K} (t) \right\|, \quad j = 1, 2,
\]

for all d.f. \( \hat{\mathcal{K}} \) with \( \sup_{r \in \mathbb{R}} \left\| \mathcal{K} (t; G) - \mathcal{K} (t) \right\| < \delta \).

For ease of exposition, we shall introduce the following notations which will be used later in the paper. Define
3. STATEMENT OF THE RESULTS

Theorem 3.1. Suppose that conditions (A1) - (A3) hold. \(E|\epsilon_i| < \infty, E|\eta_i| < \infty\) and \((\phi_i, \phi_i') \in \mathcal{K}^*\). Then for every \(\gamma > 0\),
\[
\lim_{n \to \infty} \sup_{\alpha > 3/2} \left\{ n(2\log \log n)^{1/2} \beta_n - \beta \right\} = O(n^{1/2} \log n) \quad a.s. \tag{6}
\]
\[
\lim_{n \to \infty} \sup_{\alpha > 3/2} \left\{ n(2\log \log n)^{1/2} \beta_n - \beta \right\} = O(n^{1/2} \log n) \quad a.s. \tag{7}
\]

Theorem 3.2. Suppose that conditions (A1) - (A3) hold. \(E|\epsilon_i| < \infty, E|\eta_i| < \infty\) and \((\phi_i, \phi_i') \in \mathcal{K}^*\). Then for every \(\gamma > 0\),
\[
\lim_{n \to \infty} \sup_{\alpha > 3/2} \left\{ n(2\log \log n)^{1/2} \beta_n - \beta \right\} = O(n^{1/2} \log n) \quad a.s. \tag{8}
\]
\[
\lim_{n \to \infty} \sup_{\alpha > 3/2} \left\{ n(2\log \log n)^{1/2} \beta_n - \beta \right\} = O(n^{1/2} \log n) \quad a.s. \tag{9}
\]

4. PROOFS OF THEOREMS

Throughout this paper, let \(C\) denote a generic positive constant which could take different value at each occurrence. Before the proofs of the theorems, we introduce some preliminary results.

Lemma 4.1. (Antoniadis et al. [21]). If condition (A3) holds, then

(i) \(|E_n(t, s)| = n C / (1 + n - s)|^t\) and \(|E_n(t, s)| = n C / (1 + n - s)|^t\)

for \(k \in \mathbb{N}\), where \(C_k\) is a real constant depending on \(k\).

(ii) \(\sup_{t \in \mathbb{R}} \left| \frac{E_n(t, s)}{E_n(t, s)} \right| = O(n^{1/2})\)

(iii) \(\sup_{t \in \mathbb{R}} \left| \int_{t}^{t+1} E_n(t, s) \, ds \right| \leq C\).

Lemma 4.2. (Hu and Hu [22]). Let \(\tau_n = 2^{-\alpha(k-1/2)}\) when \(1/2 < \alpha < 3/2\), \(\tau_n = \sqrt{m} - 2^{-m}\) when \(\alpha = 3/2\), \(\tau_n = 2^{-n}\) when \(\alpha > 3/2\). If conditions (A1) - (A4) hold, then
\[
\sup_{t \in \mathbb{R}} \left| \int_{t}^{t+1} E_n(t, s) \, ds \right| = O(n^{-\alpha}) + O(\tau_n)
\]
\[
\sup_{t \in \mathbb{R}} \left| \int_{t}^{t+1} g_n(t, s) \, ds \right| = O(n^{-\alpha}) + O(\tau_n)
\]

Lemma 4.3. (Qian and Cai [13]). If conditions (A1) - (A3) hold, and \(E|v|^2 \delta < \infty\) for some \(\delta > 0\), then
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \eta_i = \alpha_n^2 a.s.
\]

Lemma 4.4. (Qian and Cai [13]). If conditions (A1) - (A3) hold, and \(E|\zeta|^2 \delta < \infty\), then
\[
\max_{i \in \mathbb{N}} \left| \sum_{j=1}^{i} E_n(t, s) \right| = O(1) a.s.
\]

Lemma 4.5. (Xue [11]). If conditions (A1) - (A4) hold, then
\[
\max_{i \in \mathbb{N}} \left| \sum_{j=1}^{i} E_n(t, s) \right| = O(1) a.s.
\]

Lemma 4.6. (Gu and Lai [23]). Assume condition (A5) holds. If \(G(t) < 1\), let
\[
\tau_n = \sup \left\{ t : 1 - G(t) = n^{-1/2} \right\}
\]

where \(1/3 < \kappa < 1/2\), then
\[
\lim_{n \to \infty} \sup_{i \in \mathbb{N}} \left| \hat{g}(i) - G(i) \right| = \sup_{i \in \mathbb{N}} \left| \hat{S}(i) \sigma^{-1}(i) \right| a.s.
\]
If \(G(t) \geq 1\), then
\[
\lim_{n \to \infty} \sup_{i \in \mathbb{N}} \left| \hat{g}(i) - G(i) \right| = \sup_{i \in \mathbb{N}} \left| \hat{S}(i) \sigma^{-1}(i) \right| a.s.
\]
where
\[
\hat{S}(t) = 1 - G(t) \cdot \sigma(t) = \int_{t}^{t+1} \left( 1 - G(s) \right) (1 - F(s))^{-1} \, dG(s)
\]

Lemma 4.7. Under the conditions of theorem 1, we have
\[
n^{-\alpha/2} \left( \beta_n - \beta \right) = o(t) a.s.
\]

The proof is similar to that of lemma 1 in [17] and hence omitted here.

Lemma 4.8. (Hardle, Liang and Gao [24]). Let \(E\xi_i = 1, i = 1, 2, \ldots, n\) be independent random variables with finite variance, and \(\sup_{i \in \mathbb{N}} \left| E\xi_i \right| \leq C < \infty (r \in (0, 2))\).
Assume that \(\{v_i, i = 1, 2, \ldots, n\} \) is a sequence of real numbers such that \(\sup_{i \in \mathbb{N}} |v_i| = O(n^{-p})\) for some \(0 < p_1 < 1\) and
\[
\sum_{i=1}^{n} a_i = O(n^{p_2}) \quad \text{for} \quad p_2 \geq \max \left\{ 0, 2/r - p_1 \right\}
\]
Then
\[
\max_{1 \leq j \leq n} \left| \sum_{i=1}^{n} a_i \xi_{ij} \right| = O(n^{-s} \log n) \quad \text{for } s = (p_1 - p_2)/2 \quad \text{a.s.}
\]

**Lemma 4.9.** Let \( \{\xi_{ij}, i = 1, 2, \ldots, n\} \) be independent random variables with \( E \xi_{ij} = 0 \) and finite variances, and \( \sup_{i=1}^{n} |\xi_{ij}| \leq C < \infty \) (\( r \geq 3 \)). Assume that conditions (A1)\textendash (A4) hold. Then
\[
\max_{1 \leq j \leq n} \sum_{i=1}^{n} \int_{\mathbb{A}} E_{n}(t, s)ds \xi_{ij} = O(n^{-1/2}) \quad \text{a.s.}
\]

**Proof.** Let \( r = 3, p_1 = 2/3, p_2 = 0 \) and \( a_{ij} = \int_{\mathbb{A}} E_{n}(t, s)ds \).

The lemma follows from lemma 4.1, lemma 4.5 and lemma 4.8.

**Proof of Theorem 3.1.** Note that
\[
\beta_n^* - \beta = \left( \sum_{i=1}^{n} \hat{\xi}_{ii} \right) (I_1 + I_2)
\]
where
\[
I_1 = \sum_{i=1}^{n} \hat{\xi}_{ii} = E_\hat{\xi} = E \left( \hat{\xi}^T \hat{X}, \hat{T} \right)
\]
and then
\[
I_2 = \sum_{i=1}^{n} \hat{\xi}_{ii} - \sum_{i=1}^{n} \xi_{ii} = \sum_{i=1}^{n} \eta_{ii} = \sum_{i=1}^{n} \int_{\mathbb{A}} E_{n}(t, s)ds \xi_{ii}.
\]

Hence from lemma 4.1, lemma 4.5 and lemma 4.8, we obtain that
\[
\beta_n^* - \beta = n \left( \sum_{i=1}^{n} \hat{\xi}_{ii} \right) (n^{-1/2} + o(1)) \quad \text{as } \text{LIL of Hartman-Winter.}
\]

By Bernstein’s inequality and Bore-Cantelli lemma,
\[
\sum_{i=1}^{n} |\xi_{ij}| \leq \varepsilon n^{1/2}, \quad \text{a.s.}
\]

From (13) and (16), we have that
\[
I_1 = O(n^{1/2}), \quad \text{a.s.}, \quad I_2 = O(n^{1/2}), \quad \text{a.s.}
\]

Using the similar argument as above, by lemma 4.2, 4.4 and 4.5, we obtain that

\[
I_3 = O(n^{1/2}), \quad \text{a.s.}, \quad I_4 = O(n^{1/2}), \quad \text{a.s.}, \quad I_5 = O(n^{1/2}), \quad \text{a.s.}
\]

From lemma 4.2, it follows that
\[
I_2 \leq n \sup_{t} \left| \hat{\xi}_{t} \right| = O(n^{1/2} + O(\varepsilon n^{1/2})).
\]

From (10)-(12) and (17)-(19),
\[
n^{1/2} (\beta_n^* - \beta) = n \left( \sum_{i=1}^{n} \hat{\xi}_{ii} \right) n^{-1/2} + o(1)) \quad \text{as } \text{LIL of Hartman-Winter.}
\]

The Open Applied Mathematics Journal, 2008, Volume 2
From $E|e_i| < \infty$ and lemma 2 in [16], we have that $E|\hat{e}_i| < \infty$. Hence, by lemma 4.9,

$$I_1 = O\left(n^{-1/3} \log n \right) \text{ a.s.} \quad (24)$$

Thus (7) follows from (21)-(24). We therefore complete the proof of Theorem 3.1.

**Proof of Theorem 3.2** Note that

$$n^{1/2} \left( \hat{\beta}_n^* - \beta \right) = n^{1/2} \left( \hat{\beta}_n^* - \beta_n^* \right) + n^{1/2} \left( \beta_n^* - \beta \right)$$

The (8) follows immediately from lemma 4.7 and theorem 3.1.

Write

$$\hat{g}_n^*(t) - g(t) = \left( \hat{g}_n^*(t) - g_n^*(t) \right) + \left( g_n^*(t) - g(t) \right)$$

$$I_1 = \sum_{t \in A} E_n(t) \left( \hat{g}_n^*(t) - g_n^*(t) \right) + \sum_{t \in A} \left( g_n^*(t) - g(t) \right)$$

$$I_{11} = \sum_{t \in A} E_n(t) \left( \hat{g}_n^*(t) - g_n^*(t) \right) + \sum_{t \in A} E_n(t) \left( g_n^*(t) - g(t) \right)$$

$$+ \sum_{t \in A} E_n(t) \left( g_n^*(t) - g(t) \right)$$

$$= R_1 + R_2 \quad (25)$$

By lemma 4.1 and 4.6, we have that

$$\left| R_i \right| \leq \sup_{t \in A} \left| \hat{g}_n^*(t) - G(t) \right|$$

$$\leq C \left( n^{-1/2} \log n \right) \text{ a.s.} \quad (28)$$

In the case $G(z) > 1$, max $Z \in [a,b]$ a.s. By lemma 4.1 and 4.6,

$$\sup_i \left| R_i \right| \leq \sup_i \left[ \sum_{t \in A} E_n(t) \left| \hat{g}_n^*(t) - G(t) \right| \right]$$

$$\leq C \left( n^{-1/2} \log n \right) \text{ a.s.} \quad (29)$$

In the case $G(z) < 1$. Let $z_i = I\left( Z > \tau \right)$. Hence from lemma 4.9, we have that

$$\sup_i \left| R_i \right| \leq \sup_i \left| \hat{g}_n^*(t) - g(t) \right|$$

$$\leq C \left( n^{-1/2} \log n \right) \text{ a.s.} \quad (30)$$

Therefore the desired conclusion (9) follows from (25)-(30) and theorem 3.1.

### 5. Numerical Example

We will simulate a simple semiparametric regression model

$$y_i = X_i \beta + \cos(2\pi T_i) + e_i, \ i = 1, 2, \ldots, 64$$

where $\beta = 1, T_i = i/64, X_i = 5T_i^2 + \eta_i, \ \eta_i = N(0, 0.25)$.

Choose $\phi(t, u) = u + G(\theta(u))$ and Daubechies scaling function $\phi(t)$. By calculation, we have $\beta_n^* = 1.0158$. It closely approximates the true value of parameter $\beta$. It can be seen that our method is successful, especially in estimating the parameter. However, a further discussion of the choices the scaling function and $(\phi, \phi)$ is needed so that we can find a good method to use in practical applications.

**Acknowledgments**

*Supported by Natural Science Foundation of Education Office, Hubei (No.Q200622001).

**References**


