QUANTILE ESTIMATION FOR LEFT TRUNCATED AND RIGHT CENSORED DATA

Xian Zhou, Liuquan Sun and Haobo Ren

Hong Kong Polytechnic University, Academia Sinica and Peking University

Abstract: In this paper we study the estimation of a quantile function based on left truncated and right censored data by the kernel smoothing method. Asymptotic normality and a Berry-Esseen type bound for the kernel quantile estimator are derived. Monte Carlo studies are conducted to compare the proposed estimator with the PL-quantile estimator.

Key words and phrases: Asymptotic normality, Berry-Esseen type bound, quantile function, kernel estimation, truncated and censored data.

1. Introduction

In medical follow-up or in engineering life testing studies one may not be able to observe the variable of interest, referred to hereafter as the lifetime. Among the different forms in which incomplete data appear, right censoring and left truncation are common. Left truncation may occur if the time origin of the lifetime precedes the time origin of the study. Only those individuals that fail after the start of the study are being followed, otherwise they are left truncated. Those that are followed are further subject to right censoring during the follow-up period. See for example, an AIDS study by Struthers and Farewell (1989) where “lifetime” is the incubation period, and the truncation variable is the time from infection until entry to the study. Formally, let \((X, T, Y)\) denote random variables, where \(X\) is the lifetime with continuous distribution function (d.f.) \(F\), \(T\) is the random left truncation time with arbitrary d.f. \(G\), and \(Y\) is the random right censoring time with arbitrary d.f. \(H\). It is assumed that \(X, T, Y\) are mutually independent and, without loss of generality, that they are nonnegative. In the random left truncation and right censoring (LTRC) model one observes \((Z, T, \delta)\) if \(Z \geq T\), where \(Z = X \wedge Y = \min(X, Y)\) and \(\delta = I(X \leq Y)\) is the indicator of censoring status. When \(Z < T\) nothing is observed. Let \(\alpha = P(T \leq Z) > 0\), and let \(W\) be the d.f. of \(Z\), i.e., \(1 - W = (1 - F)(1 - H)\). Let \((Z_i, T_i, \delta_i), i = 1, \ldots, n\) be an independent and identically distributed (i.i.d.) sample of \((Z, T, \delta)\) which is observed (i.e., \(Z_i \geq T_i\)). Let \(C(x) = P(T \leq x \leq Z | T \leq Z) = \alpha^{-1}G(x)(1 - W(x-))\), and consider its empirical estimate \(C_n(x) = n^{-1} \sum_{i=1}^{n} I(T_i \leq x \leq Z_i)\).
For LTRC data, it is important to be able to obtain nonparametric estimates of various characteristics of the distribution function \( F \). Tsai, Jewell and Wang (1987) gave the nonparametric maximum likelihood estimator of \( F \) itself, called the product-limit (PL) estimator, as

\[
\hat{F}_n(x) = \begin{cases} 
1 - \prod_{Z_i \leq x} (1 - [nC_n(Z_i)]^{-1}) & , x < Z_{(n)}, \\
1 & , x \geq Z_{(n)},
\end{cases}
\]

(1.1)

where \( Z_{(n)} = \max(Z_1, \ldots, Z_n) \). The properties of \( \hat{F}_n \) have been studied by Wang (1987), Gu and Lai (1990), Lai and Ying (1991), Gijbels and Wang (1993), and Zhou (1996), among others. Nonparametric estimates of the density and hazard rate for \( F \) have been studied by Gijbels and Wang (1993), Sun and Zhou (1998), Uzunoğullari and Wang (1992), Gu (1995) and Sun (1997).

One characteristic of the distribution function \( F \) that is of interest is the quantile function. It plays an important role in various statistical applications, especially in data modeling, reliability and medical studies. Let \( Q(t) \equiv F^{-1}(t) = \inf \{ x : F(x) \geq t \} \), \( 0 < t < 1 \), be the quantile function of \( F \). We focus here on estimating the quantile function based on LTRC data.

A natural estimator of \( Q \) is the PL-quantile function defined as \( \hat{F}_n^{-1}(x) = \inf \{ u : \hat{F}_n(u) \geq x \}, 0 < x < 1 \). In case \( Q \) is a continuous function, it may be more suitable to use a smooth estimator rather than the step function \( \hat{F}_n^{-1} \), since smoothing reduces random variation in the data and allows a better display of interesting features of the lifetime distribution function. Numerous smooth quantile function estimators have been proposed for complete samples. In the censored model, Padgett (1986), Lio, Padgett and Yu (1986), Xiang (1995a, b) studied kernel-type quantile function estimators. None of their results included the case where both left truncation and right censoring are involved however. On the other hand, Gürler, Stute and Wang (1993) provide a strong representation of the empirical quantile function for left truncated data, together with asymptotic normality and the Law of Iterated Logarithm (LIL), which can be easily extended to the LTRC model. In this paper, for the LTRC model, we propose and study the kernel quantile function estimator of the form

\[
\hat{Q}_n(t) = h_n^{-1} \int_0^1 \hat{F}_n^{-1}(x)K \left( \frac{x - t}{h_n} \right) dx,
\]

(1.2)

where \( K \) is a kernel function and \( \{ h_n \}_{n \geq 1} \) is a sequence of positive bandwidths with \( h_n \to 0 \) as \( n \to \infty \).

The first aim of this article is to derive the asymptotic normality of the kernel quantile function estimator. The quality of the approximation to the distribution
of $\hat{Q}_n$ is an important issue of both theoretical and practical interest (for example, construction of confidence intervals based on the asymptotic normality). Thus we also establish a Berry-Esseen type bound of $\hat{Q}_n$.

A small Monte Carlo simulation study shows that, for certain values of the bandwidth $h_n$, $\hat{Q}_n$ performs better than $\hat{F}_n^{-1}$ in the sense of having smaller estimated mean squared errors (MSE).

For any d.f. $L$, denote the left and right endpoints of its support by $a_L = \inf\{x : L(x) > 0\}$ and $b_L = \sup\{x : L(x) < 1\}$, respectively. In the current model, as discussed by Gijbels and Wang (1993) and Zhou (1996), we assume that $a_G \leq a_W, b_G \leq b_W$ and

$$\int_{a_W}^{\infty} \frac{dF(x)}{G^2(x)} < \infty. \quad (1.3)$$

Condition (1.3) requires that the left truncation is not too heavy. As for the kernel function $K(x)$, assume

(K1) $K$ is of bounded variation with support $[-1,1]$;

(K2) $K$ is Lipschitz continuous of order one;

(K3) For some integer $r \geq 2$,

$$\frac{1}{j!} \int x^j K(x) dx = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{if } j = 1, \ldots, r-1, \\ c_r & \text{if } j = r. \end{cases}$$

These assumptions are the usual ones encountered in the kernel method of curve estimation, as in Padgett (1986) and Uzunogullari and Wang (1992).

This article is organized as follows. Main results are stated in Section 2. Some simulation results are presented in Section 3. Proofs of the results are deferred to the Appendix.

2. Main Results

We first give the asymptotic normality of $\hat{Q}_n$.

**Theorem 1.** Under (K1) and (K3), assume that $F$ is $r$ times continuously differentiable in a neighborhood of $Q(t)$ with $F'(Q(t)) = f(Q(t)) > 0$, $a_w < Q(t) < b_w$, where $r$ is the same as that in (K3). If $n^{1/2} h_n^r \to 0$ as $n \to \infty$,

$$n^{1/2}(\hat{Q}_n(t) - Q(t)) \to N(0, \sigma^2(t)) \quad \text{in distribution as } n \to \infty, \quad (2.1)$$

where the variance is

$$\sigma^2(t) = \frac{(1-t)^2}{f^2(Q(t))} \int_{a_W}^{Q(t)} \frac{dF(x)}{C(x)(1-F(x))}. \quad (2.2)$$
To apply this result to make some standard inferences in hypothesis testing and construction of confidence intervals, an estimator of the variance in (2.2) is needed and a natural nonparametric estimator is given below. Let \( W_1(x) = P(Z \leq x, \delta = 1| T \leq Z) = \alpha^{-1} \int_{-\infty}^{x} G(u)(1 - H(u-))dF(u) \). Then it can be checked that
\[
\sigma^2(t) = \frac{(1 - t)^2}{f^2(Q(t))} \int_{a_W}^{Q(t)} \frac{dW_1(x)}{C^2(x)}. \quad (2.3)
\]
Thus \( \sigma^2(t) \) can be estimated consistently by replacing the density \( f \) in (2.3) by a density estimate from Gijbels and Wang (1993), and all other unknown quantities in (2.3) by their empirical estimates.

We now give a Berry-Esseen type bound for \( \hat{Q}_n \) to assess the quality of the normal approximation in Theorem 1.

**Theorem 2.** Under (K1), (K2) and (K3), assume that \( F \) is \( r \)-time continuously differentiable in a neighborhood of \( Q(t) \) with \( F'(Q(t)) = f(Q(t)) > 0 \), \( a_W < Q(t) < b_W \). If \( a_G < a_W \) and \( nh_n^2/\log n \to \infty \) as \( n \to \infty \), there exists a positive constant \( M = M(t) \) such that for all \( n \geq 1 \),
\[
\sup_y \left| P\{n^{1/2}(\hat{Q}_n(t) - Q(t)) \leq y\sigma(t)\} - \Phi(y) \right| \leq M \left( n^{-1/2}h_n^{-1} \log n + n^{1/2}h_n^{1/2} + h_n^{1/2} \right),
\]  
\[
(2.4)
\]
where \( \sigma(t) \) is defined by (2.2).

3. Simulation Studies

A small Monte Carlo study was performed in order to provide some small-sample comparisons of the kernel estimator \( \hat{Q}_n \) with the PL-quantile \( \hat{F}_n^{-1} \) in terms of mean squared error. The study also provides some insight into the choices of \( h_n \) that might be used to estimate \( Q \) with smaller mean squared error than \( \hat{F}_n^{-1} \).

For the LTRC model, the following cases are simulated.

(i) \( F(x) = 1 - \exp(-x), H(x) = 1 - \exp(-x) \) and \( G(x) = 1 - \exp(-2x) \), with 50% censoring and 50% truncation respectively.

(ii) \( F(x) = 1 - \exp(-x), H(x) = 1 - \exp(-3x/7) \) and \( G(x) = 1 - \exp(-30x/7) \), with 30% censoring and 25% truncation respectively.

The triangular density function \( K(x) = (1 - |x|)I(|x| \leq 1) \) was used as the kernel function for the estimator \( \hat{Q}_n \). The ratio of the mean squared error of \( \hat{F}_n^{-1}(t) \) to that of the smoothed estimator \( \hat{Q}_n(t) \) were computed for various values of \( t \in (0, 1) \) and sample size \( n = 100 \). For each case, 1000 replications were done on S-PLUS on a PC-pentium 200 computer. The simulations were run for various values of \( h_n \) and the results show that \( \hat{Q}_n \) gives reasonable performance for \( h_n \leq 0.40 \) only. On the other hand, the results for \( h_n \) below 0.10 do not vary.
much. Some of the simulation results for \( h_n \) between 0.05 and 0.40 are presented in Tables 1-2 below, and these show that, for each value \( t \) listed, there is a range of window widths \( h_n \) such that the estimator \( \hat{Q}_n \) has smaller estimated mean squared error than the PL-quantile estimator. In particular, this is true for the median estimators \( \hat{Q}(0.5) \) and \( \hat{F}^{-1}(0.5) \).

From a range of \( h_n \) values, we found that \( h_n = 0.15 \) comes close to giving the smallest discrepancy between the kernel estimator and the quantile. Figures 1-2 show the plots of the quantile, kernel estimator \( \hat{Q}_n \) and the PL-quantile estimator \( \hat{F}^{-1} \) with \( h_n = 0.15 \), from which we see that \( \hat{Q}_n \) looks better than \( \hat{F}^{-1} \) except for large values of \( t \).

As one would expect for censoring, truncation and bound effect, the performance of either estimator (\( \hat{Q}_n \) or \( \hat{F}^{-1} \)) at large values of \( t \) is not as good as that for values near 0.5.

Table 1. Ratios of mean squared errors for Case (i): 50% censoring and 50% truncation (\( n = 100 \)).

<table>
<thead>
<tr>
<th>( h_n )</th>
<th>0.05</th>
<th>0.07</th>
<th>0.09</th>
<th>0.12</th>
<th>0.15</th>
<th>0.20</th>
<th>0.25</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.056</td>
<td>1.086</td>
<td>1.095</td>
<td>1.128</td>
<td>1.185</td>
<td>1.243</td>
<td>1.225</td>
<td>1.144</td>
<td>1.023</td>
<td>0.886</td>
</tr>
<tr>
<td>0.25</td>
<td>1.054</td>
<td>1.065</td>
<td>1.076</td>
<td>1.094</td>
<td>1.114</td>
<td>1.149</td>
<td>1.183</td>
<td>1.216</td>
<td>1.236</td>
<td>1.224</td>
</tr>
<tr>
<td>0.50</td>
<td>1.077</td>
<td>1.082</td>
<td>1.098</td>
<td>1.103</td>
<td>1.129</td>
<td>1.140</td>
<td>1.151</td>
<td>1.138</td>
<td>1.046</td>
<td>0.997</td>
</tr>
<tr>
<td>0.75</td>
<td>1.126</td>
<td>1.149</td>
<td>1.153</td>
<td>1.145</td>
<td>1.132</td>
<td>1.101</td>
<td>1.044</td>
<td>1.134</td>
<td>1.410</td>
<td>1.654</td>
</tr>
<tr>
<td>0.90</td>
<td>1.353</td>
<td>1.433</td>
<td>1.517</td>
<td>1.724</td>
<td>1.897</td>
<td>1.431</td>
<td>0.926</td>
<td>0.749</td>
<td>0.594</td>
<td>0.401</td>
</tr>
<tr>
<td>0.95</td>
<td>1.138</td>
<td>1.040</td>
<td>0.907</td>
<td>0.642</td>
<td>0.499</td>
<td>0.383</td>
<td>0.326</td>
<td>0.292</td>
<td>0.269</td>
<td>0.254</td>
</tr>
</tbody>
</table>

Note: Ratio=MSE(\( \hat{F}^{-1} \))/MSE(\( \hat{Q}_n \)).

Table 2. Ratios of mean squared errors for Case (ii): 30% censoring and 25% truncation (\( n = 100 \)).

<table>
<thead>
<tr>
<th>( h_n )</th>
<th>0.05</th>
<th>0.07</th>
<th>0.09</th>
<th>0.12</th>
<th>0.15</th>
<th>0.20</th>
<th>0.25</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.051</td>
<td>1.065</td>
<td>1.079</td>
<td>1.104</td>
<td>1.166</td>
<td>1.244</td>
<td>1.256</td>
<td>1.207</td>
<td>1.118</td>
<td>1.008</td>
</tr>
<tr>
<td>0.25</td>
<td>1.048</td>
<td>1.056</td>
<td>1.065</td>
<td>1.082</td>
<td>1.099</td>
<td>1.131</td>
<td>1.159</td>
<td>1.185</td>
<td>1.203</td>
<td>1.195</td>
</tr>
<tr>
<td>0.50</td>
<td>1.029</td>
<td>1.039</td>
<td>1.048</td>
<td>1.059</td>
<td>1.067</td>
<td>1.075</td>
<td>1.078</td>
<td>1.067</td>
<td>1.040</td>
<td>0.989</td>
</tr>
<tr>
<td>0.75</td>
<td>1.034</td>
<td>1.043</td>
<td>1.051</td>
<td>1.057</td>
<td>1.054</td>
<td>1.014</td>
<td>0.987</td>
<td>0.867</td>
<td>1.039</td>
<td>1.260</td>
</tr>
<tr>
<td>0.90</td>
<td>1.062</td>
<td>1.043</td>
<td>0.996</td>
<td>1.022</td>
<td>1.235</td>
<td>0.979</td>
<td>0.657</td>
<td>0.456</td>
<td>0.356</td>
<td>0.301</td>
</tr>
<tr>
<td>0.95</td>
<td>1.163</td>
<td>1.479</td>
<td>1.485</td>
<td>0.960</td>
<td>0.622</td>
<td>0.507</td>
<td>0.422</td>
<td>0.352</td>
<td>0.279</td>
<td>0.235</td>
</tr>
</tbody>
</table>

Note: Ratio=MSE(\( \hat{F}^{-1} \))/MSE(\( \hat{Q}_n \)).
Figure 1. Plots of PL-quantile and the kernel estimator for case (i) with $n = 50$ and $h_n = 0.15$.

Figure 2. Plots of PL-quantile and the kernel estimator for case (ii) with $n = 50$ and $h_n = 0.15$. 
Appendix. Proofs of Theorems

First we present preliminary results needed in the proofs of the theorems. Let
\[ L_n(x) = \int_{a_W}^{x} \frac{d|W_{1n}(u) - W_1(u)|}{C(u)} - \int_{a_W}^{x} \frac{C_n(u) - C(u)}{C^2(u)} dW_1(u). \]
From Theorem 1 of Gijbels and Wang (1993) and Theorem 2 of Zhou (1996), we have that for \( a_W \leq x \leq b < b_W \),
\[ \hat{F}_n(x) - F(x) = (1 - F(x))L_n(x) + R_n(x), \tag{A.1} \]
where
\[ \sup_{a_W \leq x \leq b} |R_n(x)| = O(n^{-1}\log^{1+\beta} n) \quad \text{a.s.} \tag{A.2} \]
with \( \beta = 0 \) when \( a_G < a_W \), and \( \beta > 1/2 \) when \( a_G = a_W \).

Introduce the PL-process \( \hat{\alpha}_n(x) = n^{-1/2}(\hat{F}_n(x) - F(x)) \). The following lemma, of interest in its own right, provides a general version of the oscillation behavior of \( \hat{\alpha}_n \).

**Lemma A.1.** Assume that \( F \) is Lipschitz continuous of order one on \([c, d]\), \( a_W < c \leq d < b_W \). Let \( \{a_n, n \geq 1\} \) be a sequence of positive constants tending to zero with \( na_n/\log n \to \infty \) as \( n \to \infty \). Then
\[ \sup_{c \leq x, y \leq d} |\hat{\alpha}_n(x) - \hat{\alpha}_n(y)| = O\left(n^{-1/2}(\log n)^{1/2}\right) + O\left(n^{-1/2}(\log n)^{1+\beta}\right) \quad \text{a.s.}, \tag{A.3} \]
where \( \beta \) is defined in (A.2).

**Proof.** It follows from (A.1) that
\[
n^{-\frac{1}{2}}(\alpha_n(x) - \alpha_n(y)) = (1 - F(x)) \int_{y}^{x} \frac{d|W_{1n}(u) - W_1(u)|}{C(u)} + (F(y) - F(x))L_n(y) - (1 - F(x)) \int_{y}^{x} \frac{C_n(u) - C(u)}{C^2(u)} dW_1(u) + (R_n(x) - R_n(y)) := \sum_{i=1}^{4} \Gamma_{ni}(x, y). \tag{A.4} \]
First, by the continuity of \( F \) together with a partitioning argument similar to that in Burke, Csörgő and Horváth (1988), it can be shown that
\[ \sup_{c \leq x, y \leq d, |x-y| \leq a_n} |\Gamma_{n1}(x, y)| = O\left(n^{-1/2}a_n^{1/2}(\log n)^{1/2}\right) \quad \text{a.s.} \tag{A.5} \]
Next, in view of (1.3), the process \( L_n(y) \) is the sum of two empirical processes over VC classes of functions with square integrable envelope on \([c, d]\), so it satisfies
Let \( \bar{\phi}(t) \) be the PL-quantile function. The proof of Theorem 1 relies on the weak asymptotic representation for the PL-quantile function \( \hat{F}_n^{-1} \) as given in (A.8) below.

Let \( a_W < Q(t_1) \leq Q(t_2) < b_W \). If \( F \) has a continuous and positive density \( f \) on \([Q(t_1) - \eta, Q(t_2) + \eta]\) for some \( \eta > 0 \), it follows from similar arguments as in Güler, Stute and Wang (1993) that, as \( n \to \infty \),

\[
\sup_{t_1 \leq t \leq t_2} \left| \hat{F}_n^{-1}(t) - Q(t) - \frac{t - \hat{F}_n(Q(t))}{f(Q(t))} \right| = o_p(n^{-1/2}). \tag{A.8}
\]

Let \( \hat{Q}_n(t) = h_n^{-1} \int_0^1 Q(x)K\left(\frac{x-t}{h_n}\right)dx \). Using (A.1) and (A.8), we get

\[
\hat{Q}_n(t) - Q(t) = n^{-1} \sum_{i=1}^n \int_{t-h_n}^{t+h_n} [-\varphi_i(Q(x))] \frac{1}{f(Q(x))h_n} K\left(\frac{x-t}{h_n}\right)dx + o_p(n^{-1/2}),
\]

where

\[
\varphi_i(x) = (1 - F(x)) \left[ I(Z_i \leq x, \delta_i = 1) - \int_{a_W}^x I(T_i \leq u \leq Z_i) \frac{C(u)}{C^2(u)} dW_1(u) \right]. \tag{A.9}
\]

Thus an application of Liapunov’s form of the Central Limit Theorem gives

\[
n^{1/2} (\hat{Q}_n(t) - Q_n(t)) \to N(0, \sigma^2(t)) \quad \text{in distribution as} \quad n \to \infty. \tag{A.10}
\]

Note that

\[
\hat{Q}_n(t) - Q(t) \sim [\hat{Q}_n(t) - \hat{Q}_n(t)] + [\hat{Q}_n(t) - Q(t)]. \tag{A.11}
\]

It follows from a Taylor expansion that

\[
\hat{Q}_n(t) - Q(t) = \frac{h_n^r}{r!} \int_{x_0}^1 x^r K(x)Q^{(r)}(t + h_n \theta x)dx
\]
for some $0 \leq \theta \leq 1$, where $Q^{(r)}$ is the $r$th derivative of $Q$. Thus
\[ n^{1/2}|\tilde{Q}_n(t) - Q(t)| = O(n^{1/2}h_n^r) \rightarrow 0 \text{ as } n \rightarrow \infty. \] (A.12)

Theorem 1 then follows immediately from (A.10)-(A.12).

The following lemma is useful in proving Theorem 2.

**Lemma A.2.** Let $U, V$ be two random variables. Then for any $a > 0$,
\[ \sup_y |P(U + V \leq y) - \Phi(y)| \leq \sup_y |P(U \leq y) - \Phi(y)| + \frac{a}{\sqrt{2\pi}} + P(|V| > a), \] (A.13)
\[ \sup_y |P(U \leq yV) - \Phi(y)| \leq \sup_y |P(U \leq y) - \Phi(y)| + P(|V - 1| > a) + a. \] (A.14)

**Proof.** The first inequality follows from Lemma 2 of Chang and Rao (1989), while the second is a consequence of Michel and Pfanzagl (1971).

In the sequel $M = M(x)$ denotes a positive constant which will vary in different contexts.

**Proof of Theorem 2.** Using a change of variable theorem (cf. Shorack and Wellner (1986, p.25)), we have
\[ \hat{Q}_n(t) - \bar{Q}_n(t) = -\int_0^\infty \left( \int_{(F(x)-t)/h_n}^{(\hat{F}_n(x)-t)/h_n} K(u)du \right) dx = S_{n1} + S_{n2}, \] (A.15)
where
\[ S_{n1} = -\int_0^\infty K \left( \frac{F(x) - t}{h_n} \right) \left( \frac{\hat{F}_n(x) - F(x)}{h_n} \right) dx, \]
\[ S_{n2} = -\int_0^\infty \left[ \int_{(F(x)-t)/h_n}^{(\hat{F}_n(x)-t)/h_n} \left( K(u) - K \left( \frac{F(x) - t}{h_n} \right) \right) du \right] dx. \]

From (A.1), write
\[ S_{n1} = n^{-1} \sum_{i=1}^n V_{i}^{(n)} + r_n \] (A.16)
with
\[ V_{i}^{(n)} = -\int_0^\infty \varphi_i(x)h_n^{-1}K \left( \frac{F(x) - t}{h_n} \right) dx, \]
\[ r_n = -h_n^{-1} \int_0^\infty K \left( \frac{F(x) - t}{h_n} \right) R_n(x) dx, \]
where $\varphi_i$ and $R_n$ are defined by (A.9) and (A.1), respectively. Standard calculations yield $E V_{i}^{(n)} = 0$,
\[ \sigma_n^2 \equiv \text{Var}(V_{i}^{(n)}) = \frac{(1-t)^2}{f^2(Q(t))} \int_{aW}^{Q(t)} \frac{dW_1(x)}{C^2(x)} + O(h_n) = \sigma^2(t) + O(h_n) \] (A.17)
and $E[V_i^{(n)}]^3 \leq d_1(1-t)^3 \int_{a_W}^{Q(t)} \frac{dW(z)}{C(t(z))} < \infty$, where $d_1 > 0$ is a constant. Thus, by the Berry-Esseen Theorem (see Chow and Teicher (1978, p.299)), for all $n \geq 1$,

$$\sup_{y} \left| P\{n^{-\frac{1}{2}} \sum_{i=1}^{n} V_i^{(n)} < y\sigma_n\} - \Phi(y) \right| \leq Mn^{-1/2}. \quad (A.18)$$

It follows from (A.14) that

$$\sup_{y} |P\{n^{-1/2} \sum_{i=1}^{n} V_i^{(n)} < y\sigma(t)\} - \Phi(y)|$$

$$\leq \sup_{y} |P\{n^{-1/2} \sum_{i=1}^{n} V_i^{(n)} < y\sigma_n\} - \Phi(y)| + P\{|\sigma_n^{-1}\sigma(t) - 1| > a\} + a \quad (A.19)$$

for any $a > 0$. It is easy to see that for any random variable $V$ and $0 < \varepsilon < 1/2$, $P\{|V - 1| > 2\varepsilon, V \neq 0\} \leq 2P\{|V - 1| > \varepsilon\}$. Hence, from the fact that for $y, z > 0$ the relation $|y - 1| > z$ holds if $|\sqrt{y} - 1| > \sqrt{z}$, and from (A.17), we have

$$P\{|\sigma_n^{-1}\sigma(t) - 1| > a\} \leq P\{|\sigma_n^{-2}\sigma^2(t) - 1| > a^2\} \leq 2P\{|\sigma_n^2 - a^2/2\} \leq 2P\{d_2h_n > a^2/2\} \quad (A.20)$$

for some constant $d_2 > 0$. In view of (A.18) – (A.20) and putting $a = (2d_2h_n)^{1/2}$, we get

$$\sup_{y} |P\{n^{-\frac{1}{2}} \sum_{i=1}^{n} V_i^{(n)} \leq y\sigma(t)\} - \Phi(y)| \leq M(n^{-1/2} + h_n^{1/2}). \quad (A.21)$$

When $a_G < a_W$, from (1.15) of Gijbels and Wang (1993) we obtain that for each $x > 0$ and $0 \leq b < b_W$,

$$P\left\{ \sup_{0 \leq y \leq b} n|R_n(y)| > x + 4\theta^{-2} \right\} \leq M(e^{-\lambda x} + (x/50)^{-2n} + e^{-\lambda x^3}), \quad (A.22)$$

where $\theta > 0$ and $\lambda > 0$ are some constants. Note that for large $n$, there exists a constant $b$ such that $0 \leq Q(t + h_nu) \leq b < b_W$ for all $|u| \leq 1$. Hence for all $n$ sufficiently large, it follows from (A.22) that

$$P\{n^{\frac{3}{2}} |R_n| \geq n^{-1/2}h_n^{-1}\log n\}$$

$$= P\left\{ \left| \int_{-1}^{1} \frac{1}{f(Q(t + h_nu))} K(u)R_n(Q(t + h_nu))du \right| \geq (nh_n)^{-1}\log n \right\}$$

$$\leq P\left\{ \sup_{0 \leq y \leq b} n|R_n(y)| \geq d_3h_n^{-1}\log n \right\} \leq Mn^{-1}, \quad (A.23)$$
where $d_3 > 0$ is a constant. For some $\tau > 0$ with $1 + \tau \leq (1 - t)/h_n$ (this holds if $n$ is large), write

$$S_{n2} = \Delta_{n1} + \Delta_{n2} + \Delta_{n3},$$

(24.24)

where $\Delta_{n1} = -h_n \int_{1+\tau}^{1} K(x) dx$, $\Delta_{n2} = -h_n \int_{1+\tau}^{1-t}/h_n K(x) dx$ and $\Delta_{n3} = -h_n \int_{1-t}/h_n K(x) dx$ with

$$K(x) = \frac{1}{f(\xi(t+h_n x))} \int_{x}^{\hat{F}_n(\xi(t+h_n x))-t}/h_n (K(u) - K(x)) du.$$

Since $K$ has bounded support $[-1,1]$,

$$|\Delta_{n2}| \leq h_n \int_{1+\tau}^{1-t}/h_n \frac{1}{f(\xi(t+h_n x))} I\left(\frac{\hat{F}_n(\xi(t+h_n x)) - t}{h_n} < 1\right)$$

$$\times \int_{x}^{\hat{F}_n(\xi(t+h_n x))-t}/h_n \left(K(u) - K(x)\right) du \left|dx\right.$$

$$\leq h_n \int_{1+\tau}^{1-t}/h_n \frac{1}{f(\xi(t+h_n x))} I\left(\frac{\hat{F}_n(\xi(t+(1+\tau)h_n)) - (t+(1+\tau)h_n)}{h_n} < -\tau\right)$$

$$\times \int_{x}^{\hat{F}_n(\xi(t+h_n x))-t}/h_n \left(K(u) - K(x)\right) du \left|dx\right.$$

(A.25)

Using Theorem 1 of Zhu (1996), we have that for $a_G < a_W \leq b < b_W$ and $\varepsilon > 0$,

$$P\left\{\sup_{a_W \leq x \leq b} |\hat{F}_n(x) - F(x)| > \varepsilon\right\} \leq d_4 \exp(-nd_5\varepsilon^2),$$

(A.26)

where $d_4$ and $d_5$ are absolute constants. Therefore, (A.25) and (A.26) imply that for all $n$ sufficiently large,

$$P\left\{n^{1/2}|\Delta_{n2}| \geq n^{-1/2}h_n^{-1} \log n\right\} \leq P\left\{n^{1/2}|\hat{F}_n(\xi(t+(1+\tau)h_n)) - (t+(1+\tau)h_n)| \geq \tau h_n\right\}$$

$$\leq d_4 \exp(-nd_5\tau^2h_n^2) \leq Mn^{-1}.$$  

(A.27)

(Recall that $nh_n^2/\log n \to \infty$, so that $nd_5\tau^2h_n^2 > \log n$ for large $n$.) Similarly,

$$P\left\{n^{1/2}|\Delta_{n3}| \geq n^{-1/2}h_n^{-1} \log n\right\} \leq Mn^{-1}.$$  

(A.28)

Since $K$ is Lipschitz continuous of order one, for some constant $d_6 > 0$, $|\Delta_{n1}| \leq d_6h_n^{-1} \sup_{|x| \leq 1+\tau} |\hat{F}_n(\xi(t+h_n x)) - (t+h_n x)|^2$. Hence, it follows from (A.26) that

$$P\left\{n^{1/2}|\Delta_{n1}| \geq n^{-1/2}h_n^{-1} \log n\right\} \leq P\left\{\sup_{|x| \leq 1+\tau} |\hat{F}_n(\xi(t+h_n x)) - (t+h_n x)|^2 \geq d_6^{-1/2}n^{-1/2}(\log n)^{1/2}\right\} \leq Mn^{-1}.\tag{29}$$
Take \( U = n^{-1/2} \sum_{i=1}^{n} V_i^{(n)} \), \( V = n^{1/2}(r_n + S_{n2}) \), and \( a_n = n^{-1/2}h_n^{-1} \log n \) in (A.13), so that \( n^{1/2}(\bar{Q}_n(t) - \bar{Q}_n(t)) = U + V \) by (A.15)–(A.16). It follows from (A.21), (A.23), (A.24) and (A.27)–(A.29) that 
\[
\sup_y |P\{n^{1/2}(\bar{Q}_n(t) - \bar{Q}(t)) \leq y\sigma(t)\} - \Phi(y)| \leq M\left(n^{-1/2}h_n^{-1} \log n + h_n^{1/2}\right).
\]
Note that \( |\Phi(x) - \Phi(y)| \leq |x - y| \), \( -\infty < x, y < \infty \). Now, using (A.11), (A.12) and Theorem 3, we get
\[
\sup_y \left| P\{n^{1/2}(\bar{Q}_n(t) - \bar{Q}(t)) \leq y\sigma(t)\} - \Phi(y) \right|
= \sup_y \left| P\{n^{1/2}(\bar{Q}_n(t) - \bar{Q}(t)) \leq y'\sigma(t)\} - \Phi(y') + \Phi(y') - \Phi(y) \right|
\leq \sup_y \left| P\{n^{1/2}(\bar{Q}_n(t) - \bar{Q}(t)) \leq y\sigma(t)\} - \Phi(y) \right| + n^{1/2}|\bar{Q}_n(t) - \bar{Q}(t)|/\sigma(t)
\leq M\left(n^{-1/2}h_n^{-1} \log n + n^{1/2}h_n^{-1/2} + h_n^{1/2}\right),
\]
where \( y' = y - n^{1/2}(\bar{Q}_n(t) - \bar{Q}(t))/\sigma(t) \). This completes the proof of Theorem 2.

Acknowledgement

This work was partially supported by an RGC Earmarked Grant No. B-Q175 and a HKPOLYU Internal Research Grant No. G-S598. We are grateful to the Editor, Associate Editor, and an anonymous referee for their valuable comments and suggestions which led to the improvement of this paper.

References


Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

E-mail: maxzhou@polyu.edu.hk

Institute of Applied Mathematics, Academia Sinica, Beijing 100080, China.

Dept of Probability and Statistics, Peking University, Beijing 100871, China.

(Received February 1999; accepted March 2000)