

A Berry-Esseen Type Bound in Kernel Density Estimation for a Random Left-Truncation Model

P. Asghari^a, V. Fakoor^{1,a}, M. Sarmad^a

^aDepartment of Statistics, Faculty of Mathematical Sciences,
Ferdowsi University of Mashhad, Iran

Abstract

In this paper we derive a Berry-Esseen type bound for the kernel density estimator of a random left truncated model, in which each datum (Y) is randomly left truncated and is sampled if $Y \geq T$, where T is the truncation random variable with an unknown distribution. This unknown distribution is estimated with the Lynden-Bell estimator. In particular the normal approximation rate, by choice of the bandwidth, is shown to be close to $n^{-1/6}$ modulo logarithmic term. We have also investigated this normal approximation rate via a simulation study.

Keywords: Asymptotic normality, Berry-Esseen, kernel density estimation, rate of convergence, left-truncation.

1. Introduction

Left truncated data are widely occurring in medical studies mostly in lifetime analysis of patients that have a particular disease. In epidemiology they are usually referred to as prevalent cases. In other research that deals with survival analysis, the data cannot be seen completely and we do not have sufficient information about the individuals before the time of recruiting the data. This is the reason that the left truncation model and the left truncation random variable appear. Truncation also occurs in industrial and insurance studies. So many works have been done to non-parametrically estimate the density function of left truncated data. A kernel estimation procedure is used in most of these works. It is of interest to obtain a rate of convergence of these non-parametric density estimators to the normal distribution known as Berry-Esseen bound.

Many works have been done in different data sampling models due to the importance and applicability of density estimation and Berry-Esseen bounds for the density estimator. Prakasa Rao (1975) considered a density estimator and presented a Berry-Esseen type bound for the estimator when the observations were from a stationary Markov process. Birkel (1988) gave uniform rates of convergence in the central limit theorem for a sequence of associated random variables. Chen (1997) gained the rate of convergence in the central limit theorem for the conditional empirical function and also for conditional sample quantiles based on kernel estimation. Isogai (1994) investigated a Berry-Esseen type bound for the kernel estimators of p^{th} order derivative of the density for independent and identically distributed random variables. Cheng (1998) established a Berry-Esseen type bound for the distribution of smooth quantile density estimators. Some works have been done in a censorship model in order to obtain a Berry-Esseen bound for a kernel density estimator by Sun and Zhu (1999); subsequently, a uniform and a non-uniform Berry-Esseen bound for a sequence of stationary random variables have

¹ Corresponding author: Assistant Professor, Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Iran. E-mail: fakoor@math.um.ac.ir

been driven by Dewan and Prakasa Rao (2002). They used this bound to obtain another bound for the expected value of the integrated mean of deviation of the kernel density estimator. About a random left truncated model (RLTM), several works have been done to estimate distributional parameters. Recently Zhou *et al.* (2006) have proposed a kernel-type estimator of the quantile function in an RLTM scheme. Ould-Saïd and Tatachak (2009) have non-parametrically estimated the mode of the distribution in RLTM using a kernel estimator. They have also investigated the strong consistency and the normality of their estimator. Liang and Baek (2008) presented uniform Berry-Esseen type bounds for a sequence of negatively associated random variables. They also used a kernel density estimator for the true density function. Liang and Uña-Álvarez (2009) have derived a Berry-Esseen bound for strong mixing censored data using a kernel estimator. For stationary processes, a Berry-Esseen bound has been driven for kernel estimators by Huang *et al.* (2011). They used a martingale approximation in their procedure. Yang *et al.* (2012) investigated a Berry-Esseen type bound for the sample quantiles of a strongly mixing sequence.

There is limited work on Berry-Esseen bounds for an RLTM and this paper intends to add to the existing literature. Here we present a Berry-Esseen type bound for the kernel density estimator in an RLTM.

Let \mathcal{P} be a finite population with large and deterministic size N . Each element of \mathcal{P} contains two independent and positive random variables that are denoted by Y and T with continuous distribution functions, respectively shown by F and G . So until now we have N independent and identically distributed random variables $(Y_i, T_i)_{1 \leq i \leq N}$. In order to correspond this population with the RLTM, suppose that T is the truncation random variable and (Y, T) is observed if $Y \geq T$, otherwise we have no information about them. So among the N random variables we only observe a limited number of them, shown by n , $n \leq N$. As a direct result n is a binomial random variable, with sample size N and success probability $\alpha = P(Y \geq T)$. In application, α is unknown. Among several estimators for α , in this paper we use the more familiar one that is proposed by He and Yang (1998).

In what follows, we focus on estimating the probability density function of Y that is denoted by $f(\cdot)$, using a kernel density estimator. In this sequel, we would also have to estimate G . Here we use its non-parametric maximum likelihood estimator (NPMLE). As for our main result, we obtain a Berry-Esseen type bound for the estimator of f .

The layout of this paper is as follows. In Section 2, we introduce our notations and we also present some preliminaries. In Section 3, we derive our main theorems and results. In Section 4 we perform a simulation study. The proofs of the theorems and corollaries of Section 2 are deferred to Section 5.

2. Preliminaries and Notation

As it is mentioned in the Introduction, in an RLTM, we observe n pairs ($n \leq N$) of $(Y_i, T_i)_{(1 \leq i \leq N)}$ if $Y_i \geq T_i$. Let (Y, T) be the generic random variables distributed identical to $(Y_i, T_i)_{(1 \leq i \leq N)}$. Y is the variable of interest and T is the left truncation random variable. Now let the conditional distribution of (Y, T) be denoted by H^* . So we have

$$\begin{aligned} H^*(y, t) &= P(Y \leq y, T \leq t | Y \geq T) \\ &= \alpha^{-1} \int_{-\infty}^y G(t \wedge u) dF(u), \end{aligned} \quad (2.1)$$

from the Introduction we know that $\alpha = P(Y \geq T)$.

So the marginal distribution function of Y is

$$F^*(y) = H^*(y, \infty) = \alpha^{-1} \int_{-\infty}^y G(u) dF(u),$$

and the marginal probability density function of Y is

$$f^*(y) = \alpha^{-1} G(y) f(y). \quad (2.2)$$

As it is discussed before, we are interested in estimating $f(\cdot)$, so from (2.2) we have

$$f(y) = \frac{\alpha}{G(y)} f^*(y). \quad (2.3)$$

Here we use the following kernel density estimator for $f(y)$

$$f_n(y) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{Y_i - y}{h_n}\right) \frac{\alpha}{G(Y_i)}, \quad (2.4)$$

in which $\{h_n\}_{n \geq 1}$ is a bandwidth sequence, such that $h_n \rightarrow 0$ as $n \rightarrow \infty$, and $K(\cdot)$ is a kernel function.

In what follows, for any distribution function L on $[0, \infty]$, let $a_L := \inf\{x > 0 : L(x) > 0\}$ and $b_L := \sup\{x > 0 : L(x) < 1\}$.

$f_n(y)$ is applicable when G is known, but in most applications, it is unknown. So we need to use an estimator. Here we use the non-parametric maximum likelihood estimator of G , modified by Lynden-Bell (see Lynden-Bell, 1971), which we refer to it by $G_n(\cdot)$. If there are no ties in the data, it is given by

$$F_n(y) = 1 - \prod_{i: Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)} \right], \quad G_n(t) = 1 - \prod_{i: T_i > t} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)} \right], \quad (2.5)$$

in which

$$C_n(y) = G_n^*(y) - F_n^*(y^-) = \frac{1}{n} \sum_{i=1}^n I_{\{T_i \leq y \leq Y_i\}}, \quad y \in [a_F, +\infty),$$

it is the empirical estimator of

$$C(y) := G^*(y) - F^*(y) = \alpha^{-1} G(y) (1 - F(y)), \quad y \in [a_F, +\infty). \quad (2.6)$$

(2.6) suggests using the following estimator for α ,

$$\alpha_n = \frac{G_n(y) [1 - F_n(y^-)]}{C_n(y)}. \quad (2.7)$$

This estimator is defined for all y such that $C_n(y) \neq 0$ and it is shown that it is not a function of y . In addition, He and Yang (1998) have shown that α_n is strongly consistent for α . According to (2.5) and (2.7), we are now in a place to present the plug-in version of $f_n(\cdot)$ which we refer to it by $\hat{f}_n(\cdot)$

$$\hat{f}_n(y) = \frac{\alpha_n}{nh_n} \sum_{i=1}^n K\left(\frac{Y_i - y}{h_n}\right) \frac{1}{G_n(Y_i)}. \quad (2.8)$$

Note that in (2.8) the sum is taken over i 's for which $G_n(Y_i) \neq 0$. Let $\sigma_n^2(y) := nh_n \text{Var}(f_n(y))$ and $\sigma^2(y) := \alpha f(y) / G(y) \int_{-1}^1 K^2(t) dt$.

3. Main Results

Before presenting the main results, we should mention some basic assumptions that are used in the following theorems and lemmas. In the sequel C is a compact set such that $C \subset \{y : y \geq a_F\}$. Here is the required definition:

Definition 1. The kernel function $K(\cdot)$ is said to be a second order kernel function if $\int K(t) dt = 1$, $\int tK(t) dt = 0$ and $\int t^2 K(t) dt > 0$.

A1 $K(\cdot)$ is a positive kernel function such that $\int_{-1}^1 K(t) dt = 1$, $K(t) = 0$ for $|t| > 1$ and it is bounded for any t within $[-1, 1]$.

A2 For the distribution functions F and G , $a_G < a_F$, $b_G \leq b_F$.

A3 (i) $nh_n \rightarrow \infty$ as $n \rightarrow \infty$.
(ii) $h_n \log \log n \rightarrow 0$ as $n \rightarrow \infty$.
(iii) $nh_n^5 \rightarrow 0$ as $n \rightarrow \infty$.

A4 $K(\cdot)$ is a second-order kernel function.

A5 $f(\cdot)$ is two times continuously differentiable for $y \in C$.

Discussion on the Assumptions. Assumption A1 is commonly used in non-parametric estimation. From Woodroffe (1985) Assumption A2 is needed to estimate F and G . It is mentioned that if $a_G \leq a_F$ then F is estimable. The uniform convergence rate of G_n to G is true for y 's that do not include a_G . This means that we should have $a_G < a_F$.

Assumptions A3(i) and A3(iii) are used for consistency of the kernel density estimator in the literature. Assumptions A3(i)–A3(iii) are used in Theorems 2 and 3 and they are needed for normal approximation in this paper. Assumptions A4 and A5 are needed to use Theorem 1 of Ould-Saïd and Tatachak (2009) in Corollary 2.

Theorem 1. Suppose that A1–A3(i) are satisfied and also f and G have bounded first derivatives in a neighborhood of y for $y \geq a_F$, then for such y 's we have

$$\sup_{x \in \mathbb{R}} \left| P \left(\sqrt{nh_n} [f_n(y) - Ef_n(y)] \leq x\sigma_n(y) \right) - \Phi(x) \right| = O \left((nh_n)^{-\frac{1}{2}} \right).$$

Theorem 2. Under the assumptions of Theorem 1 and A3(ii), for $y \geq a_F$ we have

$$\sup_{x \in \mathbb{R}} \left| P \left(\sqrt{nh_n} [\hat{f}_n(y) - Ef_n(y)] \leq x\sigma_n(y) \right) - \Phi(x) \right| = O \left((nh_n)^{-\frac{1}{2}} + (h_n \log \log n)^{\frac{1}{4}} \right).$$

Theorem 3. Suppose that A1–A5 are satisfied and let f and G have bounded first derivatives in a neighborhood of $y \geq a_F$, then for such y 's we have

$$\sup_{x \in \mathbb{R}} \left| P \left(\sqrt{nh_n} [\hat{f}_n(y) - f(y)] \leq x\sigma(y) \right) - \Phi(x) \right| = O(a_n),$$

where $a_n = (nh_n)^{-1/2} + (h_n \log \log n)^{1/4} + h_n + n^{1/2}h_n^{5/2}$.

Remark 1. When f and G are unknown, Theorem 3 cannot be used in statistical applications such as finding a confidence interval for $f(y)$ or hypothesis testing. So we have to estimate $\sigma^2(y)$. Here we propose the following estimator for $\sigma^2(y)$

$$\hat{\sigma}_n^2(y) = \frac{\alpha_n \hat{f}_n(y)}{G_n(y)} \int_{-1}^1 K^2(t) dt.$$

Use of this estimator forces a change in the rate achieved in Theorem 3. This change is discussed in the following corollaries.

In the following corollaries, we substitute A3(i) with $nh_n/\log n \rightarrow \infty$ as $n \rightarrow \infty$.

Corollary 1. *Let the assumptions of Theorem 3 be satisfied. We have*

$$\sup_{y \in \mathcal{C}} |\hat{\sigma}_n^2(y) - \sigma^2(y)| = O \left[\max \left(\sqrt{\frac{\log n}{nh_n}}, h_n^2 \right) + \sqrt{\frac{\log \log n}{n}} \right] \text{ a.s.}$$

Corollary 2. *If the assumptions of Corollary 3 are satisfied, then for $y \in \mathcal{C}$ we have*

$$\sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\hat{\sigma}_n(y)} \leq x \right) - \Phi(x) \right| = O(a_n + b_n),$$

in which $b_n = \max(\sqrt{\log n/nh_n}, h_n^2) + \sqrt{\log \log n/n}$.

Remark 2. If we choose $h_n = O(n^{-\alpha})$ for $1/3 < \alpha < 2/3$, the Berry-Esseen bound in Theorem 3 and Corollary 2 reduces to $O(n^{-\alpha/4})$ modulo logarithmic term, therefore if we choose α near $2/3$, the convergence rate approximately equals to $O(n^{-1/6})$ modulo logarithmic term as $n \rightarrow \infty$. This means that estimating $\sigma^2(y)$ does not affect the rate of convergence.

4. Simulation Study

We perform a simulation study in this section to investigate the normality of \hat{f}_n . As it is mentioned after Theorem 3, Corollary 2 is the applicable form of Theorem 3 which is the main theorem. So it is of interest to perform the simulation for this applicable form.

In order to generate a random left truncated sequence of (Y_i, T_i) we assume that Y has exponential distribution truncated by a_F with mean $1/3 + a_F$ for $a_F = 0.5, 1, 2$. T is also exponentially distributed with mean $1/\lambda$ for $\lambda > 0$. This parameter is used in order to obtain different truncation proportions that are $\alpha = 0.3, 0.7, 0.9$. We employ the kernel $K(u) = 3/4(1 - u^2)\mathbb{I}(|u| \leq 1)$.

$N = 1000$ pairs of (Y_i, T_i) are simulated and those in which $Y_i \leq T_i$ are put aside and a sample of size n is achieved. Then we evaluate

$$P \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\hat{\sigma}_n(y)} \leq x \right),$$

using Monte Carlo based on $M = 1000$ replications for $x = 1$ and $y = 3$. Let this estimate be denoted by $\hat{P}(\sqrt{nh_n}(\hat{f}_n(y) - f(y))/\hat{\sigma}_n(y) \leq x)$. Then we calculate

$$A_n(x, y, a_F) := \left| \hat{P} \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\hat{\sigma}_n(y)} \leq x \right) - \Phi(x) \right|. \quad (4.1)$$

Table 1: Simulation results

α	$A_n(1, 3, 0.5)$	$A_n(1, 3, 1)$	$A_n(1, 3, 2)$
0.3	0.1338	0.1446	0.1586
0.7	0.1156	0.0776	0.06026
0.9	0.1206	0.0766	0.0034

The bandwidth that is used to perform the simulation is what that minimizes the observed integrated square error (ISE) between the estimate \hat{f}_n and the actual density f with respect to the restriction that is assumed in Remark 2 which says that h_n should be equal to n^{-a} for $1/3 < a < 2/3$. The ISE is given by

$$\text{ISE}(\hat{f}_n, h_n; f) = \int [\hat{f}_n(y) - f(y)]^2 dy, \quad (4.2)$$

and as it is mentioned, the optimal bandwidth is

$$h_{ISE} = \underset{n^{-\frac{2}{3}} < h_n < n^{-\frac{1}{3}}}{\text{argmin}} \text{ISE}(\hat{f}_n, h_n; f). \quad (4.3)$$

The procedure of achieving h_{ISE} is as follows. \hat{f}_n would be calculated for a simulated sample and as it is mentioned, f is known. So the ISE would be a function of h_n . For each of 1000 replication of simulations, the minimizing point of this function would be the h_{ISE} . Table 1 summarizes the results as well as shows that the Berry-Esseen Theorem works better for high proportions of truncation like 90%, rather than low proportions such as 30%. It can be seen that in high proportions of truncation, when the difference between a_F and a_G increases, our normal convergence rate improves.

5. Proofs

Below we mention some lemmas that are used in the proofs of the main theorems. Note that from now on, C is a purely numerical positive constant that may change from one line to other line in the procedure of proof of a theorem.

Lemma 1. *Suppose that A1 and A2 are satisfied and let f and G have bounded first derivatives in a neighborhood of y . For $y \geq a_F$*

$$|\sigma_n^2(y) - \sigma^2(y)| = O(h_n).$$

Proof: From the definition of $\sigma_n^2(y)$ and $\sigma^2(y)$ we have

$$\begin{aligned} |\sigma_n^2(y) - \sigma^2(y)| &= \left| \frac{1}{h_n} \int K^2\left(\frac{u-y}{h_n}\right) \frac{\alpha^2 f^*(u)}{G^2(u)} du - \frac{1}{h_n} \left[\int K\left(\frac{u-y}{h_n}\right) \frac{\alpha f^*(u)}{G(u)} du \right]^2 - \frac{\alpha f(y)}{G(y)} \int_{-1}^1 K^2(t) dt \right| \\ &= \left| \int_{-1}^1 K^2(t) \frac{\alpha f(y+h_nt)}{G(y+h_nt)} dt - h_n \left[\int_{-1}^1 K(t) f(y+h_nt) dt \right]^2 - \frac{f(y)}{G(y)} \int_{-1}^1 K^2(t) dt \right| \\ &= \left| \int_{-1}^1 \alpha K^2(t) \frac{G(y) f(y+h_nt) - G(y+h_nt) f(y)}{G(y+h_nt) G(y)} dt - h_n \left[\int_{-1}^1 K(t) f(y+h_nt) dt \right]^2 \right| \\ &= O(h_n). \end{aligned} \quad (5.1)$$

Note that in the last equation we used A1 and the fact that f and G have bounded first derivatives in a neighborhood of y . \square

Proof of Theorem 1: Set

$$\begin{aligned} f_n(y) - E[f_n(y)] &= \frac{1}{nh_n} \sum_{i=1}^n \left\{ K\left(\frac{Y_i - y}{h_n}\right) \frac{\alpha}{G(Y_i)} - E\left(K\left(\frac{Y_i - y}{h_n}\right) \frac{\alpha}{G(Y_i)}\right) \right\} \\ &=: \sum_{i=1}^n Z_{ni}. \end{aligned} \quad (5.2)$$

From the Berry-Esseen Theorem (See Petrov, 1995), it can be written that

$$\sup_{x \in \mathbb{R}} \left| \Pr\left(\sqrt{nh_n} [f_n(y) - E f_n(y)] \leq x \sigma_n(y)\right) - \Phi(x) \right| \leq C \frac{n^{\frac{5}{2}} h_n^{\frac{3}{2}} E|Z_{n1}|^3}{\sigma_n^3(y)}, \quad (5.3)$$

and

$$\begin{aligned} E|Z_{n1}|^3 &\leq \frac{1}{(nh_n)^3} \left\{ E \left| K\left(\frac{Y_1 - y}{h_n}\right) \frac{\alpha}{G(Y_1)} \right|^3 + \left| E \left(K\left(\frac{Y_1 - y}{h_n}\right) \frac{\alpha}{G(Y_1)} \right) \right|^3 \right\} \\ &\leq \frac{2}{(nh_n)^3} E \left| K\left(\frac{Y_1 - y}{h_n}\right) \frac{\alpha}{G(Y_1)} \right|^3, \end{aligned} \quad (5.4)$$

and

$$\begin{aligned} E \left| K\left(\frac{Y_1 - y}{h_n}\right) \frac{\alpha}{G(Y_1)} \right|^3 &= \alpha^3 \int K^3\left(\frac{u - y}{h_n}\right) \frac{f^*(u)}{G^3(u)} du \\ &= \alpha^2 h_n \int_{-1}^1 K^3(t) \frac{f(y + h_n t)}{G^2(y + h_n t)} dt. \end{aligned} \quad (5.5)$$

Since f and G are bounded in a neighborhood of y , A1 and a little calculation yields

$$\sup_{x \in \mathbb{R}} \left| \Pr\left(\sqrt{nh_n} [f_n(y) - E f_n(y)] \leq x \sigma_n(y)\right) - \Phi(x) \right| = O\left((nh_n)^{-\frac{1}{2}}\right), \quad (5.6)$$

and Theorem 1 is proved. \square

In order to start the proof of Theorem 2, we shall state another lemma that is subsequently used.

Lemma 2. *Let X and Y be random variables. For any $a > 0$ we have*

$$\sup_{u \in \mathbb{R}} |P(X + Y \leq u) - \Phi(u)| \leq \sup_{u \in \mathbb{R}} |P(X \leq u) - \Phi(u)| + \frac{a}{\sqrt{2\pi}} + P(|Y| > a).$$

The proof of Lemma 2 can be found in Chang and Rao (1989), page 4655.

Proof of Theorem 2: Set $K_{ni} := K((Y_i - y)/h_n)$. By using Lemma 2 for an arbitrary $a > 0$, we have the following decomposition

$$\begin{aligned} &\sup_{x \in \mathbb{R}} \left| P\left(\sqrt{nh_n} [\hat{f}_n(y) - E f_n(y)] \leq x \sigma_n(y)\right) - \Phi(x) \right| \\ &\leq \sup_{x \in \mathbb{R}} \left| P\left(\sqrt{nh_n} [f_n(y) - E f_n(y)] \leq x \sigma_n(y)\right) - \Phi(x) \right| + \frac{a}{\sqrt{2\pi}} + P\left(\left| \frac{\sqrt{nh_n} [\hat{f}_n(y) - f_n(y)]}{\sigma_n(y)} \right| > a\right). \end{aligned} \quad (5.7)$$

Alternatively

$$\begin{aligned}
E \left| \frac{\sqrt{nh_n} [\hat{f}_n(y) - f_n(y)]}{\sigma_n(y)} \right| &\leq \frac{1}{\sqrt{nh_n} \sigma_n(y)} \sum_{i=1}^n E \left\{ |K_{ni}| \left| \frac{\alpha_n}{G_n(Y_i)} - \frac{\alpha}{G(Y_i)} \right| \right\} \\
&\leq \frac{n}{\sqrt{nh_n} \sigma_n(y)} E \left\{ |K_{n1}| \frac{|\alpha_n G(Y_1) - \alpha G_n(Y_1)|}{G_n(Y_1) G(Y_1)} \right\} \\
&\leq \frac{n}{\sqrt{nh_n} \sigma_n(y)} E \left\{ |K_{n1}| \frac{G(Y_1) |\alpha_n - \alpha| + \alpha |G_n(Y_1) - G(Y_1)|}{G_n(Y_1) G(Y_1)} \right\}, \quad (5.8)
\end{aligned}$$

using Woodroffe (1985) we have

$$\sup_{y \geq a_F} |G_n(y) - G(y)| = O \left(\sqrt{\frac{\log \log n}{n}} \right), \quad (5.9)$$

and also using He and Yang (1998) we have

$$|\alpha_n - \alpha| = O \left(\sqrt{\frac{\log \log n}{n}} \right), \quad (5.10)$$

now according to (5.9), (5.10) and Lemma 1, using A1 and because f is bounded in a neighborhood of y , it can be written that

$$\begin{aligned}
(5.8) &\leq C \sqrt{nh_n} \left(|\alpha_n - \alpha| + \sup_{y \geq a_F} |G_n(y) - G(y)| \right) \int_{-1}^1 |K(t)| f(y + h_n t) dt \\
&\leq C \sqrt{nh_n} O \left(\sqrt{\frac{\log \log n}{n}} \right) \\
&= O \left(\sqrt{h_n \log \log n} \right). \quad (5.11)
\end{aligned}$$

So by letting $a = (h_n \log \log n)^{1/4}$ and using (5.7), (5.11), Markov's inequality and Theorem 1 we get the result. \square

Proof of Theorem 3: Using Lemma 2, we have

$$\begin{aligned}
&\sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\sigma(y)} \leq x \right) - \Phi(x) \right| \\
&\leq \sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - E f_n(y))}{\sigma_n(y)} \leq \frac{\sigma(y)}{\sigma_n(y)} x \right) - \Phi \left(\frac{\sigma(y)}{\sigma_n(y)} x \right) \right| \\
&\quad + \sup_{x \in \mathbb{R}} \left| \Phi \left(\frac{\sigma(y)}{\sigma_n(y)} x \right) - \Phi(x) \right| + \frac{\sqrt{nh_n} |E f_n(y) - f(y)|}{\sigma(y)}. \quad (5.12)
\end{aligned}$$

Alternatively, a little calculation yields

$$\sup_{x \in \mathbb{R}} \left| \Phi \left(\frac{\sigma(y)}{\sigma_n(y)} x \right) - \Phi(x) \right| = O \left(|\sigma_n^2(y) - \sigma^2(y)| \right). \quad (5.13)$$

Under A4 and A5, a Taylor expansion gives

$$\sup_{y \in C} |f(y) - Ef_n(y)| = O(h_n^2). \quad (5.14)$$

Now by using (5.12), (5.13), (5.14) and Theorem 2 we get the result. \square

Proof of Corollary 1: By using the definition of $\hat{\sigma}_n^2(y)$ and $\sigma^2(y)$ we can write

$$\begin{aligned} & \sup_{y \in C} |\hat{\sigma}_n^2(y) - \sigma^2(y)| \\ &= \sup_{y \in C} \left| \frac{\alpha_n \hat{f}_n(y)}{\hat{G}_n(y)} - \frac{\alpha f(y)}{G(y)} \right| \int_{-1}^1 K^2(t) dt \\ &\leq \sup_{y \in C} \left| \frac{\alpha_n \hat{f}_n(y) G(y) - \alpha f(y) G_n(y)}{G_n(y) G(y)} \right| \int_{-1}^1 K^2(t) dt \\ &\leq C \left(\sup_{y \in C} G(y) |\alpha_n \hat{f}_n(y) - \alpha f(y)| + \sup_{y \in C} f(y) |G_n(y) - G(y)| \right) \int_{-1}^1 K^2(t) dt \\ &\leq C \left(\sup_{y \in C} \alpha_n |\hat{f}_n(y) - f(y)| + \sup_{y \in C} f(y) |\alpha_n - \alpha| + \sup_{y \in C} f(y) |G_n(y) - G(y)| \right) \int_{-1}^1 K^2(t) dt. \end{aligned} \quad (5.15)$$

Under A3(iii), A4 and A5, Theorem 1 of Ould-Said and Tatachak (2009) gives

$$(5.15) = O \left[\max \left(\sqrt{\frac{\log n}{nh_n}}, h_n^2 \right) + \sqrt{\frac{\log \log n}{n}} \right]. \quad (5.16)$$

\square

Proof of Corollary 2:

$$\begin{aligned} & \sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\hat{\sigma}_n(y)} \leq x \right) - \Phi(x) \right| \\ &\leq \sup_{x \in \mathbb{R}} \left| \Pr \left(\frac{\sqrt{nh_n} (\hat{f}_n(y) - f(y))}{\sigma(y)} \leq \frac{\hat{\sigma}_n(y)}{\sigma(y)} x \right) - \Phi \left(\frac{\hat{\sigma}_n(y)}{\sigma(y)} x \right) \right| + \sup_{x \in \mathbb{R}} \left| \Phi \left(\frac{\hat{\sigma}_n(y)}{\sigma(y)} x \right) - \Phi(x) \right|. \end{aligned} \quad (5.17)$$

Now by a little calculation, from Corollary 1 we have

$$\begin{aligned} \sup_{x \in \mathbb{R}} \left| \Phi \left(\frac{\hat{\sigma}_n(y)}{\sigma(y)} x \right) - \Phi(x) \right| &= O(|\hat{\sigma}_n^2(y) - \sigma^2(y)|) \\ &= O \left[\max \left(\sqrt{\frac{\log n}{nh_n}}, h_n^2 \right) + \sqrt{\frac{\log \log n}{n}} \right]. \end{aligned} \quad (5.18)$$

From (5.17), (5.18) and with the aid of Theorem 3 we get the result and Corollary 1 is proved. \square

References

- Birkel, T. (1988). On the convergence rate in the central limit theorem for associated processes, *The Annals of Probability*, **16**, 1685–1698.
- Chang, M. N. and Rao, P. V. (1989). Berry-Esseen bound for the Kaplan-Meier estimator, *Communications in Statistics - Theory Methods*, **18**, 4647–4664.
- Chen, G. (1997). Berry-Esseen-type bounds for the kernel estimator of conditional distribution and conditional quantiles, *Journal of Statistical Planning and Inference*, **60**, 311–330.
- Cheng, C. (1998). A Berry-Esseen-type theorem of quantile density estimators, *Statistics and Probability Letters*, **39**, 255–262.
- Dewan, I. and Prakasa Rao, B. L. S. (2002). Non-uniform and uniform Berry-Esseen type bounds for stationary associated sequences, *Indian Statistical Institute*.
- He, S. and Yang, G. L. (1998). Estimation of the truncation probability in the random truncation model, *The Annals of Statistics*, **26**, 1011–1028.
- Huang, C., Wang, H. and Zhang, L. (2011). Berry-Esseen bounds for kernel estimates of stationary processes, *Journal of Statistical Planning and Inference*, **141**, 1290–1296.
- Isogai, E. (1994). A Berry-Esseen-type bound for recursive estimators of a density and its derivatives, *Journal of Statistical Planning and Inference*, **40**, 1–14.
- Liang, H. and Baek, J. (2008). Berry-Esseen bounds for density estimates under NA assumption, *Metrika*, **68**, 305–322.
- Liang, H. Y. and Uña-Álvarez, J. (2009). A Berry-Esseen type bound in kernel density estimation for strong mixing censored samples, *Journal of Multivariate Analysis*, **100**, 1219–1231.
- Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars, *Monthly Notices of the Royal Astronomical Society*, **155**, 95–118.
- Ould-Saïd, E. and Tatachak, A. (2009). On the non-parametric estimation of the simple mode under random left-truncation model, *Revue Roumaine de Mathématiques Pures et Appliquées*, **54**, 243–266.
- Petrov, V. V. (1995). *Limit Theorems of Probability Theory*, Oxford University Press Inc., New York.
- Prakasa Rao, B. L. S. (1975). Berry-Esseen type bound for density estimation of stationary Markov processes, *Indian Statistical Institute*, 15–21.
- Sun, L. and Zhu, L. (1999). A Berry-Esseen type bound for Kernel Density Estimators under Random Censorship, *Acta Mathematica Sinica*, **42**, 627–636.
- Woodroffe, M. (1985). Estimating a distribution function with truncated data, *The Annals of Statistics*, **13**, 163–177.
- Yang, W., Hu, S., Wang, X. and Ling, N. (2012). The Berry-Esseen type bound of sample quantiles for strong mixing sequence, *Journal of Statistical Planning and Inference*, **142**, 660–672.
- Zhou, Y., Wu, G. and Li, D. (2006). A kernel-type estimator of a quantile function under randomly truncated data, *Acta Mathematica Scientia*, **26B**, 585–594.