Journal of Sciences, Islamic Republic of Iran 26(3): 265 - 272 (2015) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

A Berry-Esseen Type Bound for the Kernel Density Estimator of Length-Biased Data

P. Asghari, V. Fakoor*, and M. Sarmad

Department of Statistics, Faculty of Mathematical Sciences,
Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran.

Received: 12 July 2014 / Revised: 6 June 2015 / Accepted: 10 August 2015

Abstract

Length-biased data are widely seen in applications. They are mostly applicable in epidemiological studies or survival analysis in medical researches. Here we aim to propose a Berry-Esseen type bound for the kernel density estimator of this kind of data.The rate of normal convergence in the proposed Berry-Esseen type theorem is shown to be O(n^(-1/6)) modulo logarithmic term as n tends to infinity by a proper choice of the bandwidth.The results of a simulation study is also presented in this paper inorder to examine the performance of the result.

Keywords: Asymptotic normality; Berry-Esseen theorem; Kernel estimator; Rate of convergence; Length-biased.

Introduction

Length-biased data happen in different applications. If larger elements of the population (in size, length or volume) are more probable to be sampled, then this sample is called length-biased. One of the most important applications of analysing this kind of data is in survival analysis. When survival data are collected from patients with diseases such as AIDS, cancer or dementia, in most cases the initiation time of the diseases is not definite. In these cases it is obvious that any individual who lives longer, is more probable to be sampled. Thus, this sample of survival data is length-biased.

The phenomenon of length-bias is noticed by Wicksell [18] for the first time. In his research he noticed that only the cells that were larger than a particular amount, were visible in the microscope. This caused a length-biased sample of cells to be studied. Length-bias was later statistically studied by Mcfadden [12], Blumenthal [3] and Cox [5]. There are many other examples of application of length-biased data that give a good reason to work on various aspects of these kinds of data. Estimating the density function of the population of such data is of high interest. Here we use two kernel density estimators that are proposed by Jones [8] and Bhattacharyya et al. [2]. Various properties of these estimators are studied, but the rate of normal convergence is not investigated. The rate of normal convergence isusually achieved in the form of a theorem that is referred to as the Berry-Esseen type theorem in the similar works. Parzen [14] achieved a Berry-Esseen type bound for the kernel density estimator when the sample is independent and identically distributed (i.i.d.), Rosenblatt [17] investigated the normality of the kernel density estimator of a sample from a stationary Markov process and Parakasa Rao [15] proposed a Berry-Esseen type theorem for this estimator. Isogai [9] investigated a Berry-Esseen type bound for the estimator of the

* Corresponding author: Tel: +985138828605; Fax: +985138828605; Email: fakoor@math.um.ac.ir

derivatives of the density function of a random sample.

Several works have been done for a random right censorship model. Liuguan and Lixing [11] investigated the Berry-Esseen type bound for the kernel density estimator of the population density function of such data. Liang and Uña-Álvarez [10] have proposed a Berry-Esseen type theorem for the kernel density estimator when the data are strongly mixing and are also opposed to random censorship. Asghari et al. [1] have started e Berry-Esseen type theorem for the kernel density estimator of left truncated data.Investigating a Berry-Esseen type bound for the kernel density estimator of length-biased data is the main purpose of this paper.

The layout of this paper is organized as follows. In preliminary section the needed notations and some preliminaries are given. The main results of the paper are proposed in th main results section. A simulation study is performed in the simulation section. A conclusion from the main results of the paper is briefed in the conclusion section. The proofs of the main results are stated in the the proof section.

Prelimaniers

Let $\mathcal{P}$ be a population and suppose that Y is the random variable of interest in this population with density function f. It is well-known that in a length-biased model, the data should be positive. n i.i.d. random variables $Y_1,\dots,Y_n$ are sampled from this population with the condition that any individual that is larger in size, length, volume or lives longer than other individuals, has a higher probability to be sampled than other individuals. This sample is length-biased with the length-biased density function that is shown by g

$$g(y)=\frac{yf(y)}{\mu}y>0,$$

in which $\mu=\int_0^\infty uf(u)du<\infty$ is the mean of the population. According to the definition of g it can be obtained that

$$\mu^{-1}=\int_0^\infty u^{-1}g(u)du.$$

It gives the idea of using the following estimator of μ that is denoted by $\hat{\mu}$

$$\hat{\mu}=\left(\frac{1}{n}\sum_{i=1}^n\frac{1}{Y_i}\right)^{-1} \quad (1)$$

Here a Berry-Esseen type bound is obtained for two kernel density estimators that are shown by $\hat{f}_{n1}$ and $\hat{f}_{n2}$ and are defined as follows

$$\hat{f}_{n1}(y)=\frac{\hat{\mu}}{nh_n}\sum_{i=1}^n\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right),$$

$$\hat{f}_{n2}(y)=\frac{\hat{\mu}}{nh_ny}\sum_{i=1}^n K\left(\frac{Y_i-y}{h_n}\right),$$

in which $K(\cdot)$ is a kernel function. $\hat{f}_{n1}$ is defined for $y>0$ and $\hat{f}_{n2}$ is defined for $y>\varepsilon$ for an arbitrary constant $\varepsilon>0$. $\hat{f}_{n1}$ and $\hat{f}_{n2}$ are originally proposed and investigated respectively by Jones [8] and Bhattacharyya et al. [2]. In order to achieve the desired result, we need to present another versions of Jones and Bhattacharyya estimators of f, which are denoted by f_{n1} and f_{n2}. These estimators are defined as bellow

$$f_{n1}(y)=\frac{\mu}{nh_n}\sum_{i=1}^n\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right),\ y>0.$$

$$f_{n2}(y)=\frac{\mu}{nh_ny}\sum_{i=1}^n K\left(\frac{Y_i-y}{h_n}\right),\ y>\varepsilon.$$

$f_{n1}(y)$ and $f_{n2}(y)$ are usefull when μ is known. In this paper we present the Berry-Esseen type theorem for $\hat{f}_{n1}$ and $\hat{f}_{n2}$. We also compare their achieved normality rate. First we give an abstract modification for the variance of f_{n1} and f_{n2} because of their usage in the main theorems.

$$Var\big(f_{n1}(y)\big)=\frac{\mu^2}{nh_n^2}Var\left(\frac{1}{Y_1}K\left(\frac{Y_1-y}{h_n}\right)\right)$$
$$=\frac{\mu^2}{nh_n^2}\left(\int K^2\left(\frac{u-y}{h_n}\right)\frac{g(u)}{u^2}du-\left(\int K\left(\frac{u-y}{h_n}\right)\frac{g(u)}{u}du\right)^2\right)$$
$$=\frac{\mu}{nh_n}\int_{-1}^1 K^2(t)\frac{f(y+th_n)}{y+th_n}dt-\frac{1}{n}\left(\int_{-1}^1 K(t)f(y+th_n)dt\right)^2),$$

$$Var\big(f_{n2}(y)\big)=\frac{\mu^2}{nh_n^2y^2}Var\left(K\left(\frac{Y_1-y}{h_n}\right)\right)$$
$$=\frac{\mu^2}{nh_n^2y^2}\left(\int K^2\left(\frac{u-y}{h_n}\right)g(u)du-\left(\int K\left(\frac{u-y}{h_n}\right)g(u)du\right)^2\right)$$
$$=\frac{\mu}{nh_ny^2}\left(\int_{-1}^1 K^2(t)(y+th_n)f(y+th_n)dt\right)-\frac{1}{ny^2}\left(\int_{-1}^1 K(t)(y+th_n)f(y+th_n)dt\right)^2. \quad (7)$$

Let $\sigma_{n1}^2(y):=nh_nVar\big(f_{n1}(y)\big)$ and $\sigma_{n2}^2(y):=nh_nVar\big(f_{n2}(y)\big)$. From (6) and (7) it can be written that

$$\sigma_{n1}^2(y)= \quad (2)$$

$$\mu\int_{-1}^{1} K^2(t)\frac{f(y+th_n)}{y+th_n}dt - h_n\left(\int_{-1}^{1} K(t)f(y+th_n)dt\right)^2,$$

$$\sigma_{n2}^2(y) = \frac{\mu}{y^2}\int_{-1}^{1} K^2(t)(y+th_n)f(y+th_n)dt - \frac{h_n}{y^2}\left(\int_{-1}^{1} K(t)(y+th_n)f(y+th_n)dt\right)^2.$$

Another notation that is needed here is $\sigma^2(y) := \frac{\mu f(y)}{y}\int_{-1}^{1} K^2(t)dt$. $\sigma^2(y)$ will be clarified in Lemma 1.

Remark 1. *It should be noted again that $\hat{f}_{n1}(y)$, $f_{n1}(y)$ and $\sigma_{n1}^2(y)$ are defined for $y > 0$ and $\hat{f}_{n2}(y)$, $f_{n2}(y)$ and $\sigma_{n2}(y)$ are defined for $y > \varepsilon$ in which ε is anarbitrary positive constant. Therefore, in this paper wherever $\hat{f}_{n1}(y)$, $f_{n1}(y)$ and $\sigma_{n1}(y)$ are studied, y is assumed to be positive and wherever $\hat{f}_{n2}(y)$, $f_{n2}(y)$ and $\sigma_{n2}(y)$ are studied, it is assumed that $y > \varepsilon$.*

Results

Main Theorems

Before starting the main results, some assumptions are needed. Theseneeded assumptions are simillar to the assumptions that are used in Asghari et al. [1]. Here we just use the number of the assumption from Section 3 of Asghari et al. [1]. So by Assumption A1 we mean Assumption A1 from Section 3 of Asghari et al. [1].

Theorem 1. *Under Assumptions A1 and A3(i), if f is continuous in a neighbourhood of y, then we have*

$$\sup_{x\in\mathbb{R}}\left|P\left(\sqrt{nh_n}[f_{\mathrm{n}i}(y) - Ef_{\mathrm{n}i}(y)] \le x\sigma_{\mathrm{n}i}(y)\right) - \Phi(x)\right| = O\left((nh_n)^{-1/2}\right) \ for\ i = 1,2.$$

Theorem 2. *If the assumptions of Theorem 1 and A3(ii) are satisfied, we have*

$$\sup_{x\in\mathbb{R}}\left|P\left(\sqrt{nh_n}[\hat{f}_{\mathrm{n}i}(y) - Ef_{\mathrm{n}i}(y)] \le x\sigma_{\mathrm{n}i}(y)\right) - \Phi(x)\right| = O\left((nh_n)^{-1/2} + (h_n\mathrm{loglog}n)^{1/4}\right) \ for\ i = 1,2.$$

Theorem 3. *If Assumptions A1, A3-A5 are satisfied and f and G have bounded first derivatives in a neighbourhood of y, then for $y > 0$ we have*

$$\sup_{x\in\mathbb{R}}\left|P\left(\sqrt{nh_n}[\hat{f}_{\mathrm{n}i}(y) - f(y)] \le x\sigma(y)\right) - \Phi(x)\right| = O(a_n) \ for\ i = 1,2,$$

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

Remark 2. *As it can be seen from the definition of $\sigma^2(y)$ when f is unknown, $\sigma^2(y)$ is unknown too, so Theorem 3 is not applicable to inferential purposes such as creating a confidence interval or in hypotheses testing. In order to cover the applicability issue we need to estimate $\sigma^2(y)$. Two estimators are proposed for $\sigma^2(y)$. They are shown by $\hat{\sigma}_{ni}^2(y)$ for $i = 1,2$. The consistency of these proposed estimators is studied in Corollary 1. Then we present another version of Theorem 3 using this estimator instead of $\sigma^2(y)$. For $i = 1,2$ let* (8)

$$\hat{\sigma}_{ni}^2(y) := \frac{\hat{\mu}\hat{f}_{ni}(y)}{y}\int_{-1}^{1} K^2(t)dt. \quad (9)$$

In the following corollary, we substitute A3(i) with the following assumption.

A3(i)′: $h_n \to 0$ and $\frac{nh_n}{\mathrm{log}n} \to \infty$ as $n \to \infty$,

and we also have an additional assumption which is

H: There exists a small $\gamma > 0$ such that $G(y) = 0$ for all $0 < y \le \gamma$.

Corollary 1. *Let Assumptions A1, A3(i), A3(ii), A4 and A5 be satisfied,*

i: *if Assumption H is also satisfied, then for $y > 0$ we have*

$$|\hat{\sigma}_{n1}^2(y) - \sigma^2(y)| = O\left(\sqrt{\frac{\mathrm{log}n}{nh_n}} + h_n^2\right) a.s.$$

ii: for $y > \varepsilon$ we have

$$|\hat{\sigma}_{n2}^2(y) - \sigma^2(y)| = O\left(\sqrt{\frac{\mathrm{loglog}n}{nh_n}} + h_n^2\right) a.s.$$

Theorem 4. *Let Assumptions A1 and A3-A5 are satisfied,*

i: *if Assumption H is also satisfied, then for $y > 0$ we have*

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n1}(y) - f(y)\right) \le x\hat{\sigma}_{n1}(y)\right] - \Phi(x)\right| = O\left(a_n + \sqrt{\frac{\mathrm{log}n}{nh_n}}\right) a.s.$$

ii: for $y > \varepsilon$ we have

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n2}(y) - f(y)\right) \le x\hat{\sigma}_{n2}(y)\right] - \Phi(x)\right| = O\left(a_n + \sqrt{\frac{\mathrm{loglog}n}{nh_n}}\right) a.s.$$

Remark 3. *Let $h_n = O(n^{-\alpha})$ for $\frac{1}{5} < \alpha < 1$. By choosing α close to $\frac{2}{3}$, the Berry-Esseen type bound for $\hat{f}_{n1}$ and $\hat{f}_{n2}$ in Theorem 4, reduces to $O\left(n^{-\frac{1}{6}}\right)$ modulo logarithmic term as $n \to \infty$.*

Simulation study

In this section we invetigate the convergance rate of Theorem 4 via a simulation study. To simulate a length-biased sample $Y_1, \dots, Y_n$, the distribution of the population is assumed to be the gamma distribution with density fucntion $f(y) = \frac{y^3 e^{-y}}{2}$ $y > 0$.

So the length-bias density function would be $g(y) = \frac{y^4 e^{-y}}{6}$ $y > 0$.It should be noted that the distribution of the population is chosen to be the mentioned gamma distribution, so that the length-biased density satisfies Assumption H.Here 1000 samples of size n=۲00 are generated from this gamma distribution. We chose $x = 0$ for different values of y which are 0.5, 1.5, 2 and 3. The following kernel function is used

$$K(u) = \frac{3}{4}\left(1-u^2\right) I\left(|u| \le 1\right).$$

Then we evaluated $P\left[\sqrt{nh_n}\left(\hat{f}_{ni}(y) - f(y)\right) \le x\hat{\sigma}_{ni}(y)\right]$ $for\ i = 1,2$,using Monte Carlo based on 1000 replications. This estimator would be noted by $\hat{P}\left[\sqrt{nh_n}\left(\hat{f}_{ni}(y) - f(y)\right) \le x\hat{\sigma}_{ni}(y)\right]$ $for\ i = 1,2.$

Then $A_{ni}(x,y)$ $for\ i = 1,2$ is calculated. $A_{ni}(x,y)$ is defined as bellow

$$A_{ni}(x,y) := \left|\hat{P}\left[\sqrt{nh_n}\left(\hat{f}_{ni}(y) - f(y)\right) \le x\hat{\sigma}_{ni}(y)\right] - \Phi(x)\right| \quad for\ i = 1,2.$$

The used bandwidth is the bandwidth that minimizes the observed integrated square error (ISE) of $\hat{f}_{ni}(y)$ and the population density which is $f(y)$. ISE is caclulated as bellow

$$ISE\left(\hat{f}_{ni}, h_n; f\right) = \int\left[\hat{f}_{ni}(y) - f(y)\right]^2 dy,$$

and as it is mentioned $h_{ISE} = \underset{n^{-1} < h_n < n^{-\frac{2}{3}}}{\arg\min}\ ISE\left(\hat{f}_{ni}, h_n; f\right)$. Results of the performed simulation are presented in Table 1.

From Table 1, it can be concluded that for small values of y such as 0.5, the Berry-Esseen rate for $\hat{f}_{n2}(y)$ works better than$\hat{f}_{n1}(y)$. The reason appears to refer to the Assumption H. On the other hand the Berry-Esseen rate for $\hat{f}_{n1}(y)$ works better than $\hat{f}_{n2}(y)$ for higher values of y such as 1, 1.5, 2 and 3.

Table 1. Simulation results

y	$A_{n1}(0,y)$	$A_{n2}(0,y)$
0.5	0.0101	0.0007
1.5	0.0019	0.0022
2	0.0056	0.0107
3	0.0041	0.0060

Conclusion

In this paper a Berry-Esseen type bound for the kernel density estimator of length biased data is investigated . Two types of estimators are used. One is proposed by Jones [8] ($\hat{f}_{n1}$) and the other one is presented by Bhattacharya et al. [2] ($\hat{f}_{n2}$). From Theorems 3 and 4 it can be concluded that $\hat{f}_{n1}$ and $\hat{f}_{n2}$ appear to have similar normal convergence rates, but there is a principal difference between them. The results for $\hat{f}_{n1}$ are valid for all $y > 0$, but the results for $\hat{f}_{n2}$ are not valid for y's near 0.

Proofs

In this section we start the proofs to the theorems and corollary that are stated in main results section. In order to ease the procedure of the proofs, some lemmas are presented. These lemmas are stated where they are needed. From here on, let C be a positive constant that can be changed from one line to other line.

Lemma 1. *If Assumptions A1 and A3(i) are satisfied and f is continuous in a neighbourhood of y, then $\sigma_{ni}^2(y) \to \sigma^2(y)$ as $n \to \infty$ for $i = 1,2$. Furthermore, if f has bounded first derivative in a neighbourhood of y then we have*

$$|\sigma_{ni}^2(y) - \sigma^2(y)| = O(h_n),\ fori = 1,2.$$

Proof. Since f is continuous in a neighbourhood of y, it is also bounded in this neighbourhood. So for $\sigma_{n1}^2(y)$ since $h_n \to 0$ $asn \to \infty$ we have

$$\int_{-1}^{1} K^2(t)\frac{f(y+th_n)}{y+th_n}dt \to \frac{f(y)}{y}\int_{-1}^{1} K^2(t)dt,$$

$$h_n\left(\int_{-1}^{1} K(t)f(y+th_n)\right)^2 \to 0,$$

as $n \to \infty$. So using (8) we have $\sigma_{n1}^2(y) \to \sigma^2(y) asn \to \infty,$

and for $\sigma_{n2}^2(y)$ we have

$$\int_{-1}^{1} K^2(t)(y+th_n)f(y+th_n)dt \to yf(y)\int_{-1}^{1} K^2(t)dt \quad as\, n \to \infty,$$

$$h_n\left(\int_{-1}^{1} K(t)(y+th_n)f(y+th_n)dt\right)^2 \to 0 \quad as\, n \to \infty,$$

so from (9) we have
$\sigma_{n2}^2(y) \to \sigma^2(y)$ as $n \to \infty$.

Now using Assumption A1, if f has bounded first derivative in a neighbourhood of y, it can be written that

$$|\sigma_{n1}^2(y)-\sigma^2(y)| = \left|\mu\int_{-1}^{1} K^2(t)\frac{f(y+th_n)}{y+th_n}dt - h_n\left(\int_{-1}^{1} K(t)f(y+th_n)dt\right)^2 - \frac{\mu f(y)}{y}\int_{-1}^{1} K^2(t)dt\right| = O(h_n),$$

$$|\sigma_{n2}^2(y)-\sigma^2(y)| = \left|\frac{\mu}{y^2}\int_{-1}^{1} K^2(t)(y+th_n)f(y+th_n)dt - \frac{h_n}{y^2}\left(\int_{-1}^{1} K(t)(y+th_n)f(y+th_n)dt\right)^2 - \frac{\mu f(y)}{y}\int_{-1}^{1} K^2(t)dt\right| = O(h_n),$$

and the proof is completed.

Proof of Theorem 1. Let

$$W_{ni} := \frac{\mu}{nh_n}\left\{\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right) - E\left[\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right)\right]\right\}, \quad (12)$$

$$V_{ni} := \frac{\mu}{nh_n y}\left\{K\left(\frac{Y_i-y}{h_n}\right) - E\left[K\left(\frac{Y_i-y}{h_n}\right)\right]\right\},$$

so it can be written that

$$f_{n1}(y) - E[f_{n1}(y)] = \sum_{i=1}^{n} W_{ni},$$

$$f_{n2}(y) - E[f_{n2}(y)] = \sum_{i=1}^{n} V_{ni}.$$

Let $B_{n1} := \sum_{i=1}^{n} \mathrm{Var}(W_{ni})$ and $B_{n2} := \sum_{i=1}^{n} \mathrm{Var}(V_{ni})$. Now from Theorem 5.7 of Petrov [16] it can be concluded that

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(f_{n1}(y)-Ef_{n1}(y)\right) \le x\sigma_{n1}(y)\right] - \Phi(x)\right| = \sup_{x\in\mathbb{R}}\left|P\left[\frac{\sum_{i=1}^{n} W_{ni}}{\sqrt{Var\ \left(\sum_{i=1}^{n} W_{ni}\right)}} \le x\right] - \Phi(x)\right| \le \frac{C}{B_{n1}^{\frac{3}{2}}}\sum_{i=1}^{n} E|W_{ni}|^3 = \frac{C}{B_{n1}^{\frac{3}{2}}} nE|W_{n1}|^3,$$

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(f_{n2}(y)-Ef_{n2}(y)\right) \le x\sigma_{n2}(y)\right] - \Phi(x)\right| = \sup_{x\in\mathbb{R}}\left|P\left[\frac{\sum_{i=1}^{n} V_{ni}}{\sqrt{Var\ \left(\sum_{i=1}^{n} V_{ni}\right)}} \le x\right] - \Phi(x)\right| \le \frac{C}{B_{n2}^{\frac{3}{2}}}\sum_{i=1}^{n} E|V_{ni}|^3 = \frac{C}{B_{n2}^{\frac{3}{2}}} nE|V_{n1}|^3.$$

From the definition of $E|W_{n1}|^3$ and $E|V_{n1}|^3$ and by using the C_r inequality (Gut [6]) we have

$$E|W_{n1}|^3 \le \frac{C}{(nh_n)^3}E\left|\frac{1}{Y_1}K\left(\frac{y-Y_1}{h_n}\right)\right|^3 = \frac{C}{n^3h_n^2}\int_{-1}^{1} K^3(t)\frac{f(y+th_n)}{(y+th_n)^2}dt, \quad (16)$$

$$E|V_{n1}|^3 \le \frac{C}{(nh_n y)^3}E\left|K\left(\frac{y-Y_1}{h_n}\right)\right|^3 = \frac{C}{n^3h_n^2}\int_{-1}^{1} K^3(t)(y+th_n)f(y+th_n)dt,$$

so using Assumption A1, Lemma 1, (16) and (17) it can be written that

$$(14) \le \frac{C}{\left(\frac{\sigma_{n1}^2(y)}{nh_n}\right)^{\frac{3}{2}}(nh_n)^2}\int_{-1}^{1} K^3(t)\frac{f(y+th_n)}{(y+th_n)^2}dt \quad (13)$$

$$= \frac{C}{\sigma_{n1}^3(y)(nh_n)^{\frac{1}{2}}}\int_{-1}^{1} K^3(t)\frac{f(y+th_n)}{(y+th_n)^2}dt$$

$$= O\left((nh_n)^{-\frac{1}{2}}\right),$$

and

$$(15) \leq \frac{C}{\left(\frac{\sigma_{n2}^2(y)}{nh_n}\right)^{\frac{3}{2}}(nh_n)^2}\int_{-1}^{1} K^3(t)(y+th_n)f(y+th_n)dt$$
$$= \frac{C}{\sigma_{n2}^3(y)(nh_n)^{\frac{1}{2}}}\int_{-1}^{1} K^3(t)(y+th_n)f(y+th_n)dt$$
$$= O\left((nh_n)^{-\frac{1}{2}}\right).$$

In the last equalities in (18) and (19) we used Assumption A1 and the fact that f is bounded in a neighbourhood of y. □

Proof of Theorem 2. Using Lemma 2 of Asghari et al. [1], for any $a_1 > 0$ and $a_2 > 0$ it can be written that

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n1}(y) - Ef_{n1}(y)\right) \leq x\sigma_{n1}(y)\right] - \Phi(x)\right|$$
$$\leq \sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(f_{n1}(y) - Ef_{n1}(y)\right) \leq x\sigma_{n1}(y)\right] - \Phi(x)\right| + \frac{a_1}{\sqrt{2\pi}}$$
$$+P\left(\frac{\sqrt{nh_n}}{\sigma_{n1}(y)}\left|\hat{f}_{n1}(y) - f_{n1}(y)\right| > a_1\right),$$

and

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n2}(y) - Ef_{n2}(y)\right) \leq x\sigma_{n2}(y)\right] - \Phi(x)\right|$$
$$\leq \sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(f_{n2}(y) - Ef_{n2}(y)\right) \leq x\sigma_{n2}(y)\right] - \Phi(x)\right| + \frac{a_2}{\sqrt{2\pi}}$$
$$+P\left(\frac{\sqrt{nh_n}}{\sigma_{n2}(y)}\left|\hat{f}_{n2}(y) - f_{n2}(y)\right| > a_2\right),$$

on the other hand we have

$$P\left(\frac{\sqrt{nh_n}}{\sigma_{ni}(y)}\left|\hat{f}_{ni}(y) - f_{ni}(y)\right| > a_i\right) \leq \frac{\sqrt{nh_n}}{a_i\sigma_{ni}(y)}E\left|\hat{f}_{ni}(y) - f_{ni}(y)\right| \; for\; i = 1,2. \quad (20)$$

By the law of the iterated logarithm for partial sums of i.i.d. random variables, we have

$$|\hat{\mu} - \mu| = O\left(\sqrt{\frac{\text{loglog}n}{n}}\right) a.s. \quad (21)$$

Now using Assumption A1, continuousness of f in a neighbourhood of y and (23), it can be written that

$$E\left|\hat{f}_{n1}(y) - f_{n1}(y)\right| \leq \frac{1}{nh_n}\sum_{i=1}^{n} E\left[\frac{|\hat{\mu} - \mu|}{Y_i}K\left(\frac{Y_i - y}{h_n}\right)\right]$$
$$\leq C\frac{(\text{loglog}n)^{1/2}}{n^{1/2}h_n}E\left[\frac{1}{Y_1}K\left(\frac{Y_1 - y}{h_n}\right)\right]$$
$$= C\left(\frac{\text{loglog}n}{n}\right)^{1/2}\int_{-1}^{1} K(t)\frac{f(y+th_n)}{y+th_n}dt$$
$$= O\left(\sqrt{\frac{\text{loglog}n}{n}}\right), \quad (19)$$

and

$$E\left|\hat{f}_{n2}(y) - f_{n2}(y)\right| \leq \frac{1}{nh_n y}\sum_{i=1}^{n} E\left[|\hat{\mu} - \mu|K\left(\frac{Y_i - y}{h_n}\right)\right]$$
$$\leq C\frac{(\text{loglog}n)^{1/2}}{n^{1/2}h_n y}E\left[K\left(\frac{Y_1 - y}{h_n}\right)\right]$$
$$= C\frac{(\text{loglog}n)^{1/2}}{n^{1/2}y}\int_{-1}^{1} K(t)f(y+th_n)dt$$
$$= O\left(\sqrt{\frac{\text{loglog}n}{n}}\right).$$

Now by letting $a_1 = a_2 = (h_n\text{loglog}n)^{\frac{1}{4}}$ in (20) and (21), substituting (24) and (25) in (22) and by using Theorem 1 we get the result. □

Proof of Theorem 3. By triangular inequality and using Lemma 2 of Asghari et al. [1] for $a = \frac{\sqrt{nh_n}}{\sigma(y)}|Ef_{n1}(y) - f(y)|$, we have

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n1}(y) - f(y)\right) \leq x\sigma(y)\right] - \Phi(x)\right|$$
$$\leq \sup_{x\in\mathbb{R}}\left|P\left[\frac{\sqrt{nh_n}}{\sigma_{n1}(y)}\left(\hat{f}_{n1}(y) - Ef_{n1}(y)\right) \leq x\frac{\sigma(y)}{\sigma_{n1}(y)}\right] - \Phi\left(\frac{\sigma(y)}{\sigma_{n1}(y)}x\right)\right|$$
$$+\sup_{x\in\mathbb{R}}\left|\Phi\left(\frac{\sigma(y)}{\sigma_{n1}(y)}x\right) - \Phi(x)\right| + \frac{\sqrt{nh_n}}{\sqrt{2\pi}\sigma(y)}|Ef_{n1}(y) - f(y)|. \quad (22)$$

It should be noted that in (26) we used the fact that the event $\frac{\sqrt{nh_n}}{\sigma(y)}|Ef_{n1}(y) - f(y)| > a$ does not happen for the sellected a.

Also by letting $a = \frac{\sqrt{nh_n}}{\sigma(y)}|Ef_{n2}(y) - f(y)|$,in Lemma 2 of Asghari et al. [1] and using the same argument, we have (23)

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n2}(y) - f(y)\right) \leq x\sigma(y)\right] - \Phi(x)\right|$$
$$\leq \sup_{x\in\mathbb{R}}\left|P\left[\frac{\sqrt{nh_n}}{\sigma_{n2}(y)}\left(\hat{f}_{n2}(y) - Ef_{n2}(y)\right) \leq x\frac{\sigma(y)}{\sigma_{n2}(y)}\right] - \right.$$

$\Phi\left(\frac{\sigma(y)}{\sigma_{n2}(y)}x\right)\Big|$
$+\sup_{x\in\mathbb{R}}\left|\Phi\left(\frac{\sigma(y)}{\sigma_{n2}(y)}x\right)-\Phi(x)\right|+$
$\frac{\sqrt{nh_n}}{\sqrt{2\pi}\sigma(y)}|Ef_{n2}(y)-f(y)|,$

a little calculation and a usage of Lemma 1 yields

$$\sup_{x\in\mathbb{R}}\left|\Phi\left(\frac{\sigma(y)}{\sigma_{\mathrm{ni}}(y)}x\right)-\Phi(x)\right| = O(|\sigma_{\mathrm{ni}}^2(y)-\sigma^2(y)|) = O(h_n)\ for\ i=1,2.$$

On the other hand using a Taylor expansion gives
$|Ef_{\mathrm{ni}}(y)-f(y)| = O(h_n^2) for\ i=1,2,$
now by substituting (29) in (26) and (27), using Theorem 3 and (28) we get the desired result. □

Proof of Corollary 1. From the definition of $\hat{\sigma}_{n1}^2$ and σ^2 we have

$$|\hat{\sigma}_{n1}^2(y)-\sigma^2(y)| = \left|\frac{\hat{\mu}\hat{f}_{n1}(y)-\mu f(y)}{y}\right|\int_{-1}^{1}K^2(t)dt \leq C\{f(y)|\hat{\mu}-\mu|+\hat{\mu}|\hat{f}_{n1}(y)-f(y)|\}. \quad (30)$$

For the second part of (30), the triangle inequality gives

$$|\hat{f}_{n1}(y)-f(y)| \leq \sup_{y\in\mathbb{R}}|f_{n1}(y)-Ef_{n1}(y)| + |\hat{f}_{n1}(y)-f_{n1}(y)| + |Ef_{n1}(y)-f(y)| =: \mathrm{I}+\mathrm{II}+\mathrm{III}.$$

Under the assumptions of the corollary, using the method that is used in the proving procedure of Lemma 1 of Ould-Saïd and Tatachak [13], results the following

$$\mathrm{I} = O\left(\sqrt{\frac{\log n}{nh_n}}\right)a.s.$$

For II we have

$$\mathrm{II} = \left|\frac{\hat{\mu}}{nh_n}\sum_{i=1}^{n}\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right)-\frac{\mu}{nh_n}\sum_{i=1}^{n}\frac{1}{Y_i}K\left(\frac{Y_i-y}{h_n}\right)\right| \leq \frac{|\hat{\mu}-\mu|}{\gamma}\left[\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{Y_i-y}{h_n}\right)\right] = \frac{|\hat{\mu}-\mu|}{\gamma}g_n(y) \quad (33)$$
$$= O\left(\sqrt{\frac{\log\log n}{n}}\right)a.s.$$

In (33), $g_n(y)$ is a kernel density estimator of $g(\cdot)$ based on $Y_1,\ldots,Y_n$. Here we used the fact that $g_n(y)$ converges to $g(y)$. Recall that from the proof of Theorem 3 we have

$$\mathrm{III} = O(h_n^2).$$

Substituting (32), (34) and (35) in (31) gives

$$(31) = O\left(\sqrt{\frac{\log n}{nh_n}}+h_n^2\right)a.s. \quad (27)$$

Now using (23) and (36) it can be concluded that

$$(30) = O\left(\sqrt{\frac{\log n}{nh_n}}+h_n^2\right)a.s. \quad (28)$$

For $\hat{\sigma}_{n2}^2$ we have (29)

$$|\hat{\sigma}_{n2}^2(y)-\sigma^2(y)| = \left|\frac{\hat{\mu}\hat{f}_{n2}(y)-\mu f(y)}{y}\right|\int_{-1}^{1}K^2(t)dt \leq C\{f(y)|\hat{\mu}-\mu|+\hat{\mu}|\hat{f}_{n2}(y)-f(y)|\}. \quad (37)$$

For the second part of (37), by using the triangle inequality it can be written

$$|\hat{f}_{n2}(y)-f(y)| \leq |\hat{f}_{n2}(y)-E\hat{f}_{n2}(y)| + |E\hat{f}_{n2}(y)-f(y)| =: \mathrm{I}'+\mathrm{II}'.$$

Under Assumptions A1, A3(i) and A3(ii), Theorem 2 of Hall [7] gives the following

$$\mathrm{I}' = O\left(\sqrt{\frac{\log\log n}{nh_n}}\right)a.s., \quad (31)$$

and on the other hand from the proof of Theorem 3 we know that
$\mathrm{II}' = O(h_n^2),$ (32)
now using (37), (38), (39) and (40), the corollary is proved.

Proof of Theorem 4. Here we only give the proof to the first part ($\hat{f}_{n1}$). The proof to the second part ($\hat{f}_{n2}$) is analogous. From Corollary 1 we have

$$\sup_{x\in\mathbb{R}}\left|\Phi\left(x\frac{\hat{\sigma}_{n1}(y)}{\sigma(y)}\right)-\Phi(x)\right| = O(|\sigma_{n2}^2(y)-\sigma^2(y)|) = O\left(\sqrt{\frac{\log n}{nh_n}}+h_n^2\right)a.s. \quad (34)$$

By using triangular inequality, it can be concluded that

$$\sup_{x\in\mathbb{R}}\left|P\left[\sqrt{nh_n}\left(\hat{f}_{n1}(y)-f(y)\right)\leq x\hat{\sigma}_{n1}(y)\right]-\Phi(x)\right|$$

$$\leq \sup_{x\in\mathbb{R}} \left| P\left[\frac{\sqrt{nh_n}}{\sigma(y)}\left(\hat{f}_{n1}(y)-f(y)\right) \leq x\frac{\hat{\sigma}_{n1}(y)}{\sigma(y)}\right] - \Phi\left(x\frac{\hat{\sigma}_{n1}(y)}{\sigma(y)}\right)\right|$$

$$+\sup_{x\in\mathbb{R}} \left|\Phi\left(x\frac{\hat{\sigma}_{n1}(y)}{\sigma(y)}\right) - \Phi(x)\right|$$

$$= \sup_{x\in\mathbb{R}} \left| P\left[\frac{\sqrt{nh_n}}{\sigma(y)}\left(\hat{f}_{n1}(y)-f(y)\right) \leq x\right] - \Phi(x)\right| + O\left(\sqrt{\frac{\log n}{nh_n}} + h_n^2\right) a.s. \quad (42)$$

By using Theorem 3, we get the desired result. The same argument for $\hat{f}_{n2}$ gives the result.

Acknowledgement

The authors would like to sincerely thank anonymous referee for the careful reading of the manuscript**.**

References

1. Asghari P., Fakoor V., and Sarmad M. A Berry-Esseen type bound in kernel density estimation for a random left-truncation Model. *Communications for Statistical Applications and Methods* **21**(2): 115–124 (2014).
2. Bhattacharyya B. B., Franklin L. A. and Richardson G. D., A comparison of non-parametric unweighed and length-biased density estimation of fibres. *Commun. Stat-Theor. M.***17**: 3629–3644(1988).
3. Blumenthal S. Proportional sampling in life length studies. *Technometrics***9**: 205–218 (1967).
4. Chang M. N., and Rao P. V. Berry-Esseen bound for the Kaplan-Meier estimator. *Commun. Stat-Theor. M.***18**(12): 4647–4664 (1989).
5. Cox D. R. *Some sampling problems in technology, new developments in survey sampling. Edited by Johnson and Smith*, Wiley. (1969).
6. Gut A. *Probability: A graduate course*. Springer, New York. (2005).
7. Hall P. Laws of the iterated logarithm for nonparametric density estimators. *Z. Wahrscheinlichkeit.***56**: 47–61 (1981).
8. Jones M. C. Kernel density estimation for length biased data. *Biometrika***78**: 511–519 (1991).
9. Isogai E. A Berry-Esseen-type bound for recursive estimators of a density and its derivatives. *J. Stat. Plan. Infer.***40**: 1–14 (1994).
10. Liang H. Y., and Uña-Álvarez J. A Berry-Esseen type bound in kernel density estimation for strong mixing censored samples. *J. Multivariate. Anal.***100**: 1219–1231 (2009).
11. Liuquan S. and Lixing Z. A Berry-Esseen Type Bound for Kernel Density Estimators under Random Censorship. *Acta. Math. Sin*. **42**(4): 627–636 (1999).
12. Mcfadden J. A. On the length of intervals in a stationary point process. *Journal of Royal Statistical Society.***24**(2): 364–382 (1962).
13. Ould-Saïd E., and Tatachak A. On the non-parametric estimation of the simple mode under random left-truncation model. *Rev. Roum. Math. Pure. A,* **54**(3): 243–266 (2009).
14. Parzen E. On estimation of a probability density function and mode. *Ann. Math. Stat.***33**(3): 1065–1076 (1962).
15. Prakasa Rao B. L. S. Berry-Esseen type bound for density estimation of stationary Markov processes. *Indian Statistical Institute* 15–21 (1975).
16. Petrov V. V. *Limit theorems of probability theory*. Oxford University Press Inc., New York (1995).
17. Rosenblatt M. Density estimates and markov sequences. *Nonparametric Techniques in Statistical Inference* 199–213 (1970).
18. Wicksell S. D. The corpuscle problem. A mathematical study of a Biometric problem. *Biometrika.* **17**: 84–99 (1925).