mathematics

MDPI

Article

On Consistency of the Nearest Neighbor Estimator of the Density Function for *m*-AANA Samples

Xin Liu, Yi Wu, Wei Wang and Yong Zhu *

Center of Applied Mathematics, School of Big Data and Artificial Intelligence, Chizhou University, Chizhou 247000, China; liuxin86562@163.com (X.L.); czu_wangwei@163.com (W.W.)
* Correspondence: zyong@czu.edu.cn

Abstract: In this paper, by establishing a Bernstein inequality for m-asymptotically almost negatively associated random variables, some results on consistency for the nearest neighbor estimator of the density function are further established. The results generalize some existing ones in the literature. Some numerical simulations are also provided to support the results.

Keywords: nearest neighbor estimator; weak consistency; m-asymptotically almost negatively associated samples; strong consistency; uniform consistency

MSC: 62G05

check for updates

Citation: Liu, X.; Wu, Y.; Wang, W.; Zhu, Y. On Consistency of the Nearest Neighbor Estimator of the Density Function for m-AANA Samples. *Mathematics* **2023**, *11*, 4391. https://doi.org/10.3390/math11204391

Academic Editor: Ionut Florescu

Received: 19 September 2023
Revised: 13 October 2023
Accepted: 19 October 2023
Published: 23 October 2023

1. Introduction

Nearest neighbor estimators can be used for many flexible questions and data types. Let X be a random variable whose density function $f(x)$ is unknown and needs to be estimated. Let $X_1, X_2, \cdots, X_n$ be the sample drawn from population X. To estimate $f(x)$, Loftsgarden and Quesenberry [1] raised the nearest neighbour estimator $f_n(x)$ as follows:

$$f_n(x) = \frac{k_n}{2na_n(x)}, \tag{1}$$

where $1 \leq k_n \leq n$ and

$$a_n(x) = \min\{\alpha : \text{ the number of } X_i \in [x-\alpha, x+\alpha] \text{ is no less than } k_n\}.$$

Since Loftsgarden and Quesenberry [1] put forward the method of estimating the density function, many scholars showed their interest in this field. For some recent examples, Liu and Wu [2] established the Bernstein inequality to deal with the consistency results under negatively dependent samples; Lu et al. [3] investigated some results on consistency and convergence rate for this estimator based on φ-mixing samples; Liu and Zhang [4] established the consistency and asymptotic normality of the estimator based on α-mixing samples; Yang [5] established various results on the consistency of the estimator based on negatively associated (NA, in short) samples; Wang and Hu [6] obtained the corresponding results for widely orthant dependent (WOD, in short) samples, which extend and improve those of Yang [5] for NA samples and further proved the rates of strong consistency and uniformly strong consistency; Lan and Wu [7] investigated the rate of uniform strong consistency for the estimator under extended negatively dependent (END, in short) samples; and Wang and Wu [8] extended and improved the results of Lan and Wu [7] from END samles to m-extended negatively dependent (m-END, in short) samples and obtained the same rates as that of END samples.

This paper will further study this topic and extend those aforementioned results to a more general setting. Now, we are at a position to recall some concepts of dependent random variables, of which the first one is that of asymptotically almost negatively associated (AANA, in short) random variables, which was first raised by Chandra and Ghosal [9] as follows.

Definition 1. *We call a sequence $\{Z_n, n \geq 1\}$ of random variables to be AANA if there is a nonnegative sequence satisfying $\lim_{n\to\infty} q(n) = 0$ such that for all $n, l \geq 1$ and for all coordinatewise nondecreasing functions f_1 and f_2,*

$$Cov(f_1(Z_n), f_2(Z_{n+1}, Z_{n+2}, \cdots, Z_{n+l})) \leq q(n)[Var(f_1(Z_n))Var(f_2(Z_{n+1}, Z_{n+2}, \cdots, Z_{n+l}))]^{1/2}$$

whenever the variances above exist.

Since the concept of AANA random variables was put forward by Chandra and Ghosal [9], plenty of results have been established concerning this dependence structure. For instance, Kim and Ko [10] developed the Hajeck–Renyi inequality for these dependent random variables; Yuan and An [11] established some moment inequalities for maximum sums; Chandra and Ghosal [12] as well as Shen and Wu [13] proved the strong law of large numbers for weighted sums; Yuan and An [14] investigated the laws of large numbers for this dependent random variables satisfying the Cesàro alpha-integrability condition; and Wu and Wang [15] studied some results on the nearest neighbor estimator of the density function under AANA samples.

As an extension of AANA random variables, the concept of m-AANA random variables was raised by Nam et al. [16] as follows.

Definition 2. *Let m be a positive integer. We say that a sequence $\{Z_n, n \geq 1\}$ of random variables is m-AANA if there exists a nonnegative sequence $q(n) \to 0$ as $n \to \infty$ such that for all $n, l \geq m$ and for all coordinatewise nondecreasing functions f_1 and f_2,*

$$Cov(f_1(Z_n), f_2(Z_{n+m}, \cdots, Z_{n+l})) \leq q(n)[Var(f_1(Z_n))Var(f_2(Z_{n+m}, \cdots, Z_{n+l}))]^{1/2}$$

whenever the variances exist.

It is known that many multivariate distributions satisfy the NA property. The concept of AANA random variables will degenerate to that of NA random variables by taking $q(n) = 0$. It is easy to see that the m-AANA sequence is equivalent to AANA with $m = 1$. Therefore, the structure of m-AANA random variables includes AANA random variables, m-NA random variables, NA random variables, moving average processes, and independent random variables as special cases, and thus it is a more plausible assumption in realistic applications. Now, we present an example of m-AANA random variables that are not necessarily AANA.

Example 1. *Let $\{Y_n, n \geq 1\}$ be independent and identically distributed $N(0, 1)$ random variables and define $X_n = (1 + a_n^2)^{-1/2}(Y_n + a_n Y_{n+1})$, where $a_n > 0$ and $a_n \to 0$. It follows from Chandra and Ghosal [9] that $\{X_n, n \geq 1\}$ is a sequence of AANA random variables that is not NA. Now, we define for each $n \geq 1$ that $Z_{m(n-1)+1} = \cdots = Z_{mn} = X_n$ with $m \geq 2$. Then, it is easy to check that the sequence $\{Z_n, n \geq 1\}$ is m-AANA. However, it is not AANA since the condition $\lim_{n\to\infty} q(n) = 0$ is not satisfied if we take $l = 1$, for example.*

In this paper, motivated by the literature above, we first establish a Bernstein inequality for m-asymptotically almost negatively associated (m-AANA, in short) random variables, which is of interest itself. By using this inequality, we further investigate some results on the consistency of the nearest neighbor estimator under m-AANA samples. These results are generalizations of the corresponding ones of Wu and Wang [15] from AANA samples to m-AANA samples.

The layout of this paper is as follows. Some preliminary lemmas are stated in Section 2. Section 3 includes the main results, while the numerical simulations are given in Section 4 to support the theoretical results. The proofs of our main results are postponed in Section 5. The paper is concluded in Section 6. Throughout this paper, $\lfloor x \rfloor$ stands for the integer part of x. Let $\log x = \max\{1, \ln x\}$. Indicator function $I(A) = 1$ if the set A occurs or $I(A) = 0$ otherwise. $C(f) = \{x : f \text{ is continuous at } x\}$. C and c_0 stand for positive constants whose values are not necessarily the same in each appearance. All limits are taken as $n \to \infty$ unless specified otherwise.

2. Preliminary Lemmas

To prove the main results, we first provide several important lemmas in this section.

Lemma 1 (cf. [14])**.** *Suppose that $\{X_n, n \geq 1\}$ is a sequence of AANA random variables with mixing coefficients $\{q(n), n \geq 1\}$. If $f_n(\cdot), n \geq 1$ are all nondecreasing or all nonincreasing, then $\{f_n(X_n), n \geq 1\}$ is still a sequence of AANA random variables with the same mixing coefficients.*

A combination of Lemma 1 and Definition 2 yields the following lemma, which is obvious, and thus the proof is omitted.

Lemma 2. *Suppose that $\{X_n, n \geq 1\}$ is a sequence of m-AANA random variables with mixing coefficients $\{q(n), n \geq 1\}$. If $f_n(\cdot), n \geq 1$ are all nondecreasing or all nonincreasing; then, $\{f_n(X_n), n \geq 1\}$ is still a sequence of m-AANA random variables with the same mixing coefficients.*

Lemma 3 (cf. [15])**.** *Let $\{X_n, n \geq 1\}$ be a sequence of AANA random variables with zero means and mixing coefficients $\{q(n), n \geq 1\}$. Assume that $|X_n|$ is bounded by a positive number b for each $n \geq 1$. Then, a positive constant C exists such that for all $n \geq 1$ and $\varepsilon > 0$,*

$$P\left(\left|\sum_{i=1}^{n} X_i\right| \geq \varepsilon\right) \leq C\left[\sum_{k=1}^{n-1} q(k) + 1\right] \cdot \exp\left\{-\frac{\varepsilon^2}{2\sum_{i=1}^{n} EX_i^2 + \frac{2}{3}b\varepsilon}\right\}. \tag{2}$$

By virtue of Lemma 3, we can further prove the Bernstein inequality for m-AANA random variables. The lemma will play a significant role in the proof of the main results.

Lemma 4. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA random variables with zero means and mixing coefficients $\{q(n), n \geq 1\}$. Assume that $|X_n|$ is bounded by a positive number b for each $n \geq 1$. Then, a positive constant C exists such that for all $n \geq 1$ and $\varepsilon > 0$,*

$$P\left(\left|\sum_{i=1}^{n} X_i\right| \geq \varepsilon\right) \leq Cm\left[\sum_{k=1}^{n-1} q(k) + 1\right] \cdot \exp\left\{-\frac{\frac{\varepsilon^2}{m^2}}{2\sum_{i=1}^{n} X_i + \frac{2}{3m}b\varepsilon}\right\}.$$

Proof. For all sufficiently large n, positive integers $j \geq 0$ and $1 \leq l \leq m$ always exist satisfying $n = mj + l$. Without a loss of generality, we may define that $X_i = 0$ for all $n < i \leq m(j+1)$. Thus, $\sum_{i=1}^{n} X_i$ can be decomposed as

$$\sum_{i=1}^{n} X_i = \sum_{l=1}^{m} \sum_{i=0}^{j} X_{mi+l}$$

where $\{X_{mi+l}, 0 \leq i \leq j\}$ are AANA for each given $l = 1, 2, \cdots, m$. Thus, we can obtain from Lemma 3 that

$$\begin{aligned}
P(|S_n| \geq \varepsilon) &= P\left(\left|\sum_{l=1}^{m}\sum_{i=0}^{j} X_{mi+l}\right| \geq \varepsilon\right) \\
&\leq P\left(\bigcup_{l=1}^{m}\left|\sum_{i=0}^{j} X_{mi+l}\right| \geq \frac{\varepsilon}{m}\right) \\
&\leq \sum_{l=1}^{m} P\left(\left|\sum_{i=0}^{j} X_{mi+l}\right| \geq \frac{\varepsilon}{m}\right) \\
&\leq C\sum_{l=1}^{m}\left[\sum_{k=1}^{j-1} q(k)+1\right]\cdot \exp\left\{-\frac{\frac{\varepsilon^2}{m^2}}{2\times\sum_{i=0}^{j} E(X_{mi+l})^2+\frac{2}{3m}b\varepsilon}\right\} \\
&\leq Cm\left[\sum_{k=1}^{n-1} q(k)+1\right]\cdot \exp\left\{-\frac{\frac{\varepsilon^2}{m^2}}{2B_n^2+\frac{2}{3m}b\varepsilon}\right\}
\end{aligned}$$

This completes the proof of the lemma. □

Lemma 5 (cf. [5]). *Let $Z_1, Z_2, \cdots, Z_n$ follow a common distribution $F(z)$, which is continuous. For $n \geq 3$, assume that z_{ni} satisfies $F(z_{ni}) = i/n$ for each $1 \leq i \leq n-1$. Then,*

$$\sup_{-\infty<z<\infty} |F_n(z) - F(z)| \leq \max_{1\leq i\leq n-1} |F_n(z_{ni}) - F(z_{ni})| + 2/n,$$

where $F_n(z) = n^{-1}\sum_{j=1}^{n} I(Z_j < z)$ is the empirical distribution function.

Lemma 6. *Let $\{Z_n, n \geq 1\}$ be a sequence of m-AANA random variables, with $F(z)$ and $f(z)$ being the distribution function and density function, respectively. Let $\{\kappa_n, n \geq 1\}$ be a sequence of positive numbers satisfying $\kappa_n \to 0$ such that $\liminf_{n\to\infty} n\kappa_n^2/\log n \geq c_0 > 0$. Then, for any $D_0 > 0$ large enough,*

$$\sum_{n=1}^{\infty} P\left(\sup_z |F_n(z) - F(z)| > D_0\kappa_n\right) < \infty.$$

In particular,

$$\sum_{n=1}^{\infty} P\left(\sup_z |F_n(z) - F(z)| > D_0(\log n/n)^{1/2}\right) < \infty.$$

Proof. Observing that $n\kappa_n \to \infty$, we have that $2/n < D_0\kappa_n/2$ for all sufficiently large n and any positive constant D_0, the value of which will be specified later. It follows from Lemma 5 that

$$\begin{aligned}
P\left(\sup_x |F_n(x) - F(x)| > D_0\kappa_n\right) &\leq P\left(\max_{1\leq i\leq n-1} |F_n(x_{ni}) - F(x_{ni})| > D_0\kappa_n/2\right) \\
&\leq \sum_{i=1}^{n-1} P(|F_n(z_{ni}) - F(z_{ni})| > D_0\kappa_n/2). \qquad (3)
\end{aligned}$$

Let $Z_j(z_{ni}) = I(Z_j < z_{ni}) - EI(Z_j < z_{ni})$. By Lemma 2, we know that $\{Z_j(z_{ni}), j \geq 1\}$ is still a sequence of m-AANA random variables with $EZ_j(z_{ni}) = 0$, $|Z_j(z_{ni})| \leq 1$ and $E(Z_j(z_{ni}))^2 \leq 1$. Thus, by Lemma 4 we have that for all n adequately large,

$$\begin{aligned}
P(|F_n(z_{ni}) - F(z_{ni})| > D_0\kappa_n/2) &= P\left(\left|\sum_{j=1}^{n} Z_j(z_{ni})\right| > D_0 n\kappa_n/2\right) \\
&\leq C_m\left[\sum_{k=1}^{n-1} q(k) + 1\right] \cdot \exp\left\{\frac{-\frac{D_0^2 n^2 \kappa_n^2}{m^2}}{8B_n^2 + \frac{4}{3m} D_0 n\kappa_n}\right\} \\
&\leq Cn \exp\left\{-\frac{D_0^2}{9m^2} n\kappa_n^2\right\} \\
&\leq Cn \exp\left\{-\frac{c_0 D_0^2}{18m^2} \log n\right\} \\
&\leq Cn^{1-\frac{c_0 D_0^2}{18m^2}}. \qquad (4)
\end{aligned}$$

Taking D_0 sufficiently large such that $1 - \frac{c_0 D_0^2}{18m^2} < -2$, by (3) and (4) we have

$$\sum_{n=1}^{\infty} P\left(\sup_z |F_n(z) - F(z)| > D_0\kappa_n\right) \leq C\sum_{n=1}^{\infty}\sum_{j=1}^{n-1} n^{1-\frac{c_0 D_0^2}{18m^2}} < \infty.$$

This completes the proof of the lemma. □

3. Main Results

Now, we state our results one by one as follows. Denote $\chi_n = \sum_{k=1}^{n-1} q(k) + 1$. The first one concerns the weak consistency of the nearest neighbor density estimator.

Theorem 1. *Suppose that $\{X_n, n \geq 1\}$ is a sequence of m-AANA samples and $k_n/n \to 0$, $k_n^2/n \to \infty$. If*

$$\lim_{n\to\infty} \chi_n \cdot \exp\left\{-\frac{\gamma k_n^2}{n}\right\} = 0 \qquad (5)$$

for all $\gamma > 0$, then for all $x \in c(f)$,

$$f_n(x) \xrightarrow{P} f(x).$$

Remark 1. *We point out that (5) is easy to verify. For example, if $\sum_{n=1}^{\infty} q(n) < \infty$, which is frequently adopted in the literature, we have $\chi_n \leq 1 + \sum_{n=1}^{\infty} q(n) < \infty$ and thus (5) follows. Moreover, if $k_n^2/(n \log n) \to \infty$, (5) also holds without any restriction on the mixing coefficients. We give it in the following corollary.*

Corollary 1. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and $k_n/n \to 0$, $k_n^2/(n \log n) \to \infty$. Then, for all $x \in c(f)$,*

$$f_n(x) \xrightarrow{P} f(x).$$

Under some slightly stronger conditions, one can obtain the following results on complete consistency.

Theorem 2. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and $k_n/n \to 0$, $k_n^2/n \to \infty$. If*

$$\sum_{n=1}^{\infty} \chi_n \exp\left\{-\frac{\gamma k_n^2}{n}\right\} < \infty \qquad (6)$$

for all $\gamma > 0$, then for all $x \in c(f)$,

$$\sum_{n=1}^{\infty} P(|f_n(x) - f(x)| > \varepsilon) < \infty$$

for all $\varepsilon > 0$, and hence

$$f_n(x) \to f(x) \ a.s.$$

By some analogous argument to that of Corollary 1, the following conclusion can also be obtained.

Corollary 2. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and $k_n/n \to 0$, $k_n^2/(n \log n) \to \infty$. Then, for all $x \in c(f)$,*

$$\sum_{n=1}^{\infty} P(|f_n(x) - f(x)| > \varepsilon) < \infty$$

for all $\varepsilon > 0$, and hence

$$f_n(x) \to f(x) \ a.s.$$

Moreover, we can further obtain the rate of complete consistency for the nearest neighbor density estimator as follows.

Theorem 3. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and $f(x)$ satisfy the local Lipschitz condition at x and $f(x) > 0$. If $k_n = O(n^{3/4} \log^{1/4} n)$ and $\tau_n =: \sqrt{n \log n}/k_n \to 0$; then, for all sufficiently large $D > 0$,*

$$\sum_{n=1}^{\infty} P(|f_n(x) - f(x)| > D\tau_n) < \infty,$$

and hence

$$|f_n(x) - f(x)| \leq D\tau_n \ a.s.$$

By choosing $k_n = \lfloor n^{3/4} \log^{1/4} n \rfloor$ in Theorem 3, the following result follows immediately.

Corollary 3. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples, and let $f(x)$ satisfy the local Lipschitz condition at x and $f(x) > 0$. If $k_n = \lfloor n^{3/4} \log^{1/4} n \rfloor$, then for all sufficiently large $D > 0$,*

$$\sum_{n=1}^{\infty} P\Big(|f_n(x) - f(x)| > Dn^{-1/4} \log^{1/4} n\Big) < \infty,$$

and hence

$$|f_n(x) - f(x)| \leq Dn^{-1/4} \log^{1/4} n \ a.s.$$

At last, we also obtain some achievements concerning uniform consistency and the corresponding convergence rate for the estimator as follows.

Theorem 4. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and $f(x)$ be uniformly continuous. If $k_n/n \to 0$, $k_n^2/(n \log n) \to \infty$, then for all $\varepsilon > 0$,*

$$\sum_{n=1}^{\infty} P\left(\sup_x |f_n(x) - f(x)| > \varepsilon\right) < \infty,$$

and hence

$$\sup_x |f_n(x) - f(x)| \to 0 \ a.s.$$

Theorem 5. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples and let $f(x)$ satisfy the Lipschitz condition on $\mathbb{R}$. If $k_n = O(n^{2/3} \log^{1/3} n)$ and $\tau_n =: \sqrt{n \log n}/k_n \to 0$; then, for any sufficiently large $D > 0$,*

$$\sum_{n=1}^{\infty} P\left(\sup_x |f_n(x) - f(x)| > D\tau_n\right) < \infty,$$

and hence

$$\sup_x |f_n(x) - f(x)| \leq D\tau_n \ a.s.$$

By choosing $k_n = \lfloor n^{2/3} \log^{1/3} n \rfloor$ in Theorem 5, one can further obtain the corollary as follows.

Corollary 4. *Let $\{X_n, n \geq 1\}$ be a sequence of m-AANA samples, and let $f(x)$ satisfy the Lipschitz condition on $\mathbb{R}$. If $k_n = \lfloor n^{2/3} \log^{1/3} n \rfloor$, then for any sufficiently large $D > 0$,*

$$\sum_{n=1}^{\infty} P\left(\sup_x |f_n(x) - f(x)| > Dn^{-1/6} \log^{1/6} n\right) < \infty,$$

and hence

$$\sup_x |f_n(x) - f(x)| \leq Dn^{-1/6} \log^{1/6} n \ a.s.$$

Remark 2. *Yang [5], as well as Wang and Hu [6], obtained the rates $o(n^{-1/4} \log^{1/4} n \log \log n)$ a.s. of strong consistency and $o(n^{-1/6} \log^{1/6} n \log \log n)$ a.s. of uniformly strong consistency for NA samples and WOD samples, respectively. Wu and Wang [15] extended their results to AANA samples with the same rates presented in Theorems 3 and 5. Noting that the rates are sharper than those of Yang [5] and Wang and Hu [6], and AANA implies m-AANA, our results extend or improve the corresponding ones in Yang [5], Wang and Hu [6], and Wu and Wang [15].*

4. Numerical Simulation

In this section, some simple numerical simulations are carried out to verify the performance of $f_n(x)$ with a finite sample. First, we generate the AANA and m-dependent data, both of which are special cases of m-AANA, according to the following two cases, respectively.

Case 1. *Let $\{Y_n, n \geq 1\}$ be independent and identically distributed with a standard normal variable, and let $X_n = (1 + a_n^2)^{-1/2}(Y_n + a_n Y_{n+1})$ for each $n \geq 1$, where $a_n > 0$ and $a_n \to 0$. It is easy to check that $X_1, X_2, \cdots, X_n$ are AANA random variables with $X_i \sim N(0,1)$ for each $i = 1, 2, \cdots, n$.*

Case 2. *For $m \geq 2$, let $Y_n, n \geq 1$ be independent and identically distributed with a common $\chi^2_{(1)}$ variable. Let $X_n = \sum_{i=1}^{m} Y_{n+i-1}$ for each $n \geq 1$. Obviously, $X_1, X_2, \cdots, X_n$ are m-dependent and thus m-AANA random variables with $X_n \sim \chi^2_{(m)}$.*

Case 3. *For $m \geq 2$, let $\{Y_n, n \geq 1\}$ be independent and identically distributed $N(0,1)$ random variables and define $Z_n = (1 + a_n^2)^{-1/2}(Y_n + a_n Y_{n+1})$, where $a_n > 0$ and $a_n \to 0$. Now, let $X_{m(n-1)+1} = \cdots = X_{mn} = Z_n$ for each $n \geq 1$. From Example 1, one knows that $\{X_n, n \geq 1\}$ is m-AANA rather than AANA.*

In this section, we will compare the frequency polygon estimator, Epanechnikov kernel estimator (that is, the kernel $K(u) = 0.75(1 - u^2)I(|u| \leq 1)$), and histogram estimation with the nearest neighbor estimator. In the sequel, we take $m = 3$, $k_n = n^{3/4}(\log n)^{1/4}$ for the nearest neighbor estimator, the bin-width $b_n = (\log(n)/n)^{0.25}$ for the frequency polygon estimator and the histogram estimator, and the bandwidth by cross validation (CV, in short) method for the Epanechnikov kernel estimator. It is deserved to mention that k_n and b_n are chosen to achieve the optimal convergence rates. According to the above three cases, we take $n = 100, 200, 500, 1000$ and different x-values such as the peak and tail, respectively. For different x and n, we adopt the R software to calculate the four estimators for 1000 times to obtain the the absolute bias (ABias, in short) and the root mean squared error (RMSE, in short) of the four estimators. The conclusions obtained are exhibited in Tables 1–3 and Figures 1–3.

Table 1. Absolute bias and RMSE of the estimators for different x and n under Case 1.

	Estimators	$n = 100$		$n = 200$		$n = 500$		$n = 1000$	
		ABias	**RMSE**	**ABias**	**RMSE**	**ABias**	**RMSE**	**ABias**	**RMSE**
$x = -3$	nearest neighbor	0.07513	0.07521	0.06881	0.06884	0.06062	0.06064	0.05455	0.05456
	frequency	0.00996	0.05338	0.00073	0.00681	0.00049	0.00421	0.00021	0.00345
	kernel	0.00102	0.00827	0.00062	0.00597	0.00055	0.00435	0.00023	0.00306
	histogram	0.00048	0.01018	0.00026	0.00720	0.00170	0.00436	0.00028	0.00377
$x = -2$	nearest neighbor	0.06779	0.06823	0.06132	0.06162	0.05254	0.05268	0.04602	0.04612
	frequency	0.00362	0.02606	0.00361	0.01935	0.00326	0.01296	0.00232	0.01059
	kernel	0.00303	0.02625	0.00232	0.02022	0.00177	0.01357	0.00160	0.01109
	histogram	0.07034	0.03137	0.01985	0.02798	−0.01545	0.02166	0.01523	0.01837
$x = -1$	nearest neighbor	0.00113	0.02717	0.00081	0.02108	0.00053	0.01543	0.00053	0.01263
	frequency	0.00252	0.04723	0.00081	0.04873	0.00067	0.02424	0.00032	0.02708
	kernel	0.00119	0.05112	0.00238	0.03798	0.00161	0.02682	0.00136	0.02187
	histogram	0.03560	0.07545	0.00353	0.053270	0.04199	0.05295	0.00349	0.02816
$x = 0$	nearest neighbor	0.02042	0.05259	0.01371	0.04031	0.00963	0.02860	0.00854	0.02147
	frequency	0.01325	0.05293	0.01086	0.04271	0.00658	0.03040	0.00584	0.02284
	kernel	0.00741	0.06047	0.00504	0.04741	0.00359	0.03413	0.00336	0.02526
	histogram	0.01467	0.08489	0.01040	0.06492	0.00689	0.04507	0.00633	0.03474
$x = 1$	nearest neighbor	0.00106	0.02738	0.00042	0.02209	0.00015	0.01542	0.00011	0.01206
	frequency	0.00055	0.04615	0.00045	0.04985	0.00040	0.02470	0.00041	0.02743
	kernel	0.00147	0.05031	0.00066	0.03776	0.00044	0.02680	0.00035	0.02131
	histogram	0.07177	0.10530	0.08878	0.10489	0.03915	0.05692	0.06573	0.07274
$x = 2$	nearest neighbor	0.06767	0.06812	0.06132	0.06158	0.05256	0.05270	0.04601	0.04610
	frequency	0.00413	0.02620	0.00444	0.01990	0.00340	0.01307	0.00181	0.01024
	kernel	0.00377	0.02602	0.00269	0.02079	0.00214	0.01401	0.00105	0.010646
	histogram	0.05344	0.07064	0.02396	0.03920	0.02081	0.02901	0.01517	0.02162
$x = 3$	nearest neighbor	0.07521	0.07528	0.06886	0.06891	0.06056	0.06058	0.05463	0.05464
	frequency	0.00031	0.00954	0.00037	0.00685	0.00073	0.00400	0.00010	0.00338
	kernel	0.00055	0.00775	0.00052	0.00582	0.00050	0.00418	0.00018	0.00306
	histogram	0.01169	0.02231	0.00880	0.01486	0.00289	0.00709	0.00482	0.00742

In view of Tables 1–3 and Figures 1–3, we can see the same conclusion under the three cases. Firstly, as the sample size increases, the error of all estimators decreases. The nearest neighbour estimator performs a little better than the kernel estimator and histogram estimation at most points, while at the points distributed on the tail, the nearest neighbour estimator performs worse than the later ones. In summary, the nearest neighbour estimator performs better than others near the peak but worse near the tail. These results show that the estimator considered in this paper also has some superiority to other classical estimators under dependent settings.

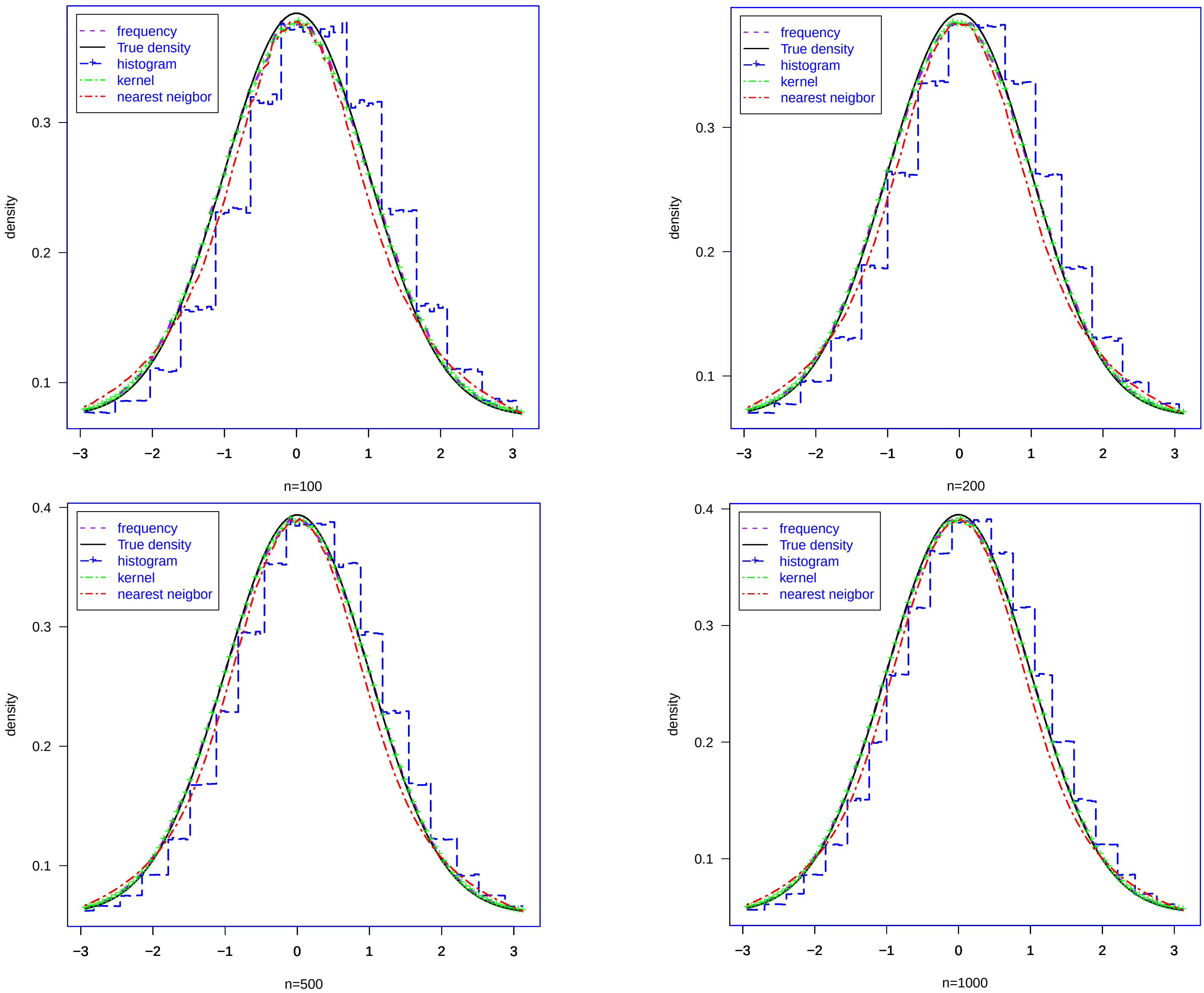

Figure 1. Comparison of different estimators for $n = 100, 200, 500, 1000$ under case 1.

Table 2. Absolute bias and RMSE of the estimators for different x and n under Case 2.

	Estimators	$n = 100$		$n = 200$		$n = 500$		$n = 1000$	
		ABias	**RMSE**	**ABias**	**RMSE**	**ABias**	**RMSE**	**ABias**	**RMSE**
$x = 0.5$	nearest neighbor	0.07937	0.08512	0.07047	0.07557	0.06200	0.06529	0.05413	0.05700
	frequency	0.05327	0.06563	0.04324	0.05447	0.03510	0.04427	0.02271	0.02801
	kernel	0.05503	0.06888	0.04300	0.05439	0.02891	0.03614	0.02245	0.02760
	histogram	0.08357	0.09896	0.07807	0.09017	0.07992	0.08971	0.02819	0.03487
$x = 1.5$	nearest neighbor	0.03234	0.04047	0.02504	0.03128	0.01759	0.02166	0.01335	0.01664
	frequency	0.04927	0.06188	0.04007	0.05044	0.03334	0.04124	0.01986	0.02491
	kernel	0.04893	0.06156	0.03944	0.04890	0.02663	0.03332	0.02040	0.02552
	histogram	0.06777	0.08469	0.04864	0.06130	0.03455	0.04433	0.02638	0.03333
$x = 3.5$	nearest neighbor	0.01586	0.01996	0.01234	0.01603	0.00910	0.01165	0.00705	0.00898
	frequency	0.04376	0.05452	0.03050	0.03800	0.02394	0.02996	0.01349	0.01714
	kernel	0.03614	0.04503	0.02806	0.03482	0.01943	0.02416	0.01475	0.01833
	histogram	0.04599	0.05764	0.03562	0.04456	0.02888	0.03656	0.02029	0.02558
$x = 5.5$	nearest neighbor	0.01592	0.01860	0.01238	0.01427	0.00930	0.01061	0.00733	0.00832
	frequency	0.02479	0.03077	0.02090	0.02603	0.01604	0.020318	0.00909	0.01148
	kernel	0.02559	0.03209	0.01918	0.02377	0.01320	0.01662	0.00985	0.01238
	histogram	0.03294	0.04127	0.02325	0.02916	0.01824	0.02381	0.01302	0.01638
$x = 7.5$	nearest neighbor	0.02162	0.02211	0.01833	0.01864	0.01465	0.01483	0.01221	0.01235
	frequency	0.01690	0.02115	0.01460	0.01878	0.01023	0.01286	0.00621	0.00784
	kernel	0.01717	0.02170	0.01262	0.01621	0.00861	0.01073	0.00669	0.00840
	histogram	0.02260	0.02991	0.01564	0.02034	0.01209	0.01539	0.00849	0.01072
$x = 9.5$	nearest neighbor	0.02283	0.02293	0.02012	0.02019	0.01680	0.01684	0.01441	0.01444
	frequency	0.01327	0.01601	0.00923	0.01197	0.00666	0.00838	0.00386	0.00495
	kernel	0.01026	0.01298	0.00823	0.01044	0.00562	0.00708	0.00416	0.00534
	histogram	0.01335	0.01611	0.00962	0.01247	0.00720	0.00915	0.00518	0.00666

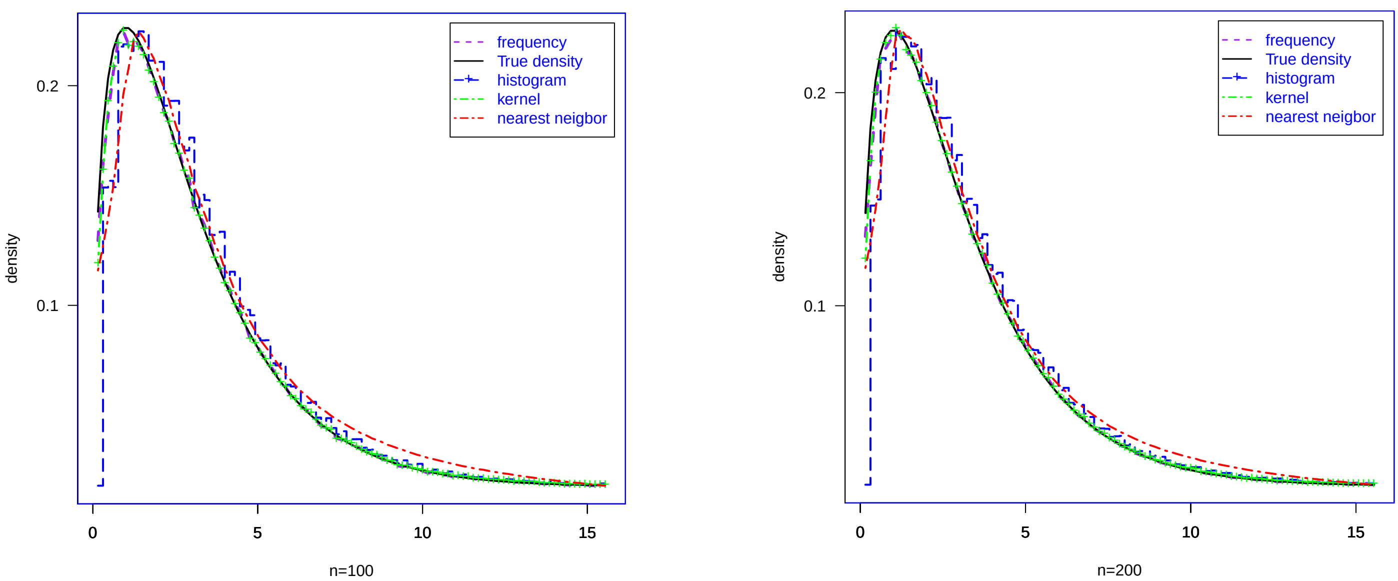

Figure 2. *Cont.*

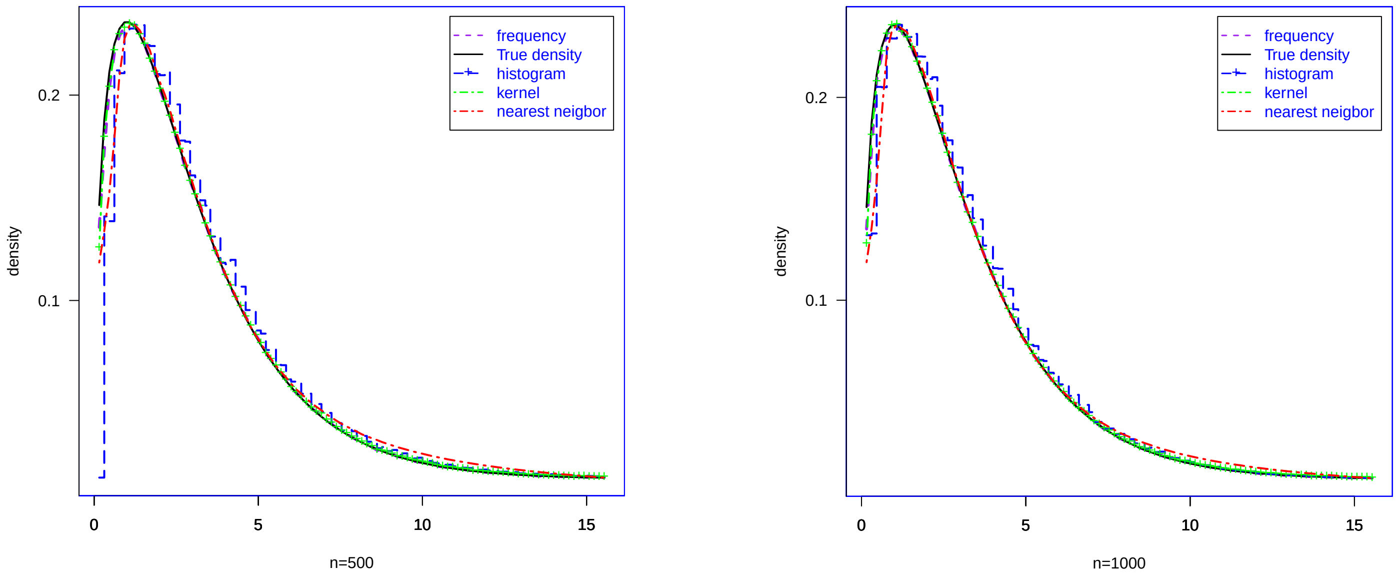

Figure 2. Comparison of different estimators for $n = 100, 200, 500, 1000$ under case 2.

Table 3. Absolute bias and RMSE of the estimators for different x and n under Case 3.

	Estimators	$n = 100$		$n = 200$		$n = 500$		$n = 1000$	
		ABias	RMSE	ABias	RMSE	ABias	RMSE	ABias	RMSE
$x = -3$	nearest neighbor	0.07610	0.07633	0.06950	0.06971	0.06052	0.06056	0.05461	0.05464
	frequency	0.00930	0.01622	0.00652	0.01128	0.00645	0.00749	0.00496	0.00588
	kernel	0.00832	0.01284	0.00710	0.01021	0.00684	0.00807	0.00444	0.00536
	histogram	0.00815	0.01663	0.00757	0.01188	0.00703	0.00814	0.00563	0.00635
$x = -2$	nearest neighbor	0.06923	0.07070	0.06118	0.06196	0.05715	0.05762	0.04669	0.04682
	frequency	0.03861	0.04881	0.02607	0.03371	0.01926	0.02511	0.01853	0.02361
	kernel	0.03808	0.04823	0.02820	0.03432	0.02276	0.03059	0.01802	0.02188
	histogram	0.04937	0.05914	0.03343	0.03839	0.02003	0.02545	0.01756	0.02350
$x = -1$	nearest neighbor	0.03631	0.04775	0.03392	0.04506	0.01997	0.02303	0.00912	0.00912
	frequency	0.06229	0.07671	0.07247	0.08903	0.02984	0.03644	0.03866	0.03866
	kernel	0.06957	0.08409	0.05274	0.06531	0.02978	0.03729	0.01678	0.01678
	histogram	0.09321	0.08409	0.07434	0.09125	0.05200	0.05696	0.03899	0.03899
$x = 0$	nearest neighbor	0.07075	0.08889	0.04916	0.06248	0.04165	0.04875	0.03089	0.03813
	frequency	0.07385	0.09206	0.05638	0.06542	0.05210	0.06610	0.03407	0.04322
	kernel	0.08064	0.10095	0.06331	0.07623	0.06388	0.08090	0.03431	0.04535
	histogram	0.11434	0.14591	0.08077	0.09885	0.06366	0.07957	0.04986	0.06241
$x = 1$	nearest neighbor	0.03835	0.05063	0.03149	0.04136	0.02109	0.02732	0.01697	0.02093
	frequency	0.06490	0.08186	0.07414	0.08815	0.03691	0.04265	0.03429	0.04344
	kernel	0.07212	0.09097	0.05512	0.06931	0.03643	0.04574	0.02777	0.03529
	histogram	0.11180	0.14016	0.10962	0.13703	0.06845	0.08603	0.06755	0.08356
$x = 2$	nearest neighbor	0.06930	0.07076	0.06486	0.06573	0.05061	0.05079	0.04127	0.04130
	frequency	0.03604	0.04665	0.02133	0.02786	0.01916	0.02315	0.01822	0.01687
	kernel	0.03724	0.04655	0.02060	0.02754	0.01864	0.02148	0.01502	0.01845
	histogram	0.08088	0.10331	0.04690	0.06046	0.02693	0.02977	0.02277	0.02503
$x = 3$	nearest neighbor	0.07593	0.07617	0.06940	0.06952	0.06054	0.06056	0.05479	0.05480
	frequency	0.00771	0.01406	0.00754	0.01303	0.00444	0.00444	0.00354	0.00390
	kernel	0.00831	0.01153	0.00781	0.01102	0.00511	0.00576	0.00350	0.00388
	histogram	0.02136	0.03534	0.01559	0.02350	0.00573	0.00655	0.01239	0.01605

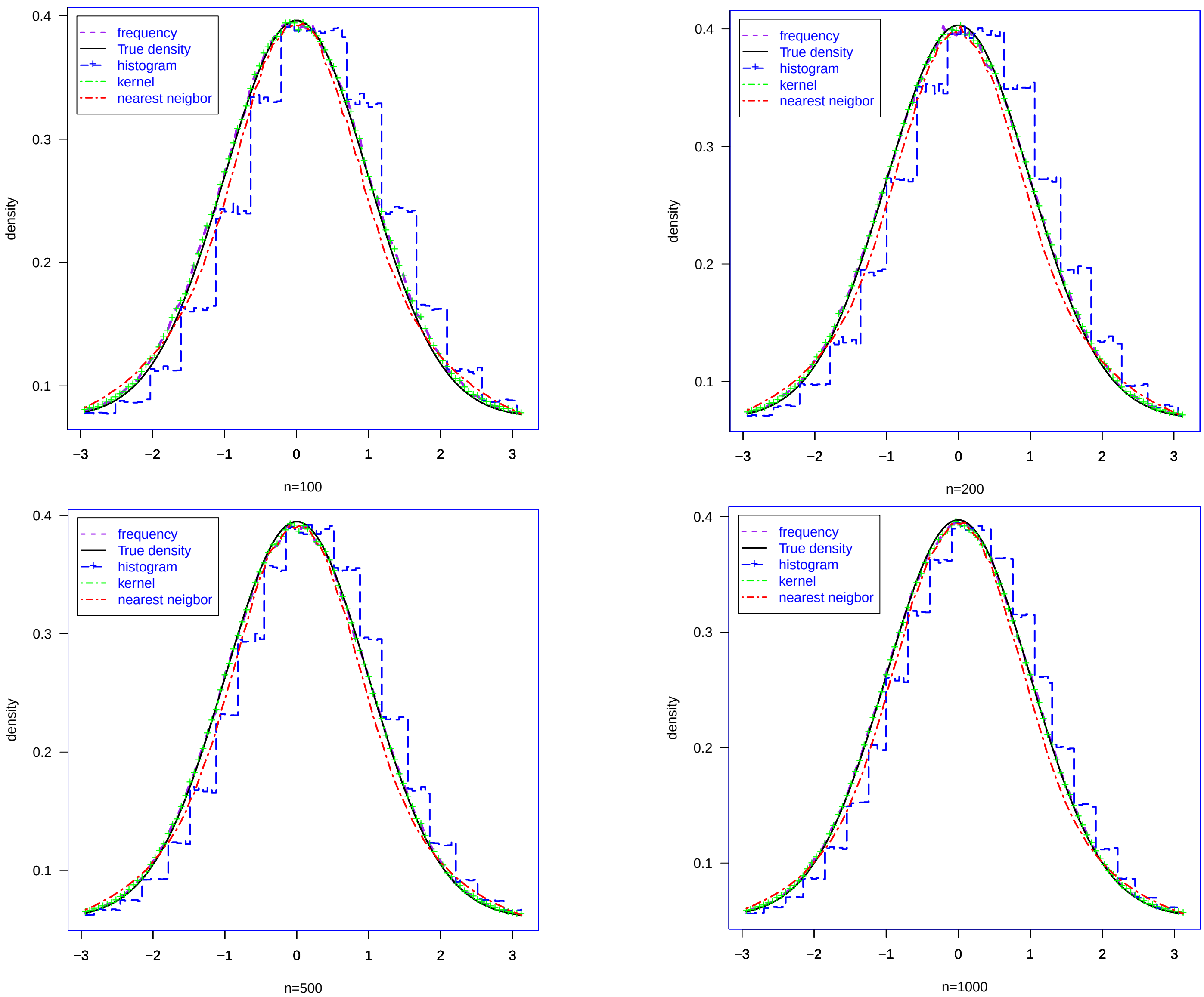

Figure 3. Comparison of different estimators for $n = 100, 200, 500, 1000$ under case 3.

5. Proof of the Main Results

The proofs are similar to those of Wu and Wang [15]. Therefore, we only present the differences in the sequel.

Proof of Theorem 1. Similar to the proof of Wu and Wang [15], we have

$$\{|f_n(x) - f(x)| > \varepsilon\} \subset A_{11x} \bigcup A_{12x} \bigcup A_{21x} \bigcup A_{22x}, \tag{7}$$

where

$$A_{11x} = \left\{|F_n(x + b_n(x)) - F(x + b_n(x))| \geq \frac{k_n}{n}\delta(x)\right\},$$

$$A_{12x} = \left\{|F_n(x - b_n(x)) - F(x - b_n(x))| \geq \frac{k_n}{n}\delta(x)\right\},$$

$$A_{21x} = \left\{ |F_n(x + c_n(x)) - F(x + c_n(x))| \geq \frac{k_n}{n}\delta(x) \right\},$$

and

$$A_{22x} = \left\{ |F_n(x - c_n(x)) - F(x - c_n(x))| \geq \frac{k_n}{n}\delta(x) \right\}$$

with $\delta(x) = \frac{\varepsilon}{8(f(x)+\varepsilon)}$.

For given x, define for each $1 \leq i \leq n,\ n \geq 1$ that

$$\xi_{ni} = I(X_i < x + b_n(x)) - EI(X_i < x + b_n(x)).$$

From Lemma 2, it is easy to see that $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nn}$ are still m-AANA random variables with $E\xi_{ni} = 0$ and $|\xi_{ni}| \leq 1$. Observe that $k_n \leq n$ and $\delta(x) \leq \frac{1}{8}$. Using Lemma 4, we have that

$$\begin{aligned} P(A_{11x}) &= P\left(|F_n(x + b_n(x)) - F(x + b_n(x))| \geq \frac{k_n}{n}\delta(x) \right) \\ &= P\left(\left| \sum_{k=1}^{n} \xi_{ni} \right| > k_n\delta(x) \right) \\ &\leq C\chi_n \cdot \exp\left\{ -\frac{k_n^2\delta^2(x)/m^2}{2B_n^2 + \frac{2}{3}k_n\delta(x)/m} \right\} \\ &\leq C\chi_n \cdot \exp\left\{ -\frac{k_n^2\delta^2(x)/m^2}{2n + \frac{1}{12}n/m} \right\} \\ &= C\chi_n \cdot \exp\left\{ -\frac{12\delta^2(x)/m}{24m+1}\frac{k_n^2}{n} \right\}. \end{aligned} \tag{8}$$

Analogously, we can also obtain the same upper bounds as in (8) for the probability of events A_{12x}, A_{21x}, and A_{22x}, respectively. Therefore, we further obtain by (5) and (7) that

$$\begin{aligned} P(|f_n(x) - f(x)| > \varepsilon) &\leq P(A_{11x}) + P(A_{12x}) + P(A_{21x}) + P(A_{22x}) \\ &\leq 4C\chi_n \cdot \exp\left\{ -\frac{12\delta^2(x)/m}{24m+1}\frac{k_n^2}{n} \right\} \to 0. \end{aligned}$$

The proof is finished. □

Proof of Corollary 1. In view of Theorem 1, we only need to verify that (5) holds. By $k_n^2/(n\log n) \to \infty$, one can obtain that

$$\exp\left\{ -\frac{\gamma k_n^2}{n} \right\} \leq \exp\{-3\log n\} = n^{-3} \tag{9}$$

for any $\gamma > 0$ and any sufficiently large n. Moreover, noticing that $q(n) \to 0$, $n_0 > 0$ exists such that $q(n) \leq 1$ for all $n > n_0$, and thus

$$\chi_n = \sum_{k=1}^{n-1} q(k) + 1 = O(n).$$

Therefore, we have by (9) that

$$\chi_n \exp\left\{ -\frac{\gamma k_n^2}{n} \right\} \leq Cn^{-2} \to 0, \tag{10}$$

which finishes the proof. □

Proof of Theorem 2. The proof is analogous to that of Theorem 1. In view of (6), one has that

$$\sum_{n=1}^{\infty} P(|f_n(x) - f(x)| > \varepsilon) \leq 4C \sum_{n=1}^{\infty} \chi_n \cdot \exp\left\{-\frac{12\delta^2(x)/m}{24m+1} \frac{k_n^2}{n}\right\} < \infty.$$

Hence, the desired result follows from the Borel–Cantelli lemma and the formula above immediately. □

Proof of Corollary 2. Similar to the proof of Corollary 1, we have by (10) that

$$\sum_{n=1}^{\infty} \left[\sum_{k=1}^{n-1} q(k) + 1\right] \cdot \exp\left\{-\frac{\gamma k_n^2}{n}\right\} \leq C \sum_{n=1}^{\infty} n^{-2} < \infty.$$

The proof is thus finished. □

Proof of Theorem 3. Analogous to the proof of Theorem 2.6 in Wu and Wang [15], we also have that

$$\{|f_n(x) - f(x)| > D\tau_n\} \subset B_{11x} \bigcup B_{12x} \bigcup B_{21x} \bigcup B_{22x}, \tag{11}$$

where

$$B_{11x} = \left\{|F_n(x+\mu_n(x)) - F(x+\mu_n(x))| \geq \frac{k_n \tau_n}{n} \cdot \frac{D}{8T}\right\},$$

$$B_{12x} = \left\{|F_n(x-\mu_n(x)) - F(x-\mu_n(x))| \geq \frac{k_n \tau_n}{n} \cdot \frac{D}{8T}\right\},$$

$$B_{21x} = \left\{|F_n(x+\nu_n(x)) - F(x+\nu_n(x))| \geq \frac{k_n \tau_n}{n} \cdot \frac{D}{8T}\right\},$$

and

$$B_{22x} = \left\{|F_n(x-\nu_n(x)) - F(x-\nu_n(x))| \geq \frac{k_n \tau_n}{n} \cdot \frac{D}{8T}\right\}$$

with $T =: \sup_x f(x) < \infty$, $D > \frac{c_1^2 L(x)}{f(x)}$ and $L(x) > 0$ depending only on x.

For each given x and $1 \leq i \leq n,\ n \geq 1$, we define

$$\eta_{ni} = I(X_i < x + \mu_n(x)) - EI(X_i < x + \mu_n(x)).$$

From Lemma 2, it is easy to see that $\eta_{n1}, \eta_{n2}, \ldots, \eta_{nn}$ are still m-AANA random variables with $E\eta_{ni} = 0$ and $|\eta_{ni}| \leq 1$. Applying Lemma 4 and noticing that $k_n \leq n$, $\tau_n \to 0$, we obtain that for all sufficiently large n,

$$
\begin{aligned}
P(B_{11x}) &= P\left(|F_n(x+\mu_n(x)) - F(x+\mu_n(x))| \geq \frac{k_n\tau_n}{n}\cdot\frac{D}{8T}\right) \\
&= P\left(\left|\sum_{i=1}^{n}\eta_{ni}\right| > k_n\tau_n\cdot\frac{D}{8T}\right) \\
&\leq C\chi_n\cdot\exp\left\{-\frac{k_n^2\tau_n^2D^2/(64T^2m^2)}{2B_n^2+\frac{D}{12Tm}k_n\tau_n}\right\} \\
&\leq Cn\exp\left\{-\frac{k_n^2\tau_n^2}{n}\cdot\frac{D^2}{128m^2T^2+\frac{16}{3}DmT}\right\} \\
&= Cn\exp\left\{-\frac{D^2}{128m^2T^2+\frac{16}{3}DmT}\log n\right\} \\
&\leq Cn^{1-\frac{D^2}{128m^2T^2+\frac{16}{3}DmT}}.
\end{aligned} \tag{12}
$$

Analogously, the probabilities of B_{12x}, B_{21x}, and B_{22x} also have the same upper bounds as in (12). Therefore, taking $D > \frac{c_0^2L(x)}{f(x)}$ such that $1 - \frac{D^2}{128m^2T^2+\frac{16}{3}DmT} < -1$, one can obtain by (11) that

$$
\begin{aligned}
\sum_{n=1}^{\infty}P(|f_n(x)-f(x)|>D\tau_n) &\leq \sum_{n=1}^{\infty}(P(B_{11x})+P(B_{12x})+P(B_{21x})+P(B_{22x})) \\
&\leq 4C\sum_{n=1}^{\infty}n^{1-\frac{D^2}{128m^2T^2+\frac{16}{3}DmT}} < \infty.
\end{aligned}
$$

This completes the proof of the theorem. □

Proof of Theorem 4. It follows from the proof of Theorem 2.9 in Wu and Wang [15] that

$$
\left(\sup_x|f_n(x)-f(x)|>\varepsilon\right) \subset \left(\sup_x|F_n(x)-F(x)| \geq \frac{\varepsilon}{8(T+\varepsilon)}\frac{k_n}{n}\right), \tag{13}
$$

where $T = \sup_x f(x) < \infty$.

On the other hand, by $k_n^2/(n\log n) \to \infty$ we have that for all sufficiently large n, $\frac{\varepsilon}{8(T+\varepsilon)}\frac{k_n}{n} \geq D_0(\log n/n)^{1/2}$. Hence, taking $\kappa_n = (\log n/n)^{1/2}$ in Lemma 6, one has by (13) that

$$
\begin{aligned}
\sum_{n=1}^{\infty}P\left(\sup_x|f_n(x)-f(x)|>\varepsilon\right) &\leq \sum_{n=1}^{\infty}P\left(\sup_x|F_n(x)-F(x)| \geq \frac{\varepsilon}{8(T+\varepsilon)}\frac{k_n}{n}\right) \\
&\leq \sum_{n=1}^{\infty}P\left(\sup_x|F_n(x)-F(x)| \geq D_0(\log n/n)^{1/2}\right) < \infty.
\end{aligned}
$$

The proof is hence finished. □

Proof of Theorem 5. It follows from the proof of Theorem 2.10 in Wu and Wang [15] that

$$
\left(\sup_x|f_n(x)-f(x)|>D\tau_n\right) \subset \left(\sup_x|F_n(x)-F(x)| \geq \frac{k_n\tau_n}{n}\cdot\frac{D}{8T}\right), \tag{14}
$$

where $D > \max\{\sqrt{4c_2^3L}, 8TD_0\}$, $T = \sup_x f(x) < \infty$, and $L > 0$ is independent of x.

Consequently, on can apply Lemma 6 with $\kappa_n = \frac{k_n\tau_n}{n} = (\log n/n)^{1/2}$ to obtain that

$$\begin{aligned}\sum_{n=1}^{\infty} P\left(\sup_x |f_n(x) - f(x)| > D\tau_n\right) &\leq \sum_{n=1}^{\infty} P\left(\sup_x |F_n(x) - F(x)| \geq \frac{k_n\tau_n}{n} \cdot \frac{D}{8T}\right) \\ &\leq \sum_{n=1}^{\infty} P\left(\sup_x |F_n(x) - F(x)| \geq D_0(\log n/n)^{1/2}\right) < \infty.\end{aligned}$$

This completes the proof of the theorem. □

6. Conclusions

In this paper, a Bernstein inequality for m-asymptotically almost negatively associated random variables is established based on that of asymptotically almost negatively associated random variables. By virtue of this inequality, some results on consistency for the nearest neighbor estimator of the density function are further obtained. The results are further extensions of existing ones in the literature. From the simulation study, we find that the nearest neighbour estimator performs better than others on the peak but worse on the tail, which encourages us to consider whether we can combine the superiorities of these estimators to construct a better method.

Author Contributions: Validation, W.W.; data curation, W.W.; writing—original draft, X.L.; writing—review & editing, Y.W.; supervision, Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Provincial Natural Science Research Project of Anhui Colleges, grant number KJ2018A0579.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are most grateful to the editor and anonymous referees for carefully reading the manuscript and for valuable suggestions that helped in improving an earlier version of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Loftsgarden, D.O.; Quesenberry, C.P. A nonparametric estimate of a multivariate density function. *Ann. Math. Stat.* **1965**, *36*, 1049–1051. [CrossRef]
2. Liu, Y.H.; Wu, Q.Y. Consistency of nearest neighbor estimator of density function for negatively dependent samples. *J. Jilin Univ.* **2012**, *50*, 1142–1145.
3. Lu, Z.L.; Ding, S.N.; Zhang, F.; Wang, R.; Wang, X.J. The consistency and convergence rate for the nearest neighbor density estimator based on φ-mixing random samples. *Commun. Stat.-Theory Methods* **2022**, *51*, 669–684. [CrossRef]
4. Liu, Y.; Zhang, Y. The consistency and asymptotic normality of nearest neighbor density estimator under φ-mixing condition. *Acta Math. Sin.* **2010**, *30*, 733–738.
5. Yang, S.C. Consistency of nearest neighbor estimator of density function for negative associated samples. *Acta Math. Appl. Sin.* **2003**, *26*, 385–395.
6. Wang, X.J.; Hu, H.S. The consistency of the nearest neighbor estimator of the density function based on WOD samples. *J. Math. Anal. Appl.* **2015**, *429*, 497–512. [CrossRef]
7. Lan, C.F.; Wu, Q.Y. Uniform Strong Consistency Rate of Nearest Neighbor Estimator of Density Function for END Samples. *J. Jilin Univ.* **2014**, *52*, 495–498.
8. Wang, W.; Wu, Y. Consistency of nearest neighbor estimator of density function for m-END samples. *Braz. J. Probab. Stat.* **2022**, *36*, 369–384. [CrossRef]
9. Chandra, T.K.; Ghosal, S. Extensions of the strong law of large numbers of Marcinkiewicz and Zygmund for dependent variables. *Acta Math. Hung.* **1996**, *71*, 327–336. [CrossRef]
10. Kim, T.; Ko, M.; Lee, I. On the strong law for asymptotically almost negatively associated random variables. *Rocky Mt. J. Math.* **2004**, *34*, 979–989. [CrossRef]
11. Yuan, D.M.; An, J. Rosenthal type inequalities for asymptotically almost negatively associated random variables and applications. *Sci. China Ser. A Math.* **2009**, *52*, 1887–1904. [CrossRef]

12. Chandra, T.K.; Ghosal, S. The strong law of large numbers for weighted averages under dependence assumptions. *J. Theor. Probab.* **1996**, *9*, 797–809. [CrossRef]
13. Shen, A.T.; Wu, R.C. Strong convergence for sequences of asymptotically almost negatively associated random variables. *Stochastics-Int. J. Probab. Stoch. Process.* **2014**, *86*, 291–303. [CrossRef]
14. Yuan, D.M.; An, J. Laws of large numbers for Cesàro alpha-integrable random variables under dependence condition AANA or AQSI. *Acta Math. Sin.* **2012**, *28*, 1103–1118. [CrossRef]
15. Wu, Y.; Wang, X.J. On Consistency of the Nearest Neighbor Estimator of the Density Function and Its Applications. *Acta Math. Sin.* **2019**, *35*, 703–720. [CrossRef]
16. Nam, T.H.; Thuy, N.T.; Hu, T.C.; Volodin, A. Maximal inequalities and strong law of large numbers for sequences of m-asymptotically almost negatively associated random variables. *Commun. Stat.-Theory Methods* **2016**, *46*, 2696–2707. [CrossRef]