Journal of Mathematical Inequalities
Volume 15, Number 3 (2021), 1287–1298
doi:10.7153/jmi-2021-15-86

ON THE UNIFORM CONSISTENCY OF FREQUENCY POLYGONS FOR ρ^-–MIXING SAMPLES

WEI WANG, HAIWU HUANG, YI WU AND KAN CHEN*

(*Communicated by X. Wang*)

Abstract. In this paper, the frequency polygon is considered as a nonparametric density estimator for ρ^--mixing samples. By the moment inequality, we prove the uniformly strong consistency of the estimator and obtain the corresponding rate under some mild conditions. The results obtained in this paper extend and improve some existing ones in the literature.

1. Introduction

At the outset, let us recall the notation of frequency polygon. Suppose that X is a random variable with a density function $f(x)$ and let $X_1, X_2, \cdots, X_n$ be the sample drawn from the population X. Consider a partition $\cdots x_{-2} < x_{-1} < x_0 < x_1 < x_2 \cdots$ of the real line into equal intervals $I_k = [(k-1)b_n, kb_n)$ of the length b_n, where b_n is the bin width. For a given $x \in \mathbb{R}$, there exists k_0 such that $(k_0 - \frac{1}{2})b_n \leqslant x < (k_0 + \frac{1}{2})b_n$. Consider two adjacent histogram bins $I_{k_0} = [(k_0-1)b_n, k_0 b_n)$ and $I_{k_1} = [k_0 b_n, k_1 b_n)$, where $k_1 = k_0 + 1$. Define $v_{k_0} = \sum_{i=1}^{n} I((k_0-1)b_n \leqslant X_i < k_0 b_n)$ and $v_{k_1} = \sum_{i=1}^{n} I(k_0 b_n \leqslant X_i < k_1 b_n)$, which are the numbers of the observations falling into the intervals mentioned above, respectively. The values of the histogram in these previous bins can be denoted by $f_{k_0} = v_{k_0} n^{-1} b_n^{-1}$ and $f_{k_1} = v_{k_1} n^{-1} b_n^{-1}$. Then, the frequency polygon $\hat{f}(x)$ can be defined as

$$\hat{f}(x) = \left(\frac{1}{2} + k_0 - \frac{x}{b_n}\right) f_{k_0} + \left(\frac{1}{2} - k_0 + \frac{x}{b_n}\right) f_{k_1} \tag{1.1}$$

for $x \in [(k_0 - \frac{1}{2})b_n, (k_0 + \frac{1}{2})b_n)$. As pointed out in Scott (1985), the frequency polygon estimator $\hat{f}(x)$ has convergence rate similar to those of kernel density estimators and greater than the rate of the histogram. As for computation, the computational effort of the frequency polygons is equivalent to the one of the histogram. For large bivariate data sets, the computational simplicity of the frequency polygons and the ease of

Mathematics subject classification (2020): 62G05.
Keywords and phrases: Uniformly strong consistency, ρ^--mixing samples, Frequency polygon.
Supported by the Key Research Project of Chaohu University (XLZ-201903, XLZ-201904), the PhD Research Initiation Fund Project of Chizhou University (2020YJRC003).
* Corresponding author.

determining exact equiprobale contours may outweigh the increased accuracy of a kernel density estimator. Since the frequency polygons has the two advantages mentioned above, it is of value and interest to investigate it further.

Recently, scholars have obtained some results on frequency polygons. For example, Carbon et al. (1997) gave the optimal bin widths asymptotically minimizing integrated mean square errors, asymptotic variance, uniformly strong consistency and the convergence rate of the frequency polygon for α-mixing processes; Carbon et al. (2010) proved asymptotic normality of the frequency polygon for random fields; Bensaid and Dabo-Niang (2010) derived the integrated mean square errors and uniformly strong rate of consistency of the estimator in continuous random fields; Xing et al. $(2015a, 2015b)$ investigated the uniformly strong consistency of the frequency polygon estimator under negatively associated samples and ψ-mixing samples, respectively. Motivated by the literature above, we further investigate the uniformly strong consistency of the frequency polygon under ρ^--mixing samples, which has not been obtained in the literature before. Furthermore, the corresponding rate of convergence is also obtained.

In the following, we recall the concept of ρ^--mixing random variables. For two nonempty disjoint sets S,T of real numbers, we define $\text{dist}(S,T)=\min\{|j-k|; j\in S, k\in T\}$.

DEFINITION 1.1. A sequence $\{Y_i, -\infty < i < +\infty\}$ is said to be ρ^--mixing, if

$$\rho^-(s) = sup\{\rho^-(S,T); S,T \subset Z, dist(S,T) \geqslant s\} \to 0 \ \ as \ \ s \to \infty,$$

where $\rho^-(S,T)=0 \vee \sup\{\text{corr}(f(X_i, i\in S), g(X_j, j\in T))\}$, and the supremum is taken over all coordinatewise increasing real functions f on $\mathbb{R}^S$ and g on $\mathbb{R}^T$.

It is obvious that $\rho^-(s) \leqslant \rho^*(s)$. It is easy to see that $\{Y_i, -\infty < i < +\infty\}$ is negatively associated (Joag-Dev and Proschan, 1983) if and only if $\rho^-(s) = 0$ for $s \geqslant 1$. So ρ^--mixing is weaker than ρ^*-mixing and can be regarded as the asymptotically negative association or negative side ρ^*-mixing. Consequently, the study of the limit properties for ρ^--mixing variables is of much interest. Since the concept of ρ^--mixing variables was introduced by Zhang and Wang (1999), many applications have been found. For example, Zhang and Wang (1999) and Zhang (2000a, 2000b) obtained moment inequalities for partial sums, the central limit theorems, the complete convergence, and the strong law of large numbers; Wang and Lu (2006) established some inequalities for the maximum of partial sums and weak convergence; Zhang (2015) established the complete moment convergence for moving-average process generated by ρ^--mixing variables; Huang et al. (2016) proved the complete convergence and complete moment convergence for weighted sums of ρ^--mixing random variables; Wang et al. (2019) investigated the Berry-Esseen bounds of weighted estimator in a nonparametric regression model for ρ^--mixing samples, and so forth.

In this work, we further study the uniformly strong consistency of the frequency polygon density estimator and obtain the corresponding rate under some mild conditions. The results obtained in this paper extend and improve some existing ones in the literature. Moreover, some numerical analysis is also presented to support the theoretical results.

Throughout this paper, we always suppose that C denotes a positive constant which only depends on some given numbers and may vary from one place to another. The bin width and the density function are denoted by b_n and $f(x)$, respectively, and the limits are taken as $n\to\infty$ unless indicated otherwise. $\lfloor x\rfloor$ stands for the integer part of x. The rest of this paper is organized as follows. Main results and some simulation results on finite sample performance of frequency polygons are detailed in Sections 2 and 3, respectively. The proofs of the results are presented in Section 4.

2. Main results

To obtain the main results, we need the following assumptions:

(A_1) $\{X_i, 1\leqslant i\leqslant n\}$ is a ρ^--mixing sample with the common density function $f(x)$.

(A_2) The bin width b_n satifies $b_n\to 0$ and $nb_n\to\infty$.

(A_3) $\{\tau_n, n\geqslant 1\}$ is a sequence of positive constants satisfying $\lim\limits_{n\to\infty}\tau_n=0$ and

$$\liminf_{n\to\infty}(n^{\delta}b_n\tau_n)>0 \quad \text{for some} \quad 0<\delta<\frac{1}{2}.$$

Based on the assumptions above, our main results can be given as follows.

THEOREM 2.1. *Suppose Assumptions $(A_1)-(A_3)$ hold. Then for any compact subset D of $\mathbb{R}$,*

$$\sup_{x\in D}|\hat f(x)-E\hat f(x)|=o(\tau_n)\ a.s.. \tag{2.1}$$

Furthermore, if $f(x)$ is differentiable for $x\in\mathbb{R}$ and $|f'(x)|\leqslant M$ for some $M>0$, then,

$$\sup_{x\in D}|E\hat f(x)-f(x)|=O(b_n). \tag{2.2}$$

Thus,

$$\sup_{x\in D}|\hat f(x)-f(x)|=o(\tau_n)+O(b_n)\ a.s.. \tag{2.3}$$

THEOREM 2.2. *Suppose Assumptions $(A_1)-(A_3)$ hold. Then for any $T>0$,*

$$\sup_{x\in[-n^T,n^T]}|\hat f(x)-E\hat f(x)|=o(\tau_n)\ a.s.. \tag{2.4}$$

Moreover, if $f(x)$ is differentiable for $x\in\mathbb{R}$, $|f'(x)|\leqslant M$ for some $M>0$ and $E|X_1|^{2/T}<\infty$, then

$$\sup_{x\in\mathbb{R}}|\hat f(x)-f(x)|=o(\tau_n)+O(b_n)\ a.s.. \tag{2.5}$$

Taking $\tau_n=n^{-\delta}b_n^{-1}$ in Theorem 2.1 and Theorem 2.2, we can get immediately the following corollary.

COROLLARY 2.1. *Suppose Assumptions* (A_1), (A_2) *are satisfied,* $f(x)$ *is differentiable for* $x \in \mathbb{R}$ *and* $|f'(x)| \leqslant M$ *for some* $M>0$*. Then*

$$\sup_{x\in D}|E\hat{f}(x)-f(x)| = O(\Psi(n)) \text{ a.s.}, \tag{2.6}$$

where $\Psi(n) = \max\{b_n, n^{-\delta}b_n^{-1}\} \to 0$ *as* $n \to \infty$*. Further if* $E|X_1|^{2/T} < \infty$ *for some* $T>0$*, then*

$$\sup_{x\in\mathbb{R}}|\hat{f}(x)-f(x)| = O(\Psi(n)) \text{ a.s.}. \tag{2.7}$$

REMARK 2.1. Comparing Corollary 2.1 with the corresponding result of Carbon et al. (1997), we have the following improvements or extensions:

(i) The conditions $f(x) \leqslant M_1$ for some $M_1>0$ and $|x|^{1/(2T)}\ f(x) \to 0$ as $x \to 0$ are not needed here.

(ii) The assumption $\sup_{(x,y)\in\mathbb{R}^2} f_{j|i}(y|x) \leqslant M_2<\infty$ for all $i<j$ and some positive constant M_2 is not required here.

(iii) The moment condition is improved from $E|X_1|^{(2+\varepsilon)/T} < \infty$ for some $\varepsilon > 0$ and $T > 0$ to $E|X_1|^{2/T} < \infty$.

(iv) The structure of the sample is extended from α-mixing to ρ^--mixing.

REMARK 2.2. Xing et al. (2015a) obtained the corresponding results for negatively associated samples by using the exponential inequality of negatively associated random variables. However, we do not know whether the exponential inequality holds for ρ^--mixing random variables or not. So we adopt the Rosenthal-type moment inequality to obtain the results. The convergence rate is slightly slower but the conditions are weaker than those of Xing et al. (2015a). Moreover, noting that ρ^--mixing includes negative association, our results generalize the corresponding results of Xing et al. (2015a).

3. Numerical simulation study

In this section, we carry out a simulation study to examine the performance of frequency polygons with ρ^--mixing samples. The simulation will be conducted under the following two cases.

Case 1. Consider the MA(1) process:

$$X_n = \varepsilon_n - \theta\varepsilon_{n-1},$$

where $\{\varepsilon_n, n \geqslant 1\}$ are independent and identically distributed and $\varepsilon_n \sim N(0,\sigma_\varepsilon^2)$. It is obvious that $E(X_n) = 0$ and $Var(X_n) = (1+\theta^2)\sigma_\varepsilon^2$, so $X_n \sim N(0,(1+\theta^2)\sigma_\varepsilon^2)$. It is easy to see that $\{X_n, n \geqslant 1\}$ is both NA and ρ^*-mixing and thus ρ^--mixing.

For the purpose of comparison of the different estimators, we consider frequency polygon estimator, Epanechnikov kernel estimator (i.e., the $K(u) = 0.75(1-u^2)I(|u| \leqslant 1)$), and histogram estimator. Take the bin widths $b_n = (\log n/n)^{1/4}$, $\theta = 0.4$, $\sigma_\varepsilon =$

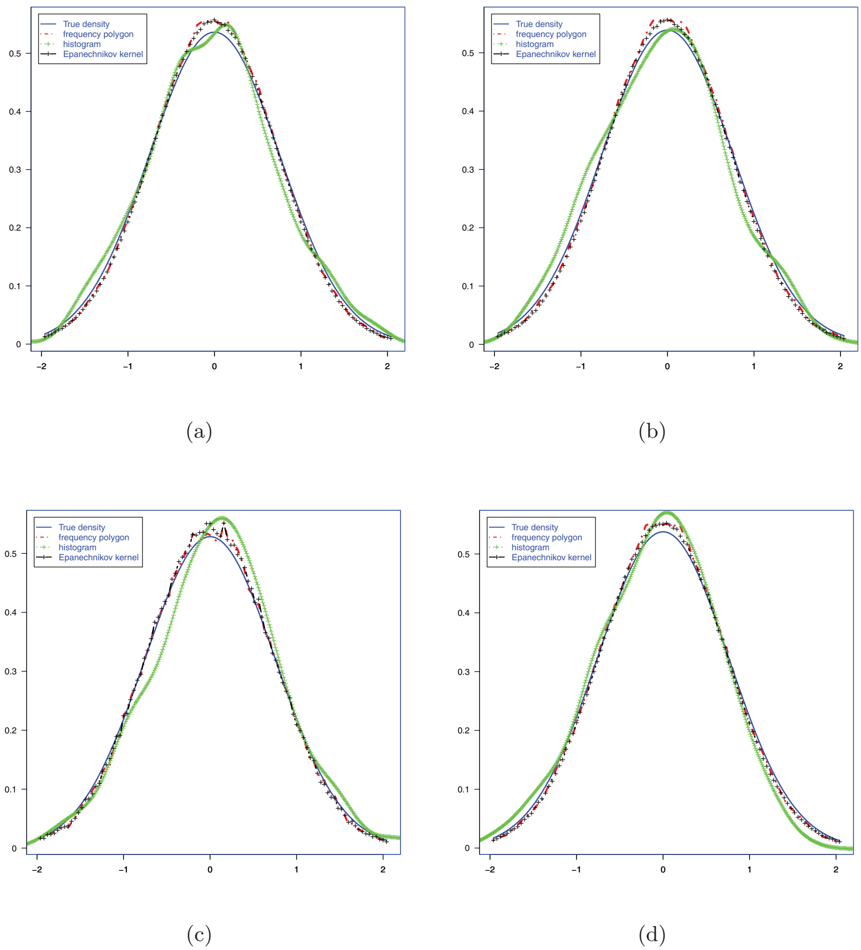

Figure 1: *Comparsion of different estimators for* $n = 100, 200, 300, 400$ *under Case 1*

0.7, and the sample sizes as $n = 100, 200, 300, 400$, respectively. We use R software to compute the estimators for 500 times to obtain the final values and then compare them with $f(x)$ in Figure 1 under different sample sizes. It reveals in Figure 1 that both the frequency polygons estimator and Epanechnikov kernel estimator perform better than the histogram estimator for each sample size, and there is no obvious difference between the frequency polygons estimator and Epanechnikov kernel estimator.

Case 2. ρ^--mixing process which is neither NA nor ρ^*-mixing. Let $\{\xi_n, n \geqslant 1\}$, $\{\eta_n, n \geqslant 1\}$ and $\{\tau_n, n \geqslant 1\}$ be three independent sequences of i.i.d. standard normal random variables. Let

$$X_n = \begin{cases} \xi_m, & \text{if } n = 2m-1 \\ -\xi_m, & \text{if } n = 2m \end{cases},$$

$$Y_n = \begin{cases} \eta_m, & \text{if } n = 2^{2m-1} \\ -\eta_m, & \text{if } n = 2^{2m} \\ \tau_n, & \text{otherwise} \end{cases},$$

and $Z_n = {X_n}^2 + Y_n$. From Zhang and Wang (1999), it follows that $\{Z_n, n \geqslant 1\}$ is ρ^--mixing. For the density function of Z_n, noting that

$$F_{Z_n}(t) = P(X_n^2 + Y_n \leqslant t) = \int_{-\infty}^{t} d\Phi(y) \int_{-\sqrt{t-y}}^{\sqrt{t-y}} d\Phi(x) = \int_{-\infty}^{t} (2\Phi(\sqrt{t-y}) - 1) d\Phi(y),$$

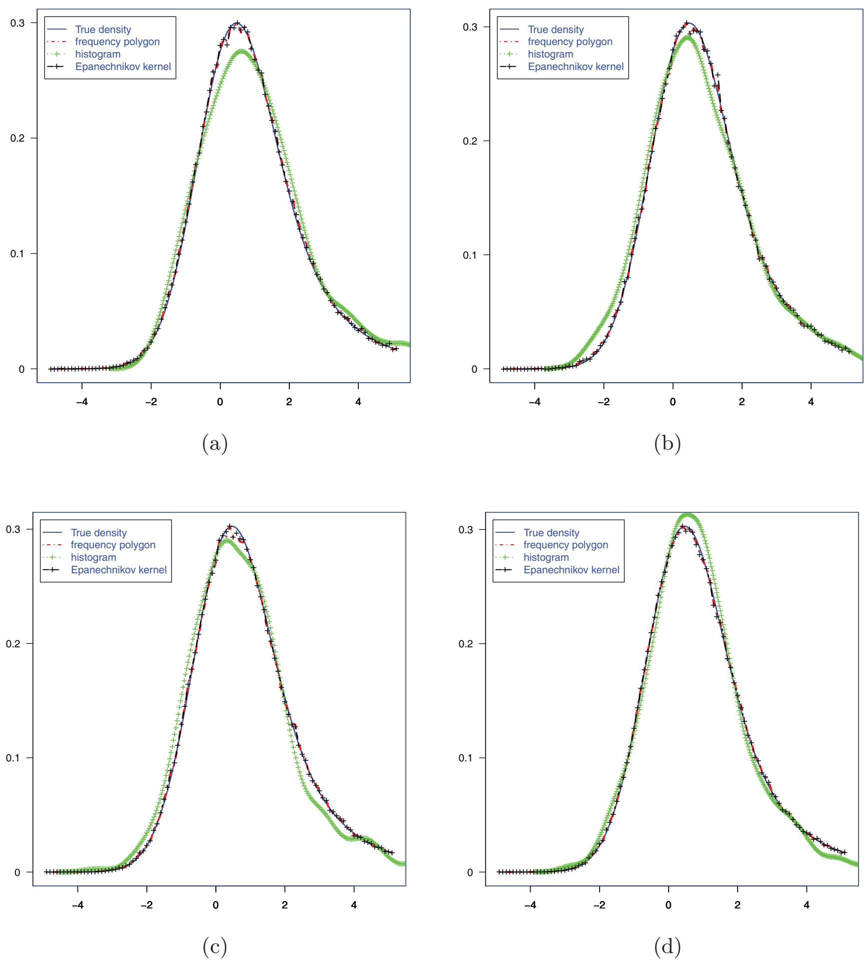

Figure 2: *Comparsion of different estimators for* $n = 100, 200, 300, 400$ *under Case 2*

thus,

$$f_{Z_n}(t) = \frac{dF_{Z_n}(t)}{dt} = \int_{-\infty}^{t} \varphi(\sqrt{t-y})(t-y)^{-\frac{1}{2}}\varphi(y)dy$$
$$= 2\int_0^{\infty} \varphi(s)\varphi(t-s^2)ds = \frac{1}{\pi}\int_0^{\infty} e^{-\frac{s^2+(t-s^2)^2}{2}}ds.$$

Other settings are the same as those in Case 1, we also use R software to compute the three estimators for 500 times to obtain the final values and then compare them with $f(x)$ in Figure 2 under different sample sizes. It can be seen that Figure 2 shows a similar conclusion as those in Case 1.

In specific, we also compute the RMSE of three estimators as presented in Table 1 under different sample sizes, respectively. We can see that in both two cases, the more the sample is, the smaller RMSE is. It also reveals that the RMSE of both the frequency polygon estimator and Epanechnikov kernel estimator are smaller than that of the histogram estimator for each sample size, and there is no obvious difference between the RMSE of the frequency polygon estimator and Epanechnikov kernel estimator.

Table 1: *The RMSE of the estimators*

Case	Estimator	$n = 100$	$n = 200$	$n = 300$	$n = 400$
	frequency polygon	0.051820	0.040587	0.035284	0.032109
1	Epanechnikov kernel	0.048170	0.038136	0.033279	0.030406
	histogram	0.103775	0.087102	0.079361	0.073087
	frequency polygon	0.002488	0.002027	0.001993	0.001464
2	Epanechnikov kernel	0.002101	0.001645	0.001597	0.001099
	histogram	0.021402	0.018259	0.016892	0.015706

4. Proofs of the main results

In this section, we first present some lemmas which are useful in proving the main results.

LEMMA 4.1. (Zhang and Wang, 1999) *Increasing functions defined on disjoint subsets of a ρ^--mixing field $\{X_k; k \in N^d\}$ with mixing coefficients $\rho^-(s)$ are also ρ^--mixing with mixing coefficients not greater than $\rho^-(s)$.*

LEMMA 4.2. (Wang and Lu, 2006) *For a positive real number $q \geqslant 2$, if $\{X_n, n \geqslant 1\}$ is a sequence of ρ^--mixing variables with $EX_i = 0, E|X_i|^q < \infty$ for every $i \geqslant 1$, then for all $n \geqslant 1$, there exists a positive constant $C = C(q, \rho^-(\cdot))$ such that*

$$E\left(\max_{1\leqslant j\leqslant n}\left|\sum_{i=1}^{j} X_i\right|^q\right) \leqslant C\left(\sum_{i=1}^{n} E|X_i|^q + \left(\sum_{i=1}^{n} EX_i^2\right)^{q/2}\right). \tag{4.1}$$

Proof of Theorem 2.1. Noting that D is a compact subset of $\mathbb{R}$, we can assume that $D = [-B, B]$, without loss of generality, where B is a positive constant. Set $D_j = [(j-$

$1/2)b_n,(j+1/2)b_n)$, where $j=-r_n,-(r_n-1),\cdots,(r_n-1),r_n$ and $r_n=\lfloor B/b_n\rfloor+1$. Since $(r_n+1/2)b_n=(\lfloor B/b_n\rfloor+3/2)b_n\geqslant(\lfloor B/b_n\rfloor+1/2)b_n=B+b_n/2>B$, $\bigcup_{j=-r_n}^{r_n}D_j$ $=[-(r_n+1/2)b_n,(r_n+1/2)b_n)\supset[-B,B]$. Thus for any $\varepsilon>0$,

$$P\left(\sup_{x\in D}|\hat{f}(x)-E\hat{f}(x)|>\varepsilon\tau_n\right)=P\left(\max_{-r_n\leqslant j\leqslant r_n}\sup_{x\in D_j}|\hat{f}(x)-E\hat{f}(x)|>\varepsilon\tau_n\right)$$
$$\leqslant\sum_{j=-r_n}^{r_n}P\left(\sup_{x\in D_j}|\hat{f}(x)-E\hat{f}(x)|>\varepsilon\tau_n\right). \tag{4.2}$$

Recall that $-\frac{1}{2}\leqslant\frac{1}{2}+j-\frac{x}{b_n}\leqslant\frac{3}{2}$ and $-\frac{1}{2}\leqslant\frac{1}{2}-j+\frac{x}{b_n}\leqslant\frac{3}{2}$ for any $x\in D_j$, which implies that $\sup_{x\in D_j}|\frac{1}{2}+j-\frac{x}{b_n}|\leqslant\frac{3}{2}$ and $\sup_{x\in D_j}|\frac{1}{2}-j+\frac{x}{b_n}|\leqslant\frac{3}{2}$. Hence, we have

$$\begin{aligned}&P\left(\sup_{x\in D_j}|\hat{f}(x)-E\hat{f}(x)|>\varepsilon\tau_n\right)\\&\leqslant P\left((|f_j-Ef_j|+|f_{j+1}-Ef_{j+1}|)>2\varepsilon\tau_n/3\right)\\&\leqslant P(|f_j-Ef_j|>\varepsilon\tau_n/3)+P\left(|f_{j+1}-Ef_{j+1}|>\varepsilon\tau_n/3\right).\end{aligned}\tag{4.3}$$

For a given j, set $\zeta_i=I\{(j-1)b_n\leqslant X_i<jb_n\}-EI\{(j-1)b_n\leqslant X_i<jb_n\}$, $i=1,2,\cdots,n$. ζ_i can be decomposed into $\zeta_i=\zeta_i(1)-\zeta_i(2)$, where $\zeta_i(1)=I\{X_i\geqslant(j-1)b_n\}-EI\{X_i\geqslant(j-1)b_n\}$ and $\zeta_i(2)=I\{X_i\geqslant jb_n\}-EI\{X_i\geqslant jb_n\}$. It follows from Lemma 4.1 that $\{\zeta_i(1),1\leqslant i\leqslant n\}$ and $\{\zeta_i(2),1\leqslant i\leqslant n\}$ are also ρ^--mixing with $|\zeta_i(1)|\leqslant 2$ and $|\zeta_i(2)|\leqslant 2$. Thus, by the Markov inequality and lemma 4.2, it follows that for any $q>2$,

$$\begin{aligned}&P\left(|f_j-Ef_j|>\varepsilon\tau_n/3\right)\\&=P\left(\left|\sum_{i=1}^n(\zeta_i(1)-E\zeta_i(1))\right|+\left|\sum_{i=1}^n(\zeta_i(2)-E\zeta_i(2))\right|>\varepsilon\tau_nnb_n/3\right)\\&\leqslant P\left(\left|\sum_{i=1}^n(\zeta_i(1)-E\zeta_i(1)\right|>\varepsilon\tau_nnb_n/6\right)+P\left(\left|\sum_{i=1}^n(\zeta_i(2)-E\zeta_i(2)\right|>\varepsilon\tau_nnb_n/6\right)\\&\leqslant Cb_n^{-q}\tau_n^{-q}n^{-q}E\left|\sum_{i=1}^n(\zeta_i(1)-E\zeta_i(1))\right|^q+Cb_n^{-q}\tau_n^{-q}n^{-q}E\left|\sum_{i=1}^n(\zeta_i(2)-E\zeta_i(2)\right|^q\\&\leqslant Cb_n^{-q}\tau_n^{-q}n^{-q}\left(\left(\sum_{i=1}^nE|\zeta_i(1)|^q+\left(\sum_{i=1}^nE\zeta_i^2(1)\right)^{\frac{q}{2}}\right)\right.\\&\quad\left.+\left(\sum_{i=1}^nE|\zeta_i(2)|^q+\left(\sum_{i=1}^nE\zeta_i^2(2)\right)^{\frac{q}{2}}\right)\right)\\&\leqslant Cb_n^{-q}\tau_n^{-q}n^{-\frac{q}{2}},\end{aligned}$$

which together with $nb_n \to \infty$ and $\liminf\limits_{n\to\infty} n^{\delta} b_n \tau_n > 0$ for some $0<\delta<\frac{1}{2}$ that for sufficiently large q,

$$\begin{aligned}\sum_{n=1}^{\infty} r_n P(|f_j - Ef_j| > \varepsilon\tau_n/3) &\leqslant C\sum_{n=1}^{\infty} b_n^{-1-q}\tau_n^{-q} n^{-\frac{q}{2}} \\ &\leqslant C\sum_{n=1}^{\infty} b_n^{-q}\tau_n^{-q} n^{-\frac{q}{2}+1} \\ &\leqslant C\sum_{n=1}^{\infty} n^{-(\frac{1}{2}-\delta)q+1} < \infty. \qquad (4.4)\end{aligned}$$

Similarly, we also have

$$\sum_{n=1}^{\infty} r_n P\left(|f_{j+1} - Ef_{j+1}| > \varepsilon\tau_n/3\right) \leqslant C\sum_{n=1}^{\infty} n^{-(\frac{1}{2}-\delta)q+1} < \infty. \qquad (4.5)$$

A combination of (4.2)-(4.5) yields that

$$\sum_{n=1}^{\infty} P\left(\sup_{x\in D} |\hat{f}(x) - E\hat{f}(x)| > \varepsilon\tau_n\right) < \infty.$$

From the Borel-Cantelli lemma, we can easily get

$$\sup_{x\in D} |\hat{f}(x) - E\hat{f}(x)| = o(\tau_n) \text{ a.s..}$$

Now, we will prove (2.2). Set

$$\eta_i(x) = \left(\frac{1}{2} + j - \frac{x}{b_n}\right) I((j-1)b_n \leqslant X_i < jb_n) + \left(\frac{1}{2} - j + \frac{x}{b_n}\right) I(jb_n \leqslant X_i < (j+1)b_n)$$

for $x \in D_j$. Then we have

$$\hat{f}(x) = \frac{1}{nb_n}\sum_{i=1}^{n} \eta_i(x). \qquad (4.6)$$

Using Taylor's expansion for $F(jb_n)$ and $F((j-1)b_n)$ around $x \in D_j$, we have

$$\begin{aligned}&P((j-1)b_n \leqslant X_i < jb_n) = F(jb_n) - F((j-1)b_n) \\ &= F(x) + f(x)(jb_n - x) + O((jb_n - x)^2) \\ &\quad -[F(x) + f(x)((j-1)b_n - x) + O(((j-1)b_n - x)^2)] \\ &= f(x)b_n + O(b_n^2).\end{aligned}$$

Similarly, $P(jb_n \leqslant X_i < (j+1)b_n) = f(x)b_n + O(b_n^2)$. Therefore,

$$\begin{aligned}E\eta_i(x) &= \left(\frac{1}{2} + j - \frac{x}{b_n}\right)[f(x)b_n + O(b_n^2)] + \left(\frac{1}{2} - j + \frac{x}{b_n}\right)[f(x)b_n + O(b_n^2)] \\ &= f(x)b_n + O(b_n^2). \qquad (4.7)\end{aligned}$$

Noting that the term $O(b_n^2)$ in the above equality is independent of x and j, we have

$$\sup_{x\in D_j} |E\hat{f}(x) - f(x)| = \sup_{x\in D_j} \left| \frac{1}{nb_n} \sum_{i=1}^{n} E\eta_i(x) - f(x) \right| = O(b_n). \tag{4.8}$$

Thus,

$$\sup_{x\in D} \left| E\hat{f}(x) - f(x) \right| = O(b_n), \tag{4.9}$$

which together with (2.1) yields (2.3). The proof is completed. □

Proof of Theorem 2.2. Set $A_n = (-\infty, -n^T) \cap (n^T, \infty)$ for $T > 0$ and define $D_j = [(j-1/2)b_n, (j+1/2)b_n)$, where $j = -r_n, -(r_n-1), \cdots, (r_n-1), r_n$ with $r_n = \lfloor n^T/b_n \rfloor + 1$. Then,

$$\begin{aligned} P\left(\sup_{x\in[-n^T, n^T]} |\hat{f}(x) - E\hat{f}(x)| > \varepsilon\tau_n \right) &= P\left(\max_{-r_n \leqslant j \leqslant r_n} \sup_{x\in D_j} |\hat{f}(x) - E\hat{f}(x)| > \varepsilon\tau_n \right) \\ &\leqslant \sum_{j=-r_n}^{r_n} P\left(\sup_{x\in D_j} |\hat{f}(x) - E\hat{f}(x)| > \varepsilon\tau_n \right). \end{aligned}$$

By the proof of Theorem 2.1, (2.4) can be derived. In what follows, we will prove (2.5).

Since $A_n = (-\infty, -n^T) \bigcup (n^T, \infty)$ and $E|X_1|^{2/T} < \infty$ for $T > 0$,

$$\begin{aligned} \sum_{n=1}^{\infty} P\left(\sup_{x\in A_n} \hat{f}(x) > n^{-1} \right) &\leqslant \sum_{n=1}^{\infty} P\left(\bigcup_{i=1}^{n} \{X_i \in A_n\} \right) \\ &\leqslant \sum_{n=1}^{\infty} \sum_{i=1}^{n} P\left(|X_i| > n^T \right) \\ &= \sum_{n=1}^{\infty} nP\left(|X_1| > n^T \right) \\ &= \sum_{n=1}^{\infty} n \sum_{j=n}^{\infty} P\left(j^T < |X_1| \leqslant (j+1)^T \right) \\ &= \sum_{j=1}^{\infty} P\left(j^T < |X_1| \leqslant (j+1)^T \right) \sum_{n=1}^{j} n \\ &\leqslant C \sum_{j=1}^{\infty} j^2 P\left(j < |X_1|^{\frac{1}{T}} \leqslant (j+1) \right) \\ &= C \sum_{j=1}^{\infty} E|X_1|^{\frac{2}{T}} I(j < |X_1|^{\frac{1}{T}} \leqslant (j+1)) \\ &\leqslant CE|X_1|^{\frac{2}{T}} < \infty. \end{aligned}$$

Therefore, we obtain that

$$\sup_{x\in A_n} \hat{f}(x) = O(n^{-1}) \ \text{ a.s..} \tag{4.10}$$

Noticing that $f(x)$ is differentiable for $x \in R$ and $|f'(x)| < M$, we have for any $x \in A_n$,

$$0 \leqslant f(x) = |x|^{-\frac{2}{T}} |x|^{\frac{2}{T}} f(x) \leqslant n^{-2}. \tag{4.11}$$

Thus, it follows that

$$\sup_{x \in A_n} f(x) = o(n^{-2}). \tag{4.12}$$

Combining (4.10) and (4.12) yields that

$$\sup_{x \in A_n} |\hat{f}(x) - f(x)| = n^{-2} = o(b_n) \ \text{ a.s.},$$

which together with (2.4) implies that (2.5) holds. The proof is complete. □

Acknowledgements. The authors are most grateful to the Editor-in-Chief and anonymous referees for carefully reading the manuscript and valuable suggestions which helped in improving an earlier version of this paper.

REFERENCES

[1] N. BESAID, S. DABO-NIANG, *Frequency polygons for continuous random fields*, Statistical Inference for Stochastic Processes, 10: 55–80 (2010).

[2] M. CARBON, B. GAREL, L. T. TRAN, *Frequency polygons for weakly dependent processes*, Statistics and Probability Letters, 33: 1–13, (1997).

[3] M. CARBON, C. FRANCQ, L. T. TRAN, *Asymptotic normality of frequency polygons for random fields*, Journal of Statistical Planning and Inference, 140 (2): 502–514, (2010).

[4] H. W. HUANG, J. Y. PENG, X. T. WU, B. WANG, *Complete convergence and complete moment convergence for arrays of rowwise ANA random variables*, Journal of Inequalities and Applications, Article ID: 72, (2016).

[5] K. JOAG-DEV, F. PROSCHAN, *Negative association of random variables with applications*, Annals of Statistics, 11 (1): 286–295, (1983).

[6] D. W. SCOTT, *Frequency polygons: theory and application*, Journal of the American Statistical Association, 80 (390): 348–354, (1985).

[7] X. J. WANG, Y. WU, S. H. HU, *The Berry–Esseen bounds of the weighted estimator in a nonparametric regression model*, Annals of the Institute of Statistical Mathematics, 71, 1143–1162, (2019).

[8] J. F. WANG, F. B. LU, *Inequalities of maximum partial sums and weak convergence for a class of weak dependent random variables*, Acta Mathematica Sinica, English Series, 22 (3): 693–700, (2006).

[9] G. D. XING, S. C. YANG, X. LIANG, *On the uniform consistency of frequency polygons for ψ-mixing samples*, Journal of the Korean Statal Society, 44 (2): 179–186, (2015).

[10] G. D. XING, S. C. YANG, *Uniformly strong consistency of frequency polygons for negatively associated samples*, Communications in Statistics-Simulation and Computation, 46 (3): 2168–2175, (2015).

[11] L. X. ZHANG, *A functional central limit theorem for asymptotically negatively dependent random fields*, Acta Mathematica Hungarica, 86 (3): 237–259, (2000).

[12] L. X. ZHANG, *Central limit theorems for asymptotically negatively associated random fields*, Acta Mathematica Sinica, English Series, 16 (4): 691–710, (2000).

[13] Y. ZHANG, *Complete moment convergence for moving average process generated by ρ^--mixing random variables*, Journal of Inequalities and Applications, Article ID: 245, (2015).

[14] L. X. ZHANG, X. Y. WANG, *Convergence rates in the strong laws of asymptotically negatively associated random fields*, Applied Mathematics-A Journal of Chinese Universities, Series B, 14 (4): 406–416, (1999).

(Received January 16, 2021)

Wei Wang
School of Big Data and Artificial Intelligence
Chizhou University
Chizhou, 247000, P. R. China

Haiwu Huang
College of Science
Guilin University of Aerospace Technology
Guilin, 541002, P. R. China

Yi Wu
School of Big Data and Artificial Intelligence
Chizhou University
Chizhou, 247000, P. R. China

Kan Chen
School of Mathematics and Statistics
Chaohu University
Hefei, 238024, P. R. China
e-mail: kanchenchu@126.com