Probability in the Engineering and Informational Sciences (2026), **40:**4 598–617
doi:10.1017/S0269964826100254

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Asymptotic behavior of frequency polygon estimation of density function for dependent samples

Yan Wang[1], Haiwu Huang[2] and Aiting Shen[3]

[1]School of Big Data and Artificial Intelligence, Chizhou University, Chizhou, P.R. China
[2]School of Science, Guilin University of Aerospace Technology, Guilin, P.R. China
[3]School of Big Data and Statistics, Anhui University, Hefei, P.R. China
Corresponding author: Aiting Shen; Email: empress201010@126.com

Keywords: frequency polygon; real data analysis; simulation study; weak consistency

Mathematical Subject Classification: 62G07; 62G20

This paper pays attention to the frequency polygon, which is constructed by connecting with straight lines the mid-bin values of a histogram. As a density estimator based on the histogram technique, the frequency polygon has the advantage of computational simplicity and has been widely used in many fields. The purpose of this article is to investigate the weak consistency, the uniformly weak consistency, and the rate of the uniformly weak consistency for frequency polygon estimation of the density function under α-mixing samples, which improve and extend the corresponding ones in the literature. In addition, the simulation study and real data analysis are also presented to verify the validity of the theoretical results based on finite samples.

1. Introduction

The convergence in probability is straight in the analysis tradition and can be thought of as variants of standard pointwise convergence. It is a kind of weak convergence, and it is also the basis of probability and statistics. We know that the weaker the requirements for convergence, the easier it is for a sequence of random variables to have a limit. Strong consistency or almost sure convergence is theoretically stricter and often overly stringent for empirical work. We point out that weak consistency or convergence in probability ensures that the estimator converges to the true parameter, which is sufficient for most real-world applications. This feature will be of great help for us to study some issues that we care about.

In recent years, scholars have paid more attention to computational cost and are more willing to choose estimation methods that require less computation. As we all know, the histogram plays an important role in data representation and density estimation, because it has simple calculations and a wide application. Therefore, the research and application of various estimations based on histogram technology are interesting. This paper will pay attention to the frequency polygon, which is constructed by connecting with straight lines the mid-bin values of a histogram. The calculation of constructing the frequency polygon is approximately equivalent to the histogram. Because the frequency polygon has these advantages, it is of great significance to study it deeply.

With regard to the frequency polygon density estimator, some results have been achieved. Some classic results are taken, for example, [8] analyzed its large sample properties for independent samples, [3] extended the discussion to the strong mixing samples and showed the integrated mean squared error and uniformly strong consistency of the estimator, Besaid and Dabo-Niang [1] studied the mean square error and uniformly strong consistency for the continuous random fields and so on. Recently, the asymptotic properties of this estimator have been considered by many scholars. Let us review the asymptotic

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (http://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

behavior of consistency, [12] studied the uniformly strong consistency of histogram density estimator and the convergence rate under strong mixing process, Kang et al. (2017) studied the asymptotic behavior of the uniformly strong consistency of the frequency polygon investigated by Scott (1985, 2015) under ϕ-mixing samples, [11] studied the uniformly complete consistency and the rates of uniformly complete consistency of frequency polygon estimation for widely orthant dependent samples under some general conditions and showed an application that obtained the uniformly complete consistency and the corresponding rates of hazard rate function estimator and so on. In particular, [13] considered a simple estimator of the density mode using the frequency polygon estimate and investigated the strong consistency of the estimator for strong mixing sequence. Let us review the asymptotic behavior of distribution, [2] established the asymptotic normality of frequency polygons for random fields, [10] focused on the uniformly asymptotic normality of frequency polygons for ψ-mixing samples, [5] focused on the Berry–Esséen bound of frequency polygons for ϕ-mixing samples, and so on.

As mentioned earlier, weak consistency plays an important role in probability limit theory. However, there is little literature on the weak consistency of the estimator under dependent samples. This paper mainly studies the weak consistency, the uniformly weak consistency, and the rate of the uniformly weak consistency for frequency polygon estimation of the density function under α-mixing samples, when $f(x)$ is defined on a compact subset of the real line or on the real line. α-mixing represents a weaker and more broadly applicable dependence structure than ψ-mixing or φ-mixing-mixing. Notably, α-mixing is particularly suitable for financial applications, as the weak dependence it captures aligns well with the empirically observed dynamics in market returns, covering a wider range of realistic process categories than its stronger counterparts. Establishing convergence properties under this weaker condition is more challenging but significantly extends the proven validity of the estimator. Let us recall the following definition of α-mixing random variables, which was introduced by [7].

Definition 1.1. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables, and let $\mathcal{F}_1^k = \sigma(X_j; 1 \leq j \leq k)$ and $\mathcal{F}_{k+n}^{\infty} = \sigma(X_j; j \geq k+n)$ be two σ-algebras. If*

$$\alpha(n) = \sup_{k \geq 1} \sup_{A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^{\infty}} |P(AB) - P(A)P(B)| \to 0, \text{ as } n \to \infty,$$

then $\{X_n, n \geq 1\}$ is called α-mixing.

In what follows, we will introduce the definition of frequency polygon in detail. Suppose that X is a random variable with a density function $f(x)$. Let $\{X_i, i \geq 1\}$ be the sample drawn from the population X, which is a sequence of identically distributed random variables with common density $f(x)$ on the real line $\mathbb{R}$. Consider a partition $\cdots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 \cdots$ of the real line into equal intervals $I_{n,j} := [(j-1)b_n, jb_n)$ of length b_n, where b_n is the binwidth and $j = 0, \pm 1, \pm 2, \cdots$. For $x \in [(j-1/2)b_n, (j+1/2)b_n)$, consider the two adjacent histogram bins $I_{n,j}$ and $I_{n,j+1}$. Denote the number of observations falling in these intervals, respectively, by $\upsilon_{n,j}$ and $\upsilon_{n,j+1}$. Then the values of the histogram in these previous bins are given by

$$f_{n,j} = \frac{\upsilon_{n,j}}{nb_n}, \; f_{n,j+1} = \frac{\upsilon_{n,j+1}}{nb_n}. \tag{1.1}$$

Thus the frequency polygon estimation of the density function $f(x)$ is defined as follows

$$f_n(x) = \left(\frac{1}{2} + j - \frac{x}{b_n}\right) f_{n,j} + \left(\frac{1}{2} - j + \frac{x}{b_n}\right) f_{n,j+1}, \tag{1.2}$$

for $x \in [(j-1/2)b_n, (j+1/2)b_n)$. In many literatures, j is usually regarded as zero. Here it is reserved for the convenience of description later.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Generally speaking, weak consistency, which ensures that the estimator converges to the true parameter value in probability, often provides a sufficient and more readily attainable guarantee of reliability for large-sample inferences. The frequency polygon, with its computational simplicity, is an attractive tool for such applications, provided its theoretical foundations under realistic dependence are firmly established. The primary contribution of this paper is to bridge this gap. We rigorously establish the weak consistency, uniformly weak consistency, and the corresponding rates of convergence for the frequency polygon density estimator under α-mixing samples. This work extends the theoretical justification for employing this computationally efficient nonparametric estimator to a comprehensive range of dependent data scenarios, reinforcing its utility in practical statistical analysis where complex dependence is present.

The organization of this paper is as follows. Section 2 contains the assumptions and main results obtained. A divergent discussion of the main results is put in Section 3. Data examples, including a simulation study and real data analysis, are provided in Sections 4 and 5, respectively. Some important lemmas are listed in Section 6, and the corresponding proofs are detailed in Section 7.

For the sake of simplicity, let $\lfloor x \rfloor$ stand for the integer part of x, let C and C_1 be positive constants, whose values may be different in different places. Denote $a_n = O(b_n)$ instead of $a_n \leq Cb_n$, and $a_n = o(b_n)$ instead of $a_n/b_n \to 0$ as $n \to \infty$, where $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ are both the sequences of positive numbers.

2. Assumptions and main results

2.1. Assumptions

Firstly, we make some basic assumptions.

(*A*1) As $n \to \infty$, $b_n \to 0$ and $nb_n/\log n \to \infty$;

(*A*2) $\{X_i, i \geq 1\}$ is a sequence of stationary α-mixing random variables. X_1 has an uniformly continuous density $f(x)$ on the real line $\mathbb{R}$ and the mixing coefficient satisfies $\alpha(n) = O(n^{-\lambda})$ for some $\lambda \geq 2$;

(*A*3) The joint density function $f_{ij}(x, y)$ of (X_i, X_j) exists and for some positive constant M, it satisfies

$$\sup_{(x,y)\in\mathbb{R}^2} |f_{ij}(x, y) - f(x)f(y)| \leq M, \text{ for any } i \neq j;$$

(*A*4) The density function $f(x)$ satisfies the Lipschitz condition of order 1, i.e., for all $x, y \in \mathbb{R}$ and some $L > 0$,

$$|f(x) - f(y)| \leq L|x - y|;$$

(*A*5) $E|X_1|^{1/T} < \infty$ for some $T > 0$.

Remark 2.1. The assumptions are general and also weaker than those assumed in the literature, such as [12], [5], [6] and [13]. (*A*1) ensures asymptotic unbiasedness and adequate bin occupancy for variance control; (*A*2) facilitates Bernstein-type inequalities for dependent data; (*A*3) restricts joint densities, which is milder than uniform boundedness often required in kernel estimation; (*A*4) is a standard Lipschitz condition; (*A*5) requires a weaker moment condition, which is essential for modeling heavier-tailed data, which significantly expands applicability to real-world datasets where higher moments may not exist. This contrasts with the stronger order-$2/T$ moment required for strong consistency in [12], for instance.

Next, we give the main results as follows.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

2.2. Main results

Theorem 2.2. *Suppose that* $(A1)--(A4)$ *are satisfied. If*

$$g(n,1) := b_n^{-1}(nb_n/\log n)^{-\lambda} \to 0, \tag{2.1}$$

then for any $x \in \mathbb{R}$,

$$|f_n(x) - f(x)| = o(1) \text{ in probability}. \tag{2.2}$$

Consider the optimal binwidth $b_n = Cn^{-1/5}$ (Carbon et al. (1979)), it is clear that $(A1)$ is satisfied. Hence, we get the following corollary.

Corollary 2.3. *Suppose that* $b_n = Cn^{-1/5}$ *and* $(A2)--(A3)$ *are satisfied. Then for any* $x \in \mathbb{R}$, *(2.2) holds.*

Theorem 2.4 *Suppose that* $(A1)--(A5)$ *are satisfied. Then we have the following two statements:*
(1) *If*

$$g(n,2) := b_n^{-2}(nb_n/\log n)^{-\lambda} \to 0, \tag{2.3}$$

then for any compact subset D of $\mathbb{R}$,

$$\sup_{x\in D}|f_n(x) - f(x)| = o(1) \text{ in probability}; \tag{2.4}$$

(2) *If* $g(n,2)n^T \to 0$, *then*

$$\sup_{x\in \mathbb{R}}|f_n(x) - f(x)| = o(1) \text{ in probability}. \tag{2.5}$$

Analogously, we have the following corollary.

Corollary 2.5. *Suppose that* $b_n = Cn^{-1/5}$ *and* $(A2)--(A5)$ *are satisfied. Then for any compact subset D of* $\mathbb{R}$, *(2.4) holds. Furthermore, if* $\lambda > \max\left\{2, \frac{2+5T}{4}\right\}$, *then (2.5) holds.*

Remark 2.6. The following theorem is the rate of the uniformly weak consistency for frequency polygon estimation of the density function under α-mixing samples.

Theorem 2.7. *Suppose that* $(A1)--(A5)$ *are satisfied. Denote* $R_n = \max\left\{b_n, \left(\frac{\log n}{nb_n}\right)^{1/2}\right\}$. *Then we have the following two statements:*
(1) *If*

$$g(n,3) := b_n^{-2}(nb_n/\log n)^{-\lambda/2+1/2} \to 0, \tag{2.6}$$

then for any compact subset D of $\mathbb{R}$,

$$\sup_{x\in D}|f_n(x) - f(x)| = O(R_n) \text{ in probability}; \tag{2.7}$$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

(2) *If* $g(n,3)n^T \to 0$, *then*

$$\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = O(R_n) \ \textit{in probability}. \tag{2.8}$$

Clearly, we have $R_n \to 0$ as $n \to \infty$, so Theorem 2.7 implies that $f_n(x)$ converges in probability to $f(x)$ with the rate R_n. Analogously, we have the following corollary.

Corollary 2.8. *Suppose that* $b_n = Cn^{-1/5}$ *and* $(A2)--(A5)$ *are satisfied. If* $\lambda > 2$, *then for any compact subset* D *of* $\mathbb{R}$, *(2.7) holds. Furthermore, if* $\lambda > 2 + \frac{5T}{2}$, *then (2.8) holds.*

Remark 2.9. From Corollary 2.3, if the binwidth $b_n = Cn^{-1/5}$, then $R_n = b_n$; if the binwidth $b_n = C(n^{-1}\log n)^{1/3}$, then $R_n = (n^{-1}\log n)^{1/3}$, which is the optimal rate of convergence for the histogram density estimation in the i.i.d. case.

Remark 2.10. Theorems 2.4 and 2.7 state the uniformly weak consistency for frequency polygon estimation of density function. Compared with the corresponding results of [12] or other related literatures establishing strong consistency under the moment condition $E|X|^{2/T} < \infty$ for some $T > 0$, it is noted that a lower moment condition is assumed here, which is $E|X|^{1/T} < \infty$ for some $T > 0$.

Remark 2.11. In fact, the weaker the requirements for convergence, the easier it is for a sequence of random variables to have a limit, we know that the existence of weak consistency can solve some issues. It is an important subject to obtain weak convergence results under weak conditions, and our results generalize and improve the existing ones in the literature.

3. Simulation study

This simulation study aims to verify the theoretical consistency of the frequency polygon estimator under finite α-mixing samples and to examine its performance across different dependence strengths.

We generate data from an ARMA(1,1) model, a standard framework for studying dependent processes. Consider the ARMA(1,1) model:

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}, \tag{3.1}$$

where $\{\varepsilon_t, t \geq 1\}$ are independent and identically distributed, $\varepsilon_i \sim N(0,\sigma_\varepsilon^2)$, $|\phi| < 1, |\theta| < 1$ and $\sigma_\varepsilon^2 = (1-\phi^2)/(1-2\phi\theta+\theta^2)$. Obviously, we know $X_t \sim N(0,1)$. We know from [4] and Yang (2023) that $\{X_t, t \geq 1\}$ is an α-mixing and stationary time series with $\alpha(k) = O(|\phi|^k)$ for the case $\phi \neq \theta$.

The specific scenarios are defined by the autoregressive parameter ϕ, which directly controls the dependence strength. We consider $\phi \in \{0, 0.1, 0.3, 0.4, 0.5, 0.6, 0.8\}$, covering a spectrum from independence ($\phi = 0$) to strong positive dependence ($\phi = 0.8$). For each ϕ, complementary moving-average parameters θ are selected to form distinct. Sample sizes are set to $n = 1250$, 2500, 5000, and 10^4 to assess the estimator's behavior as the sample increases.

[3] proved the best possible choice of the binwidth, namely $h_{FP} = 2\left(\frac{15}{49nR''(f)}\right)^{1/5}$, and $R''(f) = \frac{3}{8\sqrt{\pi}\sigma^5}$ can be calculated in our experiments, then optimal binwidth in our experiments is $b_n = h_{FP} = 2.153366n^{-1/5}\sigma$. It is called the optimal binwidth under the integrated mean squared error. The experiment is repeated $N = 100$ times in each scenario. For $x_i = -3 + 0.05(i-1)$ where $1 \leq i \leq m$ and $m = 121$, we compute the estimates $f_n(x_i)$, the global mean absolute relative deviations

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Table 1. *The values of $GMARD_f(n)$ (left) and $GR_f(n)$ (right) in different scenarios.*

ϕ	θ	$n = 1250$	$n = 2500$	$n = 5000$	$n = 10^4$	$n = 1250$	$n = 2500$	$n = 5000$	$n = 10^4$
0	0.1	0.12947	0.09294	0.07034	0.05345	0.01783	0.01487	0.01351	0.01181
0.1	0	0.13700	0.09783	0.06938	0.05257	0.01844	0.01591	0.01358	0.01168
0.1	0.3	0.11705	0.08151	0.06182	0.04965	0.01702	0.01404	0.01202	0.01089
0.3	0.5	0.11412	0.08405	0.06324	0.04690	0.01658	0.01391	0.01196	0.01099
0.4	0.3	0.13480	0.09662	0.07318	0.05327	0.01844	0.01556	0.01329	0.01205
0.4	0.6	0.10715	0.07981	0.06215	0.04926	0.01547	0.01320	0.01194	0.01098
0.5	0.4	0.12966	0.09515	0.07380	0.05250	0.01893	0.01573	0.01356	0.01153
0.6	0.4	0.11591	0.08932	0.07350	0.05722	0.01712	0.01523	0.01428	0.01326
0.8	0.6	0.12422	0.10590	0.09539	0.09587	0.01953	0.01824	0.01835	0.01999

$$GMARD_f(n) = \frac{1}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{N}\frac{|f_n(x_i,k) - f(x_i)|}{f(x_i)}, \tag{3.2}$$

and the values of

$$GR_f(n) := \frac{1}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{N}\frac{|f_n(x_i,k) - f(x_i)|}{R_n}, \tag{3.3}$$

where $f(x)$ is the density function of standard normal distribution, $f_n(x,k)$ is the estimate of density function in k-th experiment and R_n takes b_n here. $GMARD_f(n)$ measures the absolute estimation error $|f_n(\cdot) - f(\cdot)|$, while $GR_f(n)$ quantifies the relative rate at which this error shrinks in comparison to the theoretical convergence rate R_n. Specifically, the estimates $f_n(x)$ in different scenarios are shown in Figs. 1–3, the $GMARD_f(n)$ and $GR_f(n)$ obtained in different scenarios are listed in Tab. 1.

Concluding remarks

(1) Figs. 1–3 show us that when the sample size increases, the fitting effects of the three curves become better based on the same sample dependence. While the dependence enhances, the fitting effects of these three curves also become worse under the same sample size. Furthermore, as sample sizes increase, the estimation quality also correspondingly improves.

(2) Tab. 1 reveals that higher values of ϕ (e.g., $\phi = 0.8$) introduce stronger dependence and long-range dependence in the series, leading to increased estimation errors (e.g., $GMARD_f(n)$ up to 0.09587 for $n = 10^4$).

(3) Increasing θ (e.g., $\theta = 0.6$ for $\phi = 0.4$) can counteract the dependence induced by ϕ to some extent, and then reduce estimation errors (e.g., $GMARD_f(n)$ decreases from 0.13480 to 0.10715 for $n = 1250$).

(4) Weak dependence (e.g., $\phi = 0.1$ or balanced ϕ/θ pairs like $\phi = 0.3, \theta = 0.5$) yields the better performance of estimation, with lower $GMARD_f(n)$ and $GR_f(n)$ values across sample sizes.

(5) With the increasing of the sample size, the value of $GR_f(n)$ decreases. Mathematically, $GR_f(n)$ is a representation for the convergence rate of the estimation $f_n(\cdot)$, indicating that $f_n(\cdot)$ either maintains the same convergence rate as R_n or even surpasses it.

In short, $GMARD_f(n)$ decrease with the increase of sample size n, which shows that the estimation is consistent based on α-mixing samples in this case. Strong dependence reduces estimation quality, and at this time, larger samples are needed as compensation. These findings demonstrate our theoretical results.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

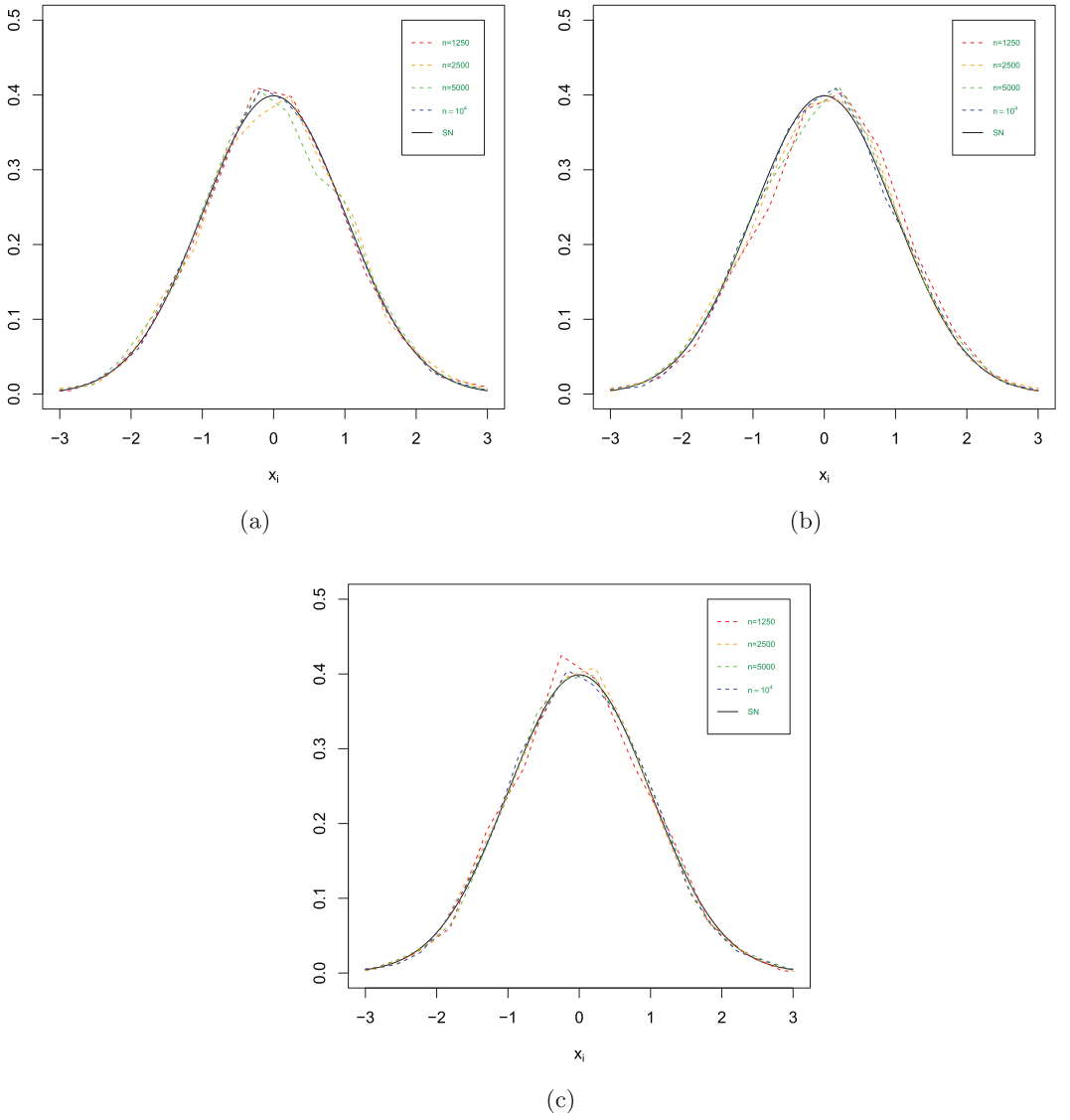


Figure 1. *The estimated values in different scenarios. (a)* $\phi = 0$ *and* $\theta = 0.1$. *(b)* $\phi = 0.1$ *and* $\theta = 0$. *(c)* $\phi = 0.1$ *and* $\theta = 0.3$.

4. Real data analysis

The statistical analysis of continuous data is a surprisingly recent development. The goal of this section is to demonstrate the practical applicability of the frequency polygon estimator for financial return data, which exhibits volatility clustering and heavy tails. As an example of financial time series data, Fig. 4 shows the rate of the daily return (or percentage change) of Dow Jones Industrial Average (DJIA) from April 21st, 2006 to April 20th, 2016. This is the typical rate of return data. The average value of the series seems to be stable, and the average rate of return is close to zero. However, the periods with great fluctuations (changes) tend to gather together. One problem in the analysis of these types of financial data is to estimate the density function, which is helpful to predict the volatility of future returns. Let d_t be the actual value of DJIA, $r_t = (d_t - d_{t-1})/d_{t-1}$ is the rate of return, then we have $1 + r_t = d_t/d_{t-1}$,

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

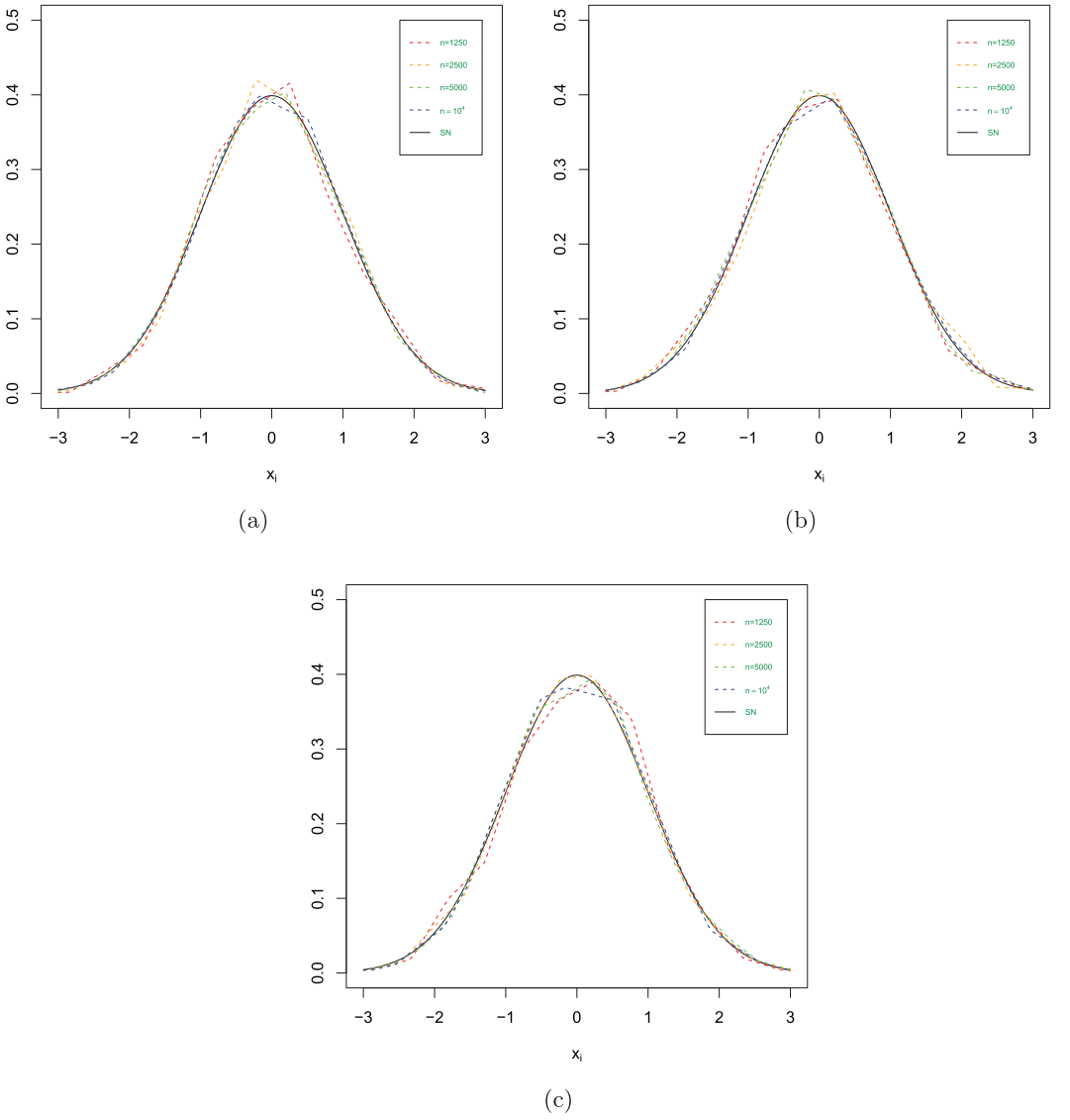


Figure 2. *The estimated values in different scenarios. (a)* $\phi = 0.3$ *and* $\theta = 0.5$. *(b)* $\phi = 0.4$ *and* $\theta = 0.3$. *(c)* $\phi = 0.4$ *and* $\theta = 0.6$.

and $\log(1 + r_t) = \log(d_t/d_{t-1}) = \log(d_t) - \log(d_{t-1}) \approx r_t$. This data set can also be obtained in the R package (astsa).

Observing Fig. 4, we note that the mean remains nearly constant, oscillating around 0, indicating that the series might be stationary. The PACF's significant spike at lag 2 (exceeding the confidence band) necessitates AR(2), while the ACF's gradual decay implies MA terms are needed. Fig. 5 suggests that the dataset could be an ARMA(2,2) process. After conducting the augmented Dickey-Fuller (ADF) unit root test, we find that the test statistic is −39.8484, which is less than the critical value of −2.58 at the 1% significance level ($-39.8484 < -2.58$). Therefore, at the 1% significance level, we reject the null hypothesis, concluding that the series is stationary. A stationary ARMA process $\{X_t, t \geq 1\}$ is α-mixing according to [4].

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

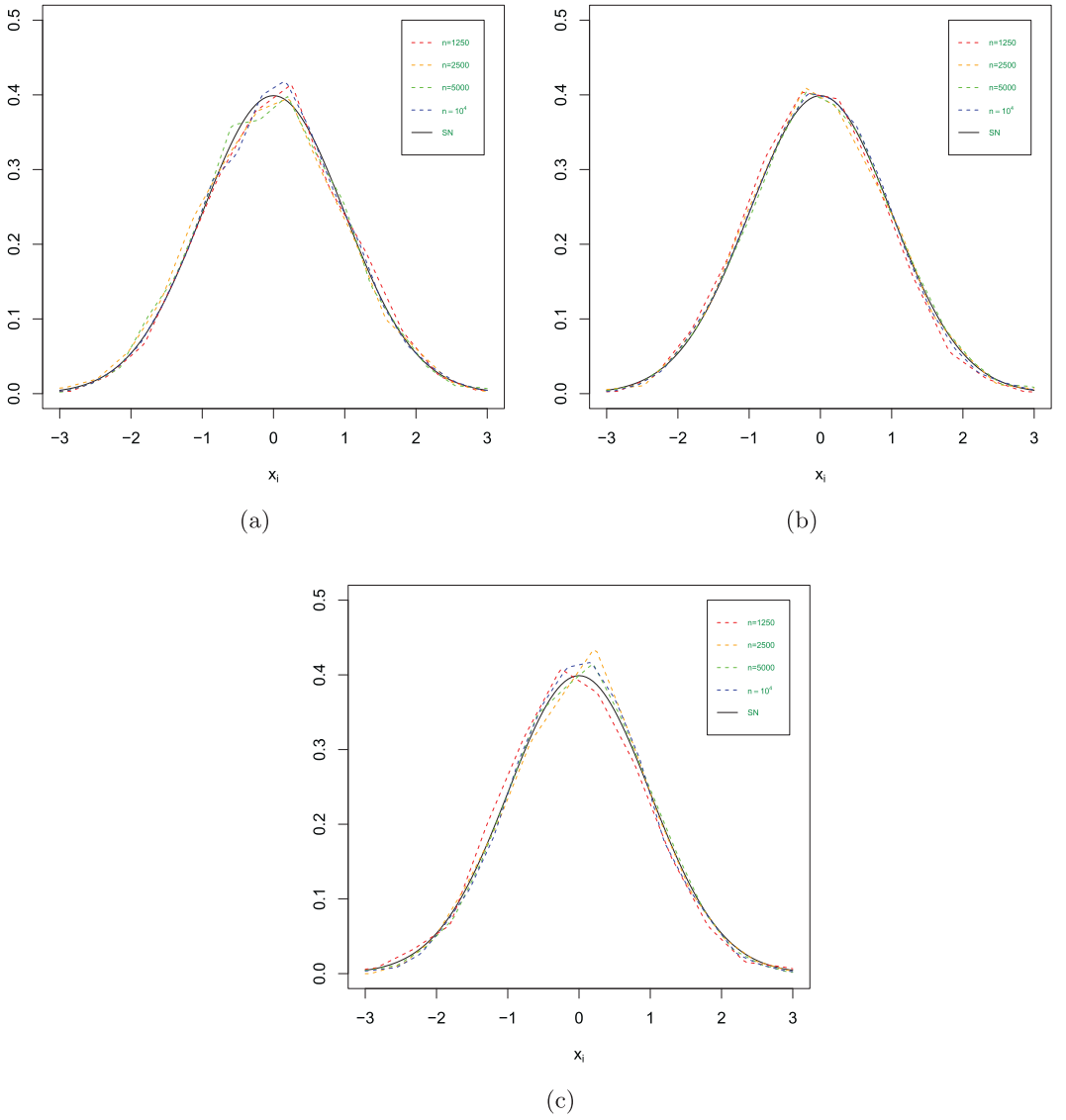


Figure 3. *The estimated values in different scenarios. (a)* $\phi = 0.5$ *and* $\theta = 0.4$. *(b)* $\phi = 0.6$ *and* $\theta = 0.4$. *(c)* $\phi = 0.8$ *and* $\theta = 0.6$.

By employing the calculation method in the Simulation study section, we can derive the frequency polygon estimation of the daily return rate (or percentage change rate) of DJIA. Figure 6 shows leptokurtic features (peaked center and fat tails) compared to the red normal distribution curve, consistent with empirical finance data.

In addition, the GARCH (generalized autoregressive conditional heteroskedasticity) model is particularly adept at capturing the variability of volatility in financial time series, and it can provide valuable insights for risk management and investment strategies. The frequency polygon estimator provides a complementary nonparametric approach for modeling return distributions. To delve deeper into comprehending the volatility pattern and structure specifically exhibited by the daily return rate of DJIA, we utilize the "rugarch" package in R to analyze the data set, and then derive the conditional variance equation of GARCH(1,1) model as follows:

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

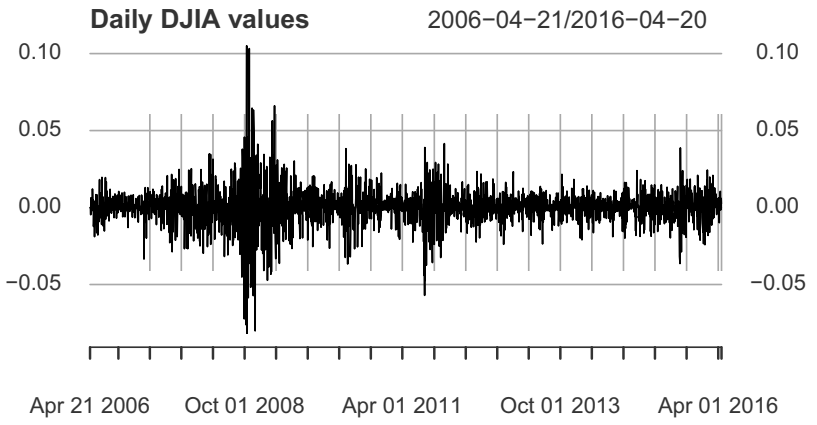


Figure 4. *The rate of the daily return (or percentage change) of DJIA from April 20th, 2006 to April 20th, 2016.*

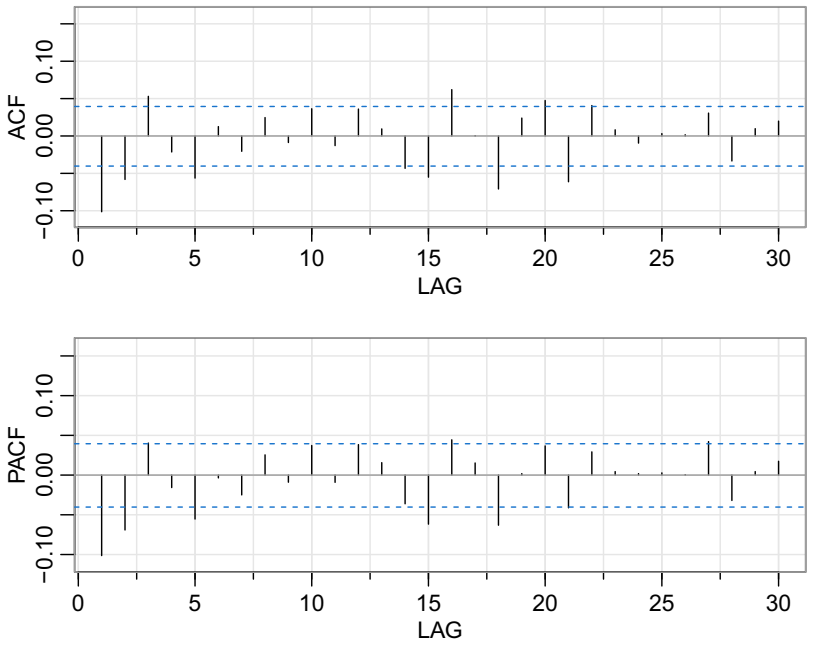


Figure 5. *The ACF and PACF plots of the daily return rate of DJIA from April 20th, 2006 to April 20th, 2016.*

$$h_t = 2.192579 \times 10^{-6} + 0.1209623\varepsilon_{t-1}^2 + 0.8616528h_{t-1}.$$

It can be noted that the aggregate value of $\alpha_1(0.1209623)$ and $\beta_1(0.8616528)$ approaches unity, which strikingly emphasizes the tenacity of volatility and the significant clustering effect within the financial time series. This observation highlights the persistent nature of market turbulence, where periods of high volatility tend to be succeeded by further volatility. The data captures the essence of volatility clustering in financial data.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

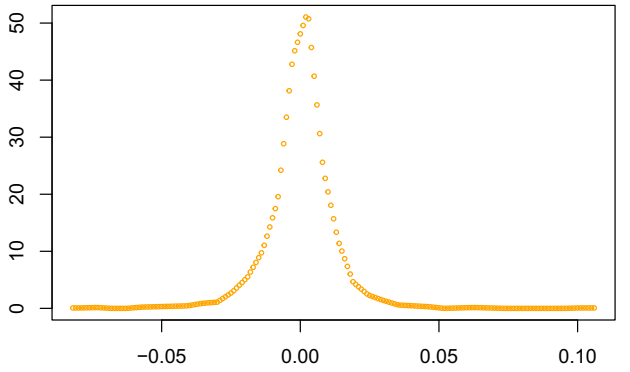


Figure 6. *The estimated density function of daily return rate of DJIA.*

5. Lemmas

Firstly, we make some notations.

Denote $a_{n,j;1} = \frac{1}{2} + j - \frac{x}{b_n}$, $a_{n,j;2} = \frac{1}{2} - j + \frac{x}{b_n}$ and $U_{n,j} = [(j-1/2)b_n, (j+1/2)b_n), j = 0, \pm 1, \pm 2, \cdots$. Then the frequency polygon estimator (1.2) may be rewritten as

$$f_n(x) = a_{n,j;1} f_{n,j} + a_{n,j;2} f_{n,j+1} \text{ for } x \in U_{n,j}. \tag{5.1}$$

Furthermore, let

$$Y_{n,j;i} = I((j-1)b_n \le X_i < jb_n), Z_{n,j;i} = Y_{n,j;i} - EY_{n,j;i},\ i \ge 1. \tag{5.2}$$

Then we have $\upsilon_{n,j} = \sum_{i=1}^{n} Y_{n,j;i}$ and

$$f_n(x) = a_{n,j;1} \frac{\sum_{i=1}^{n} Y_{n,j;i}}{nb_n} + a_{n,j;2} \frac{\sum_{i=1}^{n} Y_{n,j+1;i}}{nb_n}. \tag{5.3}$$

It is clear that the weights $a_{n,j;1}$ and $a_{n,j;2}$ satisfy

$$0 < a_{n,j;1}, a_{n,j;2} < 1 \text{ and } a_{n,j;1} + a_{n,j;2} = 1. \tag{5.4}$$

Denote $p_{nj} := EY_{n,j;i} = P((j-1)b_n \le X_i < jb_n)$. It is also clear that $EZ_{nj;i}^2 = p_{nj}(1-p_{nj})$.

This section provides several important lemmas, which play vital roles and will be used in the subsequent proofs of the main results. The first one is the moment inequality on estimating the upper bound of 2-nd moment for the partial sums of random variables.

Lemma 5.1. (cf. Yang, 2017, Lemma 3.3) *Suppose that* $(A1) - -(A3)$ *hold, and the integers a and m satisfy* $0 \le a < a+m \le n$. *Then there exists a positive constant C, which does not depend on* n, j, a *and m, such that*

$$E\left(\sum_{i=a+1}^{a+m} Z_{nj;i}\right)^2 \le Cmb_n.$$

The following one is the probability inequality for α*-mixing random variables, which is an important tool in the later proofs.*

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Lemma 5.2. (cf. Wei et al., [9], Lemma 3.3) *Let $\{\xi_i, i \geq 1\}$ be a sequence of real-valued α-mixing random variables with $E\xi_i = 0$ and $|\xi_i| \leq b < \infty$ a.s., $\{k_n, n \geq 1\}$ be positive integers such that $1 \leq k_n \leq n/2$. Then for any $\epsilon > 0$,*

$$P\left(\left|\sum_{i=1}^{n} \xi_i\right| > n\epsilon\right) \leq 4\exp\left(-\frac{n\epsilon^2}{12(3\sigma_n^2 + bk_n\epsilon)}\right) + 36b\epsilon^{-1}\alpha(k_n),$$

where $\sigma_n^2 = n^{-1}\sum_{j=1}^{2m_n+1} EV_j^2$, $m_n = \lfloor n/(2k_n)\rfloor$, $V_j = \sum_{(j-1)k_n<i\leq jk_n} \xi_i$ $(j = 1, 2, \cdots, 2m_n)$ and $V_{2m_n+1} = \sum_{i=2m_nk_n+1}^{n} \xi_i$.

In order to clarify the proofs of the main results, we provide the following two lemmas. The first one is inspired by [12].

Lemma 5.3. *If $f(x)$ is uniformly continuous in $\mathbb{R}$, then*

$$\sup_{x\in\mathbb{R}} |Ef_n(x) - f(x)| = o(1). \tag{5.5}$$

If $f(x)$ satisfies the Lipschitz condition of order 1, i.e., the condition $(A4)$, then

$$\sup_{x\in\mathbb{R}} |Ef_n(x) - f(x)| = O(b_n). \tag{5.6}$$

Proof. It is easy to see that

$$\sup_{x\in\mathbb{R}} |Ef_n(x) - f(x)| = \sup_{-\infty<j<\infty} \sup_{x\in U_{nj}} |Ef_n(x) - f(x)|.$$

Note that for $x \in U_{nj}$,

$$|Ef_n(x) - f(x)| \leq \frac{1}{b_n}\int_{(j-1)b_n}^{jb_n} |f(y) - f(x)|\, dy \text{ and } b_n \to 0.$$

If $f(x)$ is uniformly continuous in $\mathbb{R}$, then for $\epsilon > 0$, as n sufficiently large,

$$\int_{(j-1)b_n}^{jb_n} |f(y) - f(x)|\, dy \leq \epsilon b_n.$$

Hence, the desired result (5.5) is derived.

If $f(x)$ satisfies the condition $(A4)$, then as n sufficiently large, we have

$$\int_{(j-1)b_n}^{jb_n} |f(y) - f(x)|\, dy \leq Cb_n^2.$$

Hence, the desired result (5.6) is derived. □

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Lemma 5.4. *Suppose that there exists some* $T > 0$ *such that* $E|X|^{1/T} < \infty$. *Denote* $T_n = (-\infty, -n^T) \bigcup (n^T, \infty)$ *and* $R_n = \max\left\{b_n, \left(\frac{\log n}{nb_n}\right)^{1/2}\right\}$. *Then we have*

$$\sup_{x \in T_n} f_n(x) = O_P(R_n). \tag{5.7}$$

Furthermore, suppose that the density function $f(x)$ *is continuous. Then we have*

$$\sup_{x \in T_n} f(x) = O\left(n^{-(1+T)}\right). \tag{5.8}$$

Proof. In fact, if we want to get (5.7), we only need to prove that for every $\epsilon > 0$, there exist M_ϵ and N_ϵ such that for all $n > N_\epsilon$

$$P\left(\frac{\sup_{x \in T_n} f_n(x)}{R_n} > M_\epsilon\right) < \epsilon. \tag{5.9}$$

Note that

$$\sum_{k=1}^{\infty} kP\left(k^T < |X| \le (k+1)^T\right) \le CE|X|^{1/T} < \infty. \tag{5.10}$$

We can obtain by (5.10) that for sufficiently large n

$$\begin{aligned} P\left(\frac{\sup_{x \in T_n} f_n(x)}{R_n} > M_\epsilon\right) &\le P\left(\sup_{x \in T_n} f_n(x) > 0\right) \\ &\le nP(X \in T_n) = nP\left(|X| > n^T\right) \\ &= \sum_{k=n}^{\infty} nP\left(k^T < |X| \le (k+1)^T\right) \\ &\le \sum_{k=n}^{\infty} kP\left(k^T < |X| \le (k+1)^T\right) < \epsilon. \end{aligned} \tag{5.11}$$

Since $E|X|^{1/T} < \infty$ and the density function $f(x)$ is continuous, we obtain $f(x) = o\left(|x|^{-(1/T+1)}\right)$ as $|x| \to \infty$. Therefore, for sufficiently large n, we have

$$\begin{aligned} \sup_{x \in T_n} f(x) &= \sup_{|x| > n^T} f(x) \le \sup_{|x| > n^T} \left(\frac{|x|}{n^T}\right)^{1/T+1} f(x) \\ &= n^{-(1+T)} \sup_{|x| > n^T} |x|^{1/T+1} f(x) \le Cn^{-(1+T)}. \end{aligned} \tag{5.12}$$

Hence, the proof of Lemma 5.4 is completed. □

6. Proofs of main results

6.1. Proof of Theorem 2.2

Recall that the partition $\cdots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 \cdots$ of the real line $\mathbb{R}$ and equal intervals $I_{n,j} := [(j-1)b_n, jb_n)$ of length b_n, where b_n is the binwidth and $j = 0, \pm 1, \pm 2, \cdots,$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

$U_{nj} = [(j-1/2)b_n, (j+1/2)b_n)$. From Lemma 5.3, we only need to prove that for each j and any $x \in U_{nj}$,

$$|f_n(x) - Ef_n(x)| = o(1) \text{ in probability.} \tag{6.1}$$

Noting that $EZ_{nj;i} = 0$ and $|Z_{nj;i}| \le 1$, we get by Lemma 5.1 that

$$EV_j^2 = E\left(\sum_{(j-1)k_n < i \le jk_n} Z_{nj;i}\right)^2 \le Ck_nb_n,$$

and thus

$$\sigma_n^2 = n^{-1}\sum_{j=1}^{2m_n+1} EV_j^2 \le Cn^{-1}k_n(2m_n+1)b_n \le Cb_n.$$

Hence, we have that for any $\varepsilon > 0$ and $x \in U_{nj}$

$$\begin{aligned}
&P\left(|f_n(x) - Ef_n(x)| > \varepsilon\right)\\
&= P\left(\left|a_{n,j;1}\sum_{i=1}^{n}\left(Y_{n,j;i} - EY_{n,j;i}\right) + a_{n,j;2}\sum_{i=1}^{n}\left(Y_{n,j+1;i} - EY_{n,j+1;i}\right)\right| > nb_n\varepsilon\right)\\
&= P\left(\left|a_{n,j;1}\sum_{i=1}^{n}Z_{n,j;i} + a_{n,j;2}\sum_{i=1}^{n}Z_{n,j+1;i}\right| > nb_n\varepsilon\right)\\
&\le P\left(\left|\sum_{i=1}^{n}Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;1}}\right) + P\left(\left|\sum_{i=1}^{n}Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;2}}\right).
\end{aligned} \tag{6.2}$$

Note by Lemma 5.2 that

$$\begin{aligned}
P\left(\left|\sum_{i=1}^{n}Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;1}}\right) &\le P\left(\left|\sum_{i=1}^{n}Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2}\right)\\
&\le 4\exp\left\{-\frac{nb_n^2\varepsilon^2}{24(6\sigma_n^2 + k_nb_n\varepsilon)}\right\} + 72\varepsilon^{-1}b_n^{-1}\alpha(k_n)\\
&\le 4\exp\left\{-\frac{nb_n\varepsilon^2}{24(C + k_n\varepsilon)}\right\} + 72\varepsilon^{-1}b_n^{-1}\alpha(k_n)\\
&\le C\exp\left\{-\frac{\varepsilon^2}{24\left(\frac{C}{nb_n} + \frac{k_n}{nb_n}\varepsilon\right)}\right\} + Cb_n^{-1}k_n^{-\lambda}.
\end{aligned} \tag{6.3}$$

Taking $k_n = \frac{nb_n\varepsilon}{48\log n}$, we can get by (2.1) that the RHS of above inequality is less than

$$\begin{aligned}
C\exp\left\{-\frac{\varepsilon^2\log n}{24\left(\frac{\varepsilon^2}{48} + \frac{\varepsilon^2}{48}\right)}\right\} + Cb_n^{-1}\left(\frac{nb_n}{\log n}\right)^{-\lambda} &\le C\exp\{-\log n\} + Cb_n^{-(1+\lambda)}n^{-\lambda}(\log n)^{\lambda}\\
&\le Cn^{-1} + Cb_n^{-(1+\lambda)}n^{-\lambda}(\log n)^{\lambda}\\
&= Cn^{-1} + Cg(n,1) = o(1).
\end{aligned} \tag{6.4}$$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Similarly, we have

$$P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;2}}\right) = o(1), \tag{6.5}$$

which together with (6.2)–(6.4) yields (6.1).

Therefore, the proof of Theorem 2.2 is completed.

6.2. *Proof of Theorem 2.4*

Firstly, we prove (2.4). From Lemma 5.3, we only need to show that

$$\sup_{x\in D} |f_n(x) - Ef_n(x)| = o(1) \text{ in probability.} \tag{6.6}$$

Since D is a compact subset of $\mathbb{R}$, we may denote $D = [-d, d]$ for some constant $d > 0$. Let $r_n = \lfloor d/b_n \rfloor + 1$. Since $U_{nj} = [(j - 1/2)b_n, (j + 1/2)b_n)$, we have $D \subseteq \bigcup_{j=-r_n}^{r_n} U_{nj}$.

Similar to the proof of Theorem 2.2, we obtain that $EV_j^2 \le Ck_nb_n$ and $\sigma_n^2 \le Cb_n$, and thus for any $\varepsilon > 0$,

$$\begin{aligned}
&P\left(\sup_{x\in D} |f_n(x) - Ef_n(x)| > \varepsilon\right)\\
&\le P\left(\max_{-r_n\le j\le r_n} \sup_{x\in U_{nj}} |f_n(x) - Ef_n(x)| > \varepsilon\right)\\
&\le \sum_{j=-r_n}^{r_n} P\left(\sup_{x\in U_{nj}} |f_n(x) - Ef_n(x)| > \varepsilon\right)\\
&= \sum_{j=-r_n}^{r_n} P\left(\sup_{x\in U_{nj}} \left|a_{n,j;1}\sum_{i=1}^{n} Z_{n,j;i} + a_{n,j;2}\sum_{i=1}^{n} Z_{n,j+1;i}\right| > nb_n\varepsilon\right)\\
&\le \sum_{j=-r_n}^{r_n} P\left(\sup_{x\in U_{nj}} \left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;1}}\right) + \sum_{j=-r_n}^{r_n} P\left(\sup_{x\in U_{nj}} \left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2a_{n,j;2}}\right)\\
&\le \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2}\right) + \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2}\right)
\end{aligned} \tag{6.7}$$

and

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{nj;i}\right| > \frac{nb_n\varepsilon}{2}\right) \le C\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{\varepsilon^2}{24\left(\frac{C}{nb_n} + \frac{k_n}{nb_n}\varepsilon\right)}\right\} + Cb_n^{-1}k_n^{-\lambda}\right\}. \tag{6.8}$$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

It is clear that $r_n = \lfloor d/b_n \rfloor + 1 \le Cb_n^{-1}$ and $nb_n \to \infty$ as $n \to \infty$. Taking $k_n = \frac{nb_n\varepsilon}{48 \log n}$, we can obtain by (2.3) that the RHS of (6.8) is less than

$$\begin{aligned} Cr_n \exp\left\{-\frac{\varepsilon^2 \log n}{24\left(\frac{\varepsilon^2}{48} + \frac{\varepsilon^2}{48}\right)}\right\} + Cr_n b_n^{-1} k_n^{-\lambda} &\le Cr_n \exp\{-\log n\} + Cr_n b_n^{-1} k_n^{-\lambda} \\ &\le Cr_n n^{-1} + Cr_n b_n^{-1} (nb_n)^{-\lambda} (\log n)^{\lambda} \\ &\le Cb_n^{-1} n^{-1} + Cg(n,2) = o(1). \end{aligned} \tag{6.9}$$

Similarly, we have

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2}\right) = o(1), \tag{6.10}$$

which together with (6.7)–(6.9) yields (6.6).

Next, we prove (2.5). Recalling that $T_n = (-\infty, -n^T) \bigcup (n^T, \infty)$, we have

$$\begin{aligned} \sup_{x\in\mathbb{R}} |f_n(x) - f(x)| \le &\sup_{x\in[-n^T,n^T]} |f_n(x) - Ef_n(x)| + \sup_{x\in[-n^T,n^T]} |Ef_n(x) - f(x)| \\ &+ \sup_{x\in T_n} |f_n(x) - f(x)|. \end{aligned} \tag{6.11}$$

Lemma 5.3 yields that

$$\sup_{x\in[-n^T,n^T]} |Ef_n(x) - f(x)| = O(b_n) = o(1). \tag{6.12}$$

It follows from Lemma 5.4 that

$$\begin{aligned} \sup_{x\in T_n} |f_n(x) - f(x)| &\le \sup_{x\in T_n} f_n(x) + \sup_{x\in T_n} f(x) = O_P(R_n) + O\left(n^{-(1+T)}\right) \\ &= o(1), \text{ in probability.} \end{aligned} \tag{6.13}$$

So if we want to prove

$$\lim_{n\to\infty} P\left(\sup_{x\in\mathbb{R}} |f_n(x) - Ef_n(x)| > 3\varepsilon\right) = 0, \tag{6.14}$$

it suffices to prove

$$\lim_{n\to\infty} P\left(\sup_{x\in[-n^T,n^T]} |f_n(x) - Ef_n(x)| > \varepsilon\right) = 0. \tag{6.15}$$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Let $D = [-n^T, n^T]$ and $r_n = \lfloor n^T/b_n \rfloor + 1 \le Cn^T b_n^{-1}$. Similar to (6.7)–(6.8), we obtain

$$\begin{aligned} &P\left(\sup_{x\in[-n^T,n^T]} |f_n(x) - Ef_n(x)| > \varepsilon\right) \\ &\le P\left(\max_{-r_n\le j\le r_n} \sup_{x\in U_{nj}} |f_n(x) - Ef_n(x)| > \varepsilon\right) \\ &\le \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2}\right) + \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2}\right) \end{aligned} \tag{6.16}$$

and

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n\varepsilon}{2}\right) \le C\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{\varepsilon^2}{24\left(\frac{C}{nb_n} + \frac{k_n}{nb_n}\varepsilon\right)}\right\} + Cb_n^{-1}k_n^{-\lambda}\right\}. \tag{6.17}$$

Taking $k_n = \frac{nb_n\varepsilon}{48(T+1)\log n}$, the RHS of (6.17) is less than

$$\begin{aligned} Cr_n \exp\left\{-\frac{\varepsilon^2(T+1)\log n}{24\left(\frac{\varepsilon^2}{48} + \frac{\varepsilon^2}{48}\right)}\right\} + Cr_n b_n^{-1}k_n^{-\lambda} &\le Cr_n \exp\{-(T+1)\log n\} + Cr_n b_n^{-1}k_n^{-\lambda} \\ &\le Cn^T b_n^{-1} n^{-(T+1)} + Cn^T b_n^{-(2+\lambda)} n^{-\lambda} (\log n)^{\lambda} \\ &\le Cn^T b_n^{-1} n^{-(T+1)} + Cn^T g(n,2) = o(1). \end{aligned} \tag{6.18}$$

Similarly, we have

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n\varepsilon}{2}\right) = o(1), \tag{6.19}$$

which together with (6.16)–(6.18) yields (6.15).

Therefore, the proof of Theorem 2.4 is completed.

6.3. Proof of Theorem 2.7

Firstly, we prove (2.7). From Lemma 5.3, it suffices to show that

$$\sup_{x\in D} |f_n(x) - Ef_n(x)| = O(R_n) \text{ in probability.} \tag{6.20}$$

Recall that we denoted $D = [-d, d]$ for some constant $d > 0$ and set $r_n = \lfloor d/b_n \rfloor + 1$. It is easy to see that $r_n = \lfloor d/b_n \rfloor + 1 \le Cb_n^{-1}$ and $nb_n/\log n \to \infty$ as $n \to \infty$. Let C_0 be a positive constant, which will be specified later. Similar to (6.8), taking $k_n = C_0^{-1}C_1\left(\frac{nb_n}{\log n}\right)^{1/2}$ and $C_0 = \sqrt{24 \times 7C_1}$, we obtain by

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

(6.8) and (2.6) that

$$P\left(\sup_{x\in D}|f_n(x)-Ef_n(x)|>C_0R_n\right)\le P\left(\max_{-r_n\le j\le r_n}\sup_{x\in U_{nj}}|f_n(x)-Ef_n(x)|>C_0R_n\right)$$
$$\le\sum_{j=-r_n}^{r_n}P\left(\left|\sum_{i=1}^{n}Z_{n,j;i}\right|>\frac{nb_nC_0R_n}{2}\right)+\sum_{j=-r_n}^{r_n}P\left(\left|\sum_{i=1}^{n}Z_{n,j+1;i}\right|>\frac{nb_nC_0R_n}{2}\right),\tag{6.21}$$

and the former of the RHS of (6.21) is less than

$$4\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{nb_n^2C_0^2R_n^2}{24\left(6\sigma_n^2+k_nb_nC_0R_n\right)}\right\}+72C_0^{-1}R_n^{-1}b_n^{-1}k_n^{-\lambda}\right\}$$
$$\le 4\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{nb_nC_0^2R_n^2}{24\left(6C_1+k_nC_0R_n\right)}\right\}+CR_n^{-1}b_n^{-1}k_n^{-\lambda}\right\}$$
$$\le 4\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{C_0^2R_n^2\log n}{24\left[6C_1\frac{\log n}{nb_n}+C_1\left(\frac{\log n}{nb_n}\right)^{1/2}R_n\right]}\right\}+CR_n^{-1}b_n^{-1}k_n^{-\lambda}\right\}$$
$$=4\sum_{j=-r_n}^{r_n}\left\{\exp\left\{-\frac{C_0^2R_n^2\log n}{24\left[6C_1R_n^2+C_1R_n^2\right]}\right\}+CR_n^{-1}b_n^{-1}k_n^{-\lambda}\right\}$$
$$\le 4\sum_{j=-r_n}^{r_n}\left\{\exp\{-\log n\}+C\left(\frac{\log n}{nb_n}\right)^{-1/2}b_n^{-1}\left(\frac{nb_n}{\log n}\right)^{-\lambda/2}\right\}$$
$$\le Cn^{-1}b_n^{-1}+Cb_n^{-2}\left(\frac{nb_n}{\log n}\right)^{-\lambda/2+1/2}=Cn^{-1}b_n^{-1}+Cg(n,3)=o(1).\tag{6.22}$$

Similarly, the latter of the RHS of (6.21) tends to zero as n tends to infinity, that is,

$$\sum_{j=-r_n}^{r_n}P\left(\left|\sum_{i=1}^{n}Z_{n,j+1;i}\right|>\frac{nb_nC_0R_n}{2}\right)=o(1),\tag{6.23}$$

which together with (6.21)–(6.22) yields (6.20).

Next, we prove (2.8). Recall that $T_n=(-\infty,-n^T)\bigcup(n^T,\infty)$ and

$$\sup_{x\in\mathbb{R}}|f_n(x)-f(x)|\le\sup_{x\in[-n^T,n^T]}|f_n(x)-Ef_n(x)|+\sup_{x\in[-n^T,n^T]}|Ef_n(x)-f(x)|$$
$$+\sup_{x\in T_n}|f_n(x)-f(x)|.\tag{6.24}$$

Similar to (6.12)–(6.13) in the proof of Theorem 2.4(2), if we want to prove

$$\lim_{n\to\infty}P\left(\sup_{x\in\mathbb{R}}|f_n(x)-Ef_n(x)|>3C_0R_n\right)=0,\tag{6.25}$$

it suffices to prove

$$\lim_{n\to\infty}P\left(\sup_{x\in[-n^T,n^T]}|f_n(x)-Ef_n(x)|>C_0R_n\right)=0.\tag{6.26}$$

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

Let $D = [-n^T, n^T]$ and $r_n = \lfloor n^T/b_n \rfloor + 1 \le Cn^T b_n^{-1}$. Similar to (6.21)–(6.22), we obtain by taking $k_n = C_0^{-1} C_1 \left(\frac{nb_n}{\log n}\right)^{1/2}$ and $C_0 = \sqrt{24 \times 7(T+1)C_1}$ that

$$P\left(\sup_{x\in[-n^T,n^T]} |f_n(x) - Ef_n(x)| > C_0 R_n\right)$$
$$\le \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n C_0 R_n}{2}\right) + \sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n C_0 R_n}{2}\right), \tag{6.27}$$

and

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j;i}\right| > \frac{nb_n C_0 R_n}{2}\right) \le 4 \sum_{j=-r_n}^{r_n} \left\{\exp\left\{-\frac{C_0^2 R_n^2 \log n}{24\left[6C_1 R_n^2 + C_1 R_n^2\right]}\right\} + CR_n^{-1} b_n^{-1} k_n^{-\lambda}\right\}$$
$$\le 4 \sum_{j=-r_n}^{r_n} \left\{\exp\{-(T+1)\log n\} + Cb_n^{-1}\left(\frac{nb_n}{\log n}\right)^{-\lambda/2+1/2}\right\}$$
$$\le Cn^T b_n^{-1} n^{-(T+1)} + Cn^T g(n,3) = o(1). \tag{6.28}$$

Analogously, we have

$$\sum_{j=-r_n}^{r_n} P\left(\left|\sum_{i=1}^{n} Z_{n,j+1;i}\right| > \frac{nb_n C_0 R_n}{2}\right) = o(1), \tag{6.29}$$

which together with (6.27)–(6.28) yields (6.26).

Therefore, the proof of Theorem 2.7 is completed.

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

Funding. This work was supported by the National Natural Science Foundation of China (12471248, 12301181, 12361031, 12201600), the Natural Science Youth Project of Anhui Provincial Department of Education (2025AHGXZK40092), and the Chizhou University High-Level Talent Research Start-up Fund (CZ2024YJRC05).

Conflict of interest statement. We declare that we have no conflict of interest.

References

[1] Besaid, N., & Dabo-Niang, S. (2010). Frequency polygons for continuous random fields. *Statistical Inference for Stochastic Processes* 13: 55–80.
[2] Carbon, M., Francq, C., & Tran, L.T. (2010). Asymptotic normality of frequency polygons for random fields. *Journal of Statistical Planning and Inference* 140(2): 502–514.
[3] Carbon, M., Garel, B., & Tran, L.T. (1997). Frequency polygons for weakly dependent processes. *Statistics and Probability Letters* 33 1–13.
[4] Fan, J.Q., & Yao, Q.W. (2003). *Nonlinear Time series: Nonparametric and Parametric methods*. New York: Springer-Verlag.
[5] Huang, G.J., & Xing, G.D. (2017). On the Berry-Esséen bound of frequency polygons for ϕ-mixing samples. *Journal of Inequalities and Applications* 2017, Article ID 65, 11 pages.
[6] Kang, S.L., Abbasi, R., Wang, H.L., Xu, L.X., Wang, X.F., & Tang, Y.Y. (2018). Asymptotic behaviour of frequency polygons under ϕ-mixing samples. *Communications in Statistics-Theory and Methods* 47(21): 5369–5377.
[7] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*. 42: 43–47.
[8] Scott, D.W. (1985). Frequency polygons: theory and application. *Journal of the American Statistical Association* 80(390): 348–354.

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press

[9] Wei, X.L., Yang, S.C., Yu, K.M., Yang, X., & Xing, G.D. (2010). Bahadur representation of linear kernel quantile estimator of VaR under α-mixing assumptions. *Journal of Statistical Planning and Inference* 140(7): 1620–1634.
[10] Wen, Y.Z., & Xing, G.D. (2017). On the uniformly asymptotic normality of frequency polygons for ψ-mixing samples. *Communications in Statistics-Theory and Methods* 46(21): 10416–10425.
[11] Wu, Y., & Wang, X.J. (2022). Uniformly complete consistency of frequency polygon estimation for dependent samples and an application. *Statistics* 56(6): 1270–1289.
[12] Yang, X. (2017). Uniform strong consistency of histogram density estimation for dependent process. *Communications in Statistics-Theory and Methods* 46(4): 1709–1719.
[13] Younso, A. (2023). Consistency of the frequency polygon estimators of density mode for strongly mixing processes. *Communications in Statistics-Theory and Methods* 52(7): 2182–2197.

Cite this article: Wang Y., Huang H. and Shen A. (2026). Asymptotic behavior of frequency polygon estimation of density function for dependent samples. *Probability in the Engineering and Informational Sciences* 40(4): 598–617. https://doi.org/10.1017/S0269964826100254

https://doi.org/10.1017/S0269964826100254 Published online by Cambridge University Press