Appl. Math. Inf. Sci. **7**, No. 1L, 5-9 (2013)

Applied Mathematics & Information Sciences
An International Journal

Uniformly Consistency of the Cauchy-Transformation Kernel Density Estimation Underlying Strong Mixing

Wann-Jyi Horng

Department of Hospital and Health Care Administration, Chia Nan University of Pharmacy & Science, No. 60, Erh-Jen RD., Sec.1, Jen-Te, Tainan, 71710, Taiwan

Received: 7 Sep. 2012, Revised: 17 Oct. 2012, Accepted: 28 Dec. 2012
Published online: 1 Feb. 2013

Abstract: In this paper, one uses the idea of Cauchy-transformation to construct a Cauchy-transformation kernel density estimator underlying the condition of strong mixing. The uniformly strong consistency and convergence rates of the proposed estimator are obtained underlying the papers of Cai and Roussas [1] and Kim and Lee [6]. The proposed estimator can improve the boundary effects of the empirical (or uniform)-transformation kernel density estimator in the boundary area. Besides, the proposed estimator can also be applied to estimate the hazard function.

Keywords: kernel density estimator, Cauchy-transformation, empirical distribution, convergence rate, strong mixing.

1 Introduction

Let X_1,...,X_n be a strong mixing (α-mixing) sequence of random variables. Suppose that the X_1 have a distribution function F(x) and probability density function $f(x)$. Given the mixing coefficientα(n) and α(n)$\rightarrow$ 0 as n$\rightarrow \infty$. The α-mixing coefficients of the sequence $\{X_k\}$ of random variables is defined by
$\alpha(n) = \sup_{k=1,2,....} \sup \{|P(A\bigcap B) - P(A)P(B)| : A \in \Im_1^k, B\Im_{n+k}^\infty\}$
where$\Im_m^n = \sigma(X_t : m \leq t < n)$, $-\infty \leq m < n \leq \infty$, here$\sigma$is the$\sigma$algebra. The strong mixing conditions can also refer to the paper of Bradley[17].

In the literature, various density estimators for $f(x)$ underlying the condition of strong mixing, based on a random sample $X_1,\ldots,X_n$ have been proposed and their properties are studied, for examples, the papers of Schuster [13], Roussas [12], Cai and Roussas [1], Kim and Cox [4], Liebscher [7], Tae and Cox [15], Kim and Cox [5], Kim and Lee [6] and Hansen [16]. A kernel density estimator for f(x) is defined as follows:

$$\hat{f}_n(x;h) = \frac{1}{nh}\sum_{j=1}^{n} K\left\{\frac{x-X_j}{h}\right\}, \qquad (1.1)$$

where K is a (symmetric) kernel function and h=$h(n)$ is called the bandwidth. Estimator (1.1) is called the kernel density estimator (KDE) of Parzen [9] and Rosenblatt [10].

Ruppert and Cline [11] propose a empirical-transformation KDE which it is defined as follows:

$$\hat{f}_{RC}(x;h) = \frac{1}{nh}\sum_{j=1}^{n} K\left\{\frac{F_n(x)-F_n(X_j)}{h}\right\} * \hat{f}_n(x;h), \qquad (1.2)$$

where$F_n(x)$is the empirical distribution function of the data to estimate F(x) and $\hat{f}_n(x;h)$ is defined as above. When point F(x) nears the boundary regions of [0,1], the empirical-transformation kernel density estimator (Ruppert and Cline [11]) will be suffered to the problem of boundary effects, for example, the expectation of $\hat{f}_{RC}(x;h)$is not equal f(x), as F(x)$\in$ $[0,ch), 0 \leq c < 1$. Some boundary modification methods have been proposed by, for examples, Jones [3] and M$\ddot{u}$ller [8].

In this paper, by the ideas of Cauchy-transformation method and the estimator (1.2), we propose a Cauchy-transformation KDE which it does not have the boundary effects problem of the empirical-transformation KDE. The proposed estimator is defined as follows:

$$\tilde{f}_T(x;h) = \frac{1}{nh}\sum_{j=1}^{n} K\left\{\frac{T(x)-T(X_j)}{h}\right\} * T^{(1)}(x), \qquad (1.3)$$

where$T(x) = \tan[\pi(F(x)-0.5)]$, $F(x)$ is the distribution function of X and$q(\cdot)$ denotes the Cauchy density of

* Corresponding author e-mail: hwj7902@mail.chna.edu.tw

T(x). T(x) is continuous and differentiable at the point x. And q($T(\bullet)$) is continuously differentiable. The estimator (1.3) is called the ideal Cauchy-transformation KDE. Here researcher uses the estimator of $\tilde{f}_T$(x; h) to estimate f(x). We will study the uniformly consistency and convergence rates of estimator (1.3).

When the function $F(x)$ (or $f(x)$) is unknown, the estimator (1.3) will be not a practical estimator, therefore, we need further to estimate $F(x)$. The corrected estimator of (1.3) is given by

$$\hat{f}_T(x;h) \quad = \quad \frac{1}{nh}\sum_{j=1}^{n} K\left\{\frac{\hat{T}(x)-\hat{T}(X_j)}{h}\right\}\hat{T}^{(1)}(x), \qquad (1.4)$$

where $\hat{T}(x) = \tan[\pi(F_n(x) - 0.5)]$, $F_n(x)$is the empirical distribution function of the data.

This paper is organized as follows: In section 2, we state the uniformly consistency of estimator (1.4) and gives the explicit formula for the convergence rate. In section 3, we give the proofs of the asymptotic results.

2 Uniformly Consistency

In this section, one states the main results of estimator (1.4). Before one state the main results one will give the following assumptions:

(A1) X_1, ...,X_n is α-mixing sequence with the probability density function $f(x)$, $f(x)$ is bounded on its domain, and continuous at the point x and $f^{(m)}(x)$ exists, for $m \geq 2$.

(A2)$K(u)$ is bounded variation. $K(\cdot)$ is the kernel function and satisfies

(i) $\lim\sup_{|u|\to\infty} \left| u^{k+2}K(u) \right| < \infty$, for each $k\geq 0$.

(ii) $\int \left|K^{(1)}(u)\right| du < \infty$.

(A3) $\sum_{n=1}^{\infty} \frac{\log n}{n}(\log\log n)^{1+r}\alpha(n)$ converges for some r>0, the condition of Cai and Roussas [1].

Let us now give the main results of the estimator (1.4). In the following Theorem 2.1, we give the uniformly strong convergence of the estimator (1.4) at the point $x\in(-\infty,\infty)$.

Theorem 2.1. Assume that conditions (A1)-(A3) are satisfied. And suppose that $h\to 0$ and $nh\to\infty$, as $n\to\infty$, then we have

$$\sup_{x\in(-\infty,\infty)} \left| \hat{f}_T(x;h) - f(x) \right| \to 0, \qquad (2.1)$$

almost surely, as $n\to\infty$.

Remark 1: The estimator (1.4) has the strong convergence property as that of the traditional kernel density estimator. The proposed estimator does not have the boundary effects problem of the empirical-transformation KDE, the kernel function does not need to use the boundary kernel function, the boundary problem can also refer to the paper of M$\ddot{u}$ller [8].

Remark 2: As the estimator of F(x) is

$$\hat{F}(x;h) = \int_{-\infty}^{x} \hat{f}_n(t;h)dt,$$

here$\hat{f}_n(x;h)$is the KDE of Parzen (1962), the asymptotic result of Theorem 2.1 is also held.

If $f^{(2)}(x)$is uniformly continuous functions on (-∞,∞), then, we have the following theorem about the convergence rate of the estimator$\hat{f}_T(x;h)$.

Theorem 2.2. Under the assumptions of Theorem 2.1 and assume that $f^{(2)}(x)$is uniformly continuous, then we have

$$\sup_{x\in(-\infty,\infty)} \left| \hat{f}_T(x;h) - f(x) \right| \quad = \quad O\left(h^2 + \frac{\sqrt{log\log n}}{\sqrt{nh}} \right) \qquad (2.2)$$

Remark 3: From Theorem 2.2, the convergence rate is constructed for the estimator (1.4). The choice of the smoothing parameter h can refer to, for example, the book of Hrdle [2].

An application of Theorem 2.1-2, one can use the proposed estimator to estimate the hazard function. The hazard function is defined as follows:

$$H(x) = \frac{f(x)}{1-F(x)} = \frac{f(x)}{\bar{F}(x)}, \qquad (2.3)$$

where $\bar{F}(x) = 1 - F(x)$.

The estimator of H(x) is defined as follows:

$$\hat{H}(x;h) = \frac{\hat{f}_T(x;h)}{1-F_n(x)} = \frac{\hat{f}_T(x;h)}{\bar{F}_n(x)}, \qquad (2.4)$$

where $\hat{f}_T(x;h)$and$F_n(x)$are defined as above.

Let us now give the asymptotic results of the estimator (2.4). In the following Theorem 2.3-4, one also provides the uniformly strong consistency, as$x \in (-\infty,\infty)$.

Theorem 2.3. Assume that conditions (A1)-(A3) are satisfied. And suppose that $h\to 0$ and $nh\to\infty$, as $n\to\infty$, then we have

$$\sup_{x\in(-\infty,\infty)} \left| \hat{H}(x;h) - H(x) \right| \to 0, \qquad (2.5)$$

almost surely, as $n\to\infty$.

Theorem 2.4. Under the assumptions of Theorem 2.3 and assume that $f^{(2)}(x)$is uniformly continuous, then we have

$$\sup_{x\in(-\infty,\infty)} \left| \hat{H}(x;h) - H(x) \right| \quad = \quad O\left(h^2 + \frac{\sqrt{log\log n}}{\sqrt{nh}} \right) \qquad (2.6)$$

3 Proofs

In this section our main purpose is to prove Theorem 2.1-4. The four lemmas below will be used to justify the theorems.

Lemma 3.1. Let $X_1,\ldots,X_n$ be a stationary α-mixing sequence of real-valued random variables with distribution function F(x) and mixing coefficient α(n) satisfying (A3), and let $F_n(x)$ be the empirical distribution function based on the segment X_1,...,X_n. Then

$$\sup_x |F_n(x) - F(x)| \to 0, \tag{3.1}$$

almost surely, as $n\to\infty$.

Proof: The proof follows from the Theorem 1 in the book of Tucker [14] and the Corollary 2.1 in the paper of Cai and Roussas [1].

Lemma 3.2. Let $X_1,\ldots,X_n$ be a stationary α-mixing sequence of real-valued random variables with distribution function F(x) and mixing coefficient α(n) =O($n^{-\delta}$),for some $\delta > 3$, and let $F_n(x)$ be the empirical distribution function based on the segment $X_1,\ldots,X_n$. And let F(x) satisfies the Lipschitz condition $|F(x)-F(y)| \le b_0\,|x-y|$, for some constant b_0. Then

$$\mathrm{P}\{\limsup_x[(\tfrac{n}{2\log\log n})^{1/2}\,|F_n(x)-F(x)|] = \tfrac{1}{2}\}=1, \tag{3.2}$$

almost surely, as $n\to\infty$.

Proof: The proof follows from the Theorem 3.2 in the paper of Cai and Roussas [1].

Lemma 3.3. Assume that the conditions (A1)-(A2) are satisfied. Then we have

$$E\,\tilde{f}_T(x;h) - f(x) \;=\; O(\,h^2\,), \tag{3.3}$$

for $x\in$(-∞,∞).

Proof: By the calculation of expectation, as $x\in$(-∞,∞), we have

$$\begin{aligned}
&E\,\tilde{f}_T(x;h)\\
&= \int\ \tfrac{1}{h}\,K\left\{\tfrac{T(x)-T(y)}{h}\right\} f(y)T^{(1)}(x)dy\\
&= \int\ \tfrac{1}{h}\,K\left\{\tfrac{T(x)-T(y)}{h}\right\} q(T(y))T^{(1)}(y)dyT^{(1)}(x)\\
&= \int \tfrac{1}{h}\,K\left\{\tfrac{T(x)-u}{h}\right\} q(u)duT^{(1)}(x)\\
&= \int K\{w\}\,q(T(x)-wh)dwT^{(1)}(x)\\
&= q(T(x))T^{(1)}(x) + \tfrac{\mu_2}{2}h^2q^{(2)}(u)|_{u=T(x)}T^{(1)}(x) + o(h^2)\\
&= f(x) + O(h^2),
\end{aligned}$$

from techniques of Calculus underlying the conditions (A1)-(A2) and q(T(•)) is continuously differentiable. Therefore, the proof of Lemma 3.3 is proved.

Lemma 3.4. Under the conditions (A1)-(A2), then we have

$$\sup_x \left|\ \tilde{f}_T(x;h) - E\tilde{f}_T(x;h)\ \right|\ ,=\ O(\,\frac{\sqrt{log\log n}}{\sqrt{n}h}\,) \tag{3.4}$$

almost surely, for all x on real line.

Proof: We know that

$$\sup_x \left|\ \tilde{f}_T(x;h) - E\tilde{f}_T(x;h)\ \right|\ =$$

$$\sup_x \left|\ \tfrac{1}{h}\int K\left\{\tfrac{T(x)-T(y)}{h}\right\}T^{(1)}(x)dF_n(y) - \tfrac{1}{h}\int K\left\{\tfrac{T(x)-T(y)}{h}\right\}T^{(1)}(x)dF(y)\ \right|$$

$$\le \sup_x \tfrac{1}{h}\int |F_n(y)-F(y)|\,dK\left\{\tfrac{T(x)-T(y)}{h}\right\}T^{(1)}(x) \le \tfrac{\nu}{h}\sup_x|F_n(x)-F(x)|, \tag{3.5}$$

by the techniques of calculus and the conditions (A1)and A(2), where ν is the variation of $K\{\bullet\}T^{(1)}(x)$.

From Lemma 3.2 and (3.5), we have

$$\sup_x \left|\ \tilde{f}_T(x;h) - E\tilde{f}_T(x;h)\ \right|\ =\ O(\,\frac{\sqrt{log\log n}}{\sqrt{n}h}\,) \tag{3.6}$$

almost surely, as $n \to \infty$.

Proof of Theorem 2.2. We know that

$$\begin{aligned}
&\sup_x \left|\ \hat{f}_T(x;h) - f(x)\ \right|\\
&\le \sup_x\left|\hat{f}_T(x;h) - \tilde{f}_T(x;h)\right|\\
&\quad + \sup_x\left|\tilde{f}_T(x;h) - f(x)\right|
\end{aligned}$$

=I(1)+I(2),

From I(1), we have

$$\begin{aligned}
&\sup_x \left|\ \hat{f}_T(x;h) - \tilde{f}_T(x;h)\ \right|\\
&\le \sup_x \tfrac{1}{nh}\textstyle\sum_{j=1}^n \left|K\left\{\tfrac{T(x)-T(X_j)}{h}\right\}(\hat{T}^{(1)}(x) - T^{(1)}(x))\right|\\
&+ \sup_x \tfrac{1}{nh}\textstyle\sum_{j=1}^n \left|K\left\{\tfrac{\hat{T}(x)-\hat{T}(X_j)}{h}\right\} - K\left\{\tfrac{T(x)-T(X_j)}{h}\right\}\right|\hat{T}^{(1)}(x)\\
&\le \sup_x\left|\hat{T}^{(1)}(x) - T^{(1)}(x)\right|\frac{1}{nh}\sum_{j=1}^n\left|K\left\{\frac{T(x)-T(X_j)}{h}\right\}\right|\\
&\quad + C_0\sup_x\left|\hat{T}^{(1)}(x) - T^{(1)}(x)\right|\\
&\bullet\frac{1}{nh}\sum_{j=1}^n\left|K^{(1)}\left\{\frac{T(x)-T(X_j)}{h}\right\}(\frac{x-X_j}{h})\right|\hat{T}^{(1)}(x)\\
&\le C_1\sup_x|F_n(x)-F(x)| * \frac{1}{nh}\sum_{j=1}^n\left|K\left\{\frac{T(x)-T(X_j)}{h}\right\}\right|\\
&+C_2\sup_x|F_n(x)-F(x)| * \tfrac{1}{nh}\textstyle\sum_{j=1}^n\left|K^{(1)}\left\{\tfrac{T(x)-T(X_j)}{h}\right\}(\tfrac{x-X_j}{h})\right|\left|\hat{T}^{(1)}(x)\right|
\end{aligned}$$

=I(3)+I(4),

for some constant C_0,$C_1$and C_2.

By I(3), we have

$$\frac{1}{nh}\sum_{j=1}^n\left|K\left\{\frac{T(x)-T(X_j)}{h}\right\}\right|$$

$$\to q(T(x)\int|K(w)|dw < \infty \tag{3.7}$$

By I(4), we have

$$\frac{1}{nh}\sum_{j=1}^{n}\left|K^{(1)}\left\{\frac{T(x)-T(X_j)}{h}\right\}(\frac{x-X_j}{h})\right|\left|\hat{T}^{(1)}(x)\right|$$

$$\to f(x)\int\left|wK^{(1)}(w)\right|dw<\infty, \tag{3.8}$$

where f(x)=q(T(x))$T^{(1)}(x)$.

By (3.7) and (3.8), we have

$$\sup_x\left|\hat{f}_T(x;h)-\tilde{f}_T(x;h)\right|\le C_3\sup_x|F_n(x)-F(x)| \tag{3.9}$$

for some constant C_3.

From I(2), we have

$$\sup_x\left|\tilde{f}_T(x;h)-f(x)\right|$$

$$\le\sup_x\left|\tilde{f}_T(x;h)-E\tilde{f}_T(x;h)\right|+\sup_x\left|E\tilde{f}_T(x;h)-f(x)\right| \tag{3.10}$$

From (3.9), (3.10) and Lemma 3.3-4, the proof of Theorem 2.2 is completed.

Proof of Theorem 2.1. By (3.5) and Lemma 3.1, we have

$$\sup_x\mid\tilde{f}_T(x;h)-E\tilde{f}_T(x;h)\mid\to 0, \tag{3.11}$$

almost surely, as $h\to 0$, $n\to\infty$**.**

From (3.10), (3.11) and Lemma 3.3, the proof of Theorem 2.1 is proved.

Proof of Theorem 2.3. By (2.3) and (2.4), we have

$$\hat{H}(x;h)-H(x)=\frac{1}{\bar{F}(x)\bar{F}_n(x)}\cdot*\left\{\bar{F}(x)[\hat{f}_T(x;h)-f(x)]+f(x)[F_n(x)-F(x)]\right\} \tag{3.12}$$

From (3.12), we obtain

$$\sup_x\mid\hat{H}(x;h)-H(x)\mid$$

$$\le C_4\sup_x\mid\hat{f}_T(x;h)-f(x)\mid+C_5\sup_x\mid F_n(x)-F(x)\mid, \tag{3.13}$$

for some constants $C_4$and C_5.

By (3.13), Lemma 3.1 and Theorem 2.1, the Theorem 2.3 can be proved.

Proof of Theorem 2.4. From Theorem 2.3 and Lemma 3.2**,** the Theorem 2.4 can be proved, the details are omitted.

4 Conclusions

In this paper, one uses the technique of Cauchy-transformation method to construct a Cauchy-transformation kernel density estimator underlying the condition of strong mixing. The uniformly strong consistency and convergence rates of the proposed estimator are obtained underlying the papers of Cai and Roussas [1] and Kim and Lee [6]. The empirical result shows that the proposed estimator can improve the boundary effects of the uniform-transformation kernel density estimator in the boundary area (Ruppert and Cline [11]). Besides, the proposed estimator can also be applied to estimate the hazard function and others density function.

References

[1] Z. Cai and G.G. Roussas, *Uniform strong estimation underα-mixing, with rates, Statistics & Probability Letters* **15**, 47-55 (1992)

[2] W. Hrdle, *Smoothing Techniques with Implementation in S*, Springer-Verlag, New York (1991)

[3] M.C. Jones, *Simple boundary correction for kernel density estimation, Statistics Computing* **3**, 135-146 (1993)

[4] T.Y. Kim and D.D. Cox, *Uniform strong consistency of kernel density estimators under dependence, Statistics & Probability Letters* **26**, 179-185 (1996)

[5] T.Y. Kim and D.D. Cox, *Convergence rates for average square errors in kernel smoothing estimators, Journal Noparametric Statistics* **13**, 209-228 (2001)

[6] T.Y. Kim and S. Lee, *Kernel density estimator for strong mixing processes*, Working paper (2004)

[7] E. Liebscher, *Strong convergence of sums of α-mixing random variables with applications to density estimation, Stochastic Process. Appl.* **65**, 69-80 (1996)

[8] H.G. Müller, *On the boundary kernel method for nonparametric curve estimation near endpoints, Scandinavian Journal of Statistics* **20**, 313-328 (1993)

[9] E. Parzen, *On estimation of a probability density function and mode, Annals Mathematical Statistics* **33**, 1065-1076 (1962)

[10] M. Rosenblatt, *Remarks on some non-parametric estimates of a density function, Annals Mathematical Statistics* **27**, 832-837 (1956)

[11] D. Ruppert and D.B.H. Cline, *Transformation kernel density estimation: bias reduction by empirical transformation, Annals Statistics* **22**, 185-210 (1994)

[12] G.G. Roussas, *Nonparametric estimation in mixing sequences of random variables, J. Statist. Plann. Inference* **18**, 135-149 (1988)

[13] E.F. Schuster, *Estimation of a probability density function and its derivatives, Annals mathematical Statistics* **40**, 1187-1195(1969)

[14] H.G. Tucker, *A Graduate Course in Probability*, Academic Press, StateplaceNew York (1967)

[15] Y.K. Tae and D.D. Cox, *Uniform strong consistency of kernel density estimator under dependence, Statist. & Probab. Lett*. **26**, 179-185 (1996)

[16] B.E. Hansen, *Uniform convergence rates for kernel estimation with dependence data, Econometric Theory* **24**, 726-748 (2008)

[17] R.C. Bradley, *Basic properties of strong mixing conditions: A Survey and some open questions. Probab. Surveys* **2**, 107-144 (2005)

Wann-Jyi Horng received the MS degree in Applied Mathematics from National Cheng Kung University in 1989, and the PhD degree in Statistics science, from the department of Institute Statistics, National Tsing Hua University in 1995. He is currently an associate professor in Chia Nan University. His research interests are in the areas of applied statistics, applied mathematics, and financial data analysis.