DENSITY ESTIMATION FOR UNIFORM MIXING PROCESS

Abdul-al, Khaled I.
Department of Statistics, Colorado State University

Siddiqui, M. M.
Department of Mathematical Sciences, University of Petroleum and Minerals

https://doi.org/10.5109/13379

出版情報：Bulletin of informatics and cybernetics. 22 (1/2), pp.79-93, 1986-03. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics Vol. 22, No. 1~2, 1986

DENSITY ESTIMATION FOR UNIFORM MIXING PROCESS

By

Khaled I. ABDULAL*, ** and M. M. SIDDIQUI*

Abstract

Let $\{x_n\}$ be a stationary uniform sequence of random variables having a probability density function $f(x)$. Based on the first n observations an estimate of $f(x)$ is given by

$$f_n(x) = (na_n)^{-1} \sum_{j=1}^{n} K(a_n^{-1}(x - X_j))$$

where $K(y)$ is a known probability density function. Asymptotic properties of $f_n(x)$ have been studied.

1. Introduction

Suppose that the sequence $\{X_n\}$ is stationary, then X_n's have the same probability density function $f(x)$ and distribution function $F(x)$. Based on the first n observations an estimate of $f(x)$ is given by

$$f_n(x) = \frac{1}{na_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{a_n}\right) \tag{1.1}$$

where $K(y)$ is known p.d.f. satisfying the following conditions

$$\begin{aligned} &\text{(i)} \quad \sup_{-\infty<y<\infty} K(y) < \infty \\ &\text{(ii)} \quad \lim_{|y|\to\infty} |yK(y)| = 0 \end{aligned} \tag{1.2}$$

and $\{a_n\}$ is a sequence of positive real numbers such that

$$\lim_{n\to\infty} a_n = 0. \tag{1.3}$$

Estimate $f_n(x)$ of $f(x)$ based on a sample of independent observations have been considered by many authors, notably we mention Parzen [14], Leadbetter [10], Nadarya [12, 13], Murthy [11], Yamato [20], and Davies [6]. Roussas [17, 18] and Rosenblatt [16] have studied the asymptotic properties of $f_n(x)$ when the observations are assumed to be sampled from a stationary Markov process. The purpose of this paper is to study the asymptotic properties of the estimate $f_n(x)$ under the uniform mixing condition. Results obtained here generalize those of Parzen [14] and Nadarya [13] for the i.i.d. case, and also those of Roussas [17, 18] for stationary Markov sequences.

* Department of Statistics Colorado State University

** Department of Mathematical Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia

In Section 3, asymptotic unbiasedness, consistency, and uniform consistency (weak as well as strong) will be studied for $f_n(x)$. Asymptotic normality of the estimate $f_n(x)$ will be established in section 4.

2. Preliminaries

Let $\{X_n\}$ be a stationary sequence of random variables defined on a probability space (Ω, B, P). For $a \leqq b$ define $\sigma(a, b)$ as the σ-field generated by the random variables $X_a, \cdots, X_b$ and define $\sigma(a, \infty)$ as the σ-field generated by $X_a X_{a+1}, \cdots$.

We shall say that the sequence $\{X_n\}$ is uniform mixing if, for each $m (m \geqq 1)$ and for each $n (n \geqq 1)$, $A \in \sigma(1, m)$ and $B \in \sigma(m+n, \infty)$ together imply that

$$|P(AB) - P(A)P(B)| \leqq \alpha(n) P(A) \tag{2.1}$$

where $\alpha(n)$, $n=1, 2, \cdots$ is a nonnegative function of integers such that

$$\lim_{n \to \infty} \alpha(n) = 0. \tag{2.2}$$

The following Lemma will play a central role in this paper.

LEMMA 2.1. *Let $\{X_n\}$ be a uniform mixing stationary sequence and let the random variables η_1 and η_2 be measurable with respect to $\sigma(1, m)$ and $\sigma(m+n, \infty)$ respectively.*
(a) If $p>1$ and $q>1$, are two real numbers such that $1/p+1/q=1$, and if $E|\eta_1|^p < \infty$ and $E|\eta_2|^q < \infty$, then

$$|E[\eta_1 \eta_2] - E[\eta_1] E[\eta_2]| \leqq 2\{\alpha(n) E|\eta_1|^p\}^{1/p} \{E|\eta_2|^q\}^{1/q}. \tag{2.3}$$

(b) If $|\eta_i| \leqq C_i < \infty$ a.s., $i=1, 2$, then

$$|E[\eta_1 \eta_2] - E[\eta_1] E[\eta_2]| \leqq 2\alpha(n) C_1 C_2. \tag{2.4}$$

PROOF. Can be found in Billingsley, pp. 170-171.

REMARK 1. If the random variables η_1 and η_2 are complex, then separating the real and imaginary parts, we again arrive at part b of Lemma 2.1 with 2 replaced by 4.

3. Asymptotic Unbiasedness, Consistency, and Uniform Consistency

The following Lemmas play a central role in studying the asymptotic properties of $f_n(x)$.

LEMMA 3.1. *Suppose $h(y)$ is a Borel measurable function satisfying the conditions*

$$\begin{aligned} &\text{(i)} \quad \sup_{-\infty < y < \infty} |h(y)| < \infty \\ &\text{(ii)} \quad \int |h(y)| dy < \infty \\ &\text{(iii)} \quad \lim_{|y| \to \infty} |y h(y)| = 0 \end{aligned} \tag{3.1}$$

Let $g(y)$ satisfy the condition

$$\int |g(y)|dy<\infty. \tag{3.2}$$

Let $\{a_n\}$ *be a sequence of positive constants satisfying* (1.3). *Define*

$$g_n(x)=\frac{1}{a_n}\int h\Big(\frac{y}{a_n}\Big)g(x-y)dy. \tag{3.3}$$

Then at every point x *of continuity of* $g(\cdot)$

$$\lim_{n\to\infty} g_n(x)=g(x)\int h(y)dy. \tag{3.4}$$

PROOF. This Lemma is Theorem 1A in Parzen [14].

LEMMA 3.2. *Let* $h(y)$ *and* $g(y)$ *as in Lemma* 3.1, *then for any* $c\geqq 0$,

$$\lim_{n\to\infty}\int\frac{1}{a_n}\left|h\Big(\frac{y}{a_n}\Big)\right|^{1+c}g(x-y)dy=g(x)\int|h(y)|^{1+c}dy \tag{3.5}$$

at every x *of continuity of* $g(\cdot)$.

PROOF. Note that if

$$\operatorname*{Sup}_{-\infty<y<\infty}|h(y)|<\infty \text{ and } \int|h(y)|dy<\infty$$

then for every $c\geqq 0$

$$\int|h(y)|^{1+c}dy<\infty \tag{3.6}$$

the rest of the proof follows along the lines of the proof of Lemma 3.1.

LEMMA 3.3. *Assume that* $h(y)$ *satisfies condition* (3.1). *If* $g(x)$ *is uniformly continuous, then*

$$\lim_{n\to\infty}\operatorname*{Sup}_{-\infty<x<\infty}\left|g_n(x)-g(x)\int h(y)dy\right|=0. \tag{3.7}$$

PROOF. It is similar to the proof of Theorem 1A in Parzen [14] (modified to take account of the uniform continuity of $g(x)$).

THEOREM 3.4. (Asymptotic Unbiasedness). *Let* $\{X_n\}$ *be a stationary sequence, and let* $f_n(x)$ *be given by* (1.1). *Suppose* $K(y)$ *satisfies condition* (1.2) *and the constants* a_n *satisfy* (1.3). *If* $f(x)$ *is continuous, then at all points* x

$$\lim_{n\to\infty} Ef_n(x)=f(x). \tag{3.8}$$

PROOF. By stationarity, we have

$$Ef_n(x)=E\frac{1}{a_n}K\Big(\frac{x-X_1}{a_n}\Big)=\int\frac{1}{a_n}K\Big(\frac{y}{a_n}\Big)f(x-y)dy$$

which converges to $f(x)$ as $n\to\infty$ at all points x by an application of Lemma 3.1.

LEMMA 3.5. *Assume that the conditions of Theorem* 3.4 *are satisfied. If* $f(x)$ *is uniformly continuous, then*

$$\lim_{n\to\infty}\operatorname*{Sup}_{-\infty<x<\infty}|Ef_n(x)-f(x)|=0. \tag{3.9}$$

PROOF. By stationarity

$$\sup_{\infty<x<\infty} |Ef_n(x)-f(x)| = \sup_{-\infty<x<\infty} \left| E\frac{1}{a_n}K\left(\frac{x-X_1}{a_n}\right)-f(x)\right|$$

$$= \sup_{-\infty<x<\infty} \left|\int \frac{1}{a_n}K\left(\frac{y}{a_n}\right)f(x-y)dy-f(x)\right|$$

which converges to 0 as $n\to\infty$ by an application of Lemma 3.3.

The asymptotic behavior of the covariance of $f_n(x)$ at two points x and y will be determined before and consistency is considered.

THEOREM 3.6. *Let $\{X_n\}$ be a stationary sequence of r.v.'s which satisfies the uniform mixing condition and let $f_n(x)$ be defined as in* (1.1). *Suppose that the following conditions hold:*

(*i*) *$f(x)$ is continuous and bounded,*

(*ii*) $\sum_{j=1}^{\infty} \alpha^{1/2}(j)<\infty$

(*iii*) *$K(y)$ satisfies condition* (1.2),

(*iv*) *the constants a_n satisfy* (1.3), *and*

(*v*) *$f_j(x, y)$ (the joint density of X_1 and X_j, $j=2, 3, \cdots$) are continuous, bounded and*

$$\sum_{j\neq 1} |f_j(x, y)-f(x)f(y)| \leq M<\infty \qquad \forall x \text{ and } y.$$

Then at all points x and y

$$\lim_{n\to\infty} na_n \operatorname{Cov}[f_n(x), f_n(y)] = \begin{cases} f(x)\int K^2(z)dz & \text{if } x=y \\ 0 & \text{if } x\neq y. \end{cases} \tag{3.10}$$

PROOF. The proof resembles that of the proof of the asymptotic covariance of $f_n(x)$ at the points x and y in Rosenblatt [16].

By stationarity

$$\begin{aligned} \operatorname{Cov}[f_n(x), f_n(y)] &= \frac{1}{na_n^2}\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_1}{a_n}\right)\right] \\ &\quad + \frac{2}{n^2a_n^2}\sum_{j=2}^{n}(n-j+1)\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_j}{a_n}\right)\right] \\ &= I_{n1}+I_{n2}. \end{aligned} \tag{3.11}$$

Consider I_{n1}, then

$$\begin{aligned} \operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_1}{a_n}\right)\right] &= a_n\int K(z)K\left(\frac{y-x}{a_n}+z\right)f(x-a_nz)dz \\ &\quad - a_n^2\int K(z)f(x-a_nz)dz\int K(z)f(y-a_nz)dz \\ &= a_n^2\int K(z)\frac{1}{a_n}K\left(\frac{y-x}{a_n}+z\right)f(x-a_nz)dz \\ &\quad - a_n^2\int K(z)f(x-a_nz)dz\int K(z)f(y-a_nz)dz \end{aligned} \tag{3.12}$$

$$\cong \begin{cases} a_n f(x)\int K^2(z)dz & \text{if } x=y \\ 0(a_n^2) & \text{if } x\neq y \end{cases}$$

as $n\to\infty$, because $f(x)$ is continuous and bounded.

The terms in I_{n2} have the following asymptotic behavior,

$$\begin{aligned} &\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_j}{a_n}\right)\right] \\ &\quad = a_n^2 \iint K(z_1)K(z_2)f_j(x-a_nz_1, y-a_nz_2)dz_2dz_1 \\ &\quad\quad - a_n^2\int K(z)f(x-a_nz)dz\int K(z)f(y-a_nz)dz \\ &\quad \cong a_n^2\{f_j(x, y)-f(x)f(y)\} \end{aligned} \tag{3.13}$$

as $n\to\infty$, because the joint density functions $f_j(x, y)$ are continuous and bounded functions. We shall now get a bound on (3.13) under the assumptions that the sequence $\{X_n\}$ satisfies the uniform mixing condition with mixing coefficient $\alpha(n)$ such that

$$\sum_{j=1}^{\infty}\alpha^{1/2}(j)<\infty.$$

From (2.3), we have

$$\begin{aligned} \left|\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_j}{a_n}\right)\right]\right| &\leqq 2\left\{\alpha(j-1)E\left|K\left(\frac{x-X_1}{a_n}\right)\right|^2E\left|K\left(\frac{y-X_1}{a_n}\right)\right|^2\right\}^{1/2} \\ &\cong 2\alpha^{1/2}(j-1)a_n\sqrt{f(x)f(y)}\int K^2(z)dz \end{aligned} \tag{3.14}$$

for sufficiently large n. Inequality (3.14) implies that

$$\begin{aligned} &\frac{1}{n}\sum_{j=2}^{n}(n-j+1)\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right), K\left(\frac{y-X_j}{a_n}\right)\right] \\ &\quad \leqq 2a_n\sum_{j=1}^{\infty}\alpha^{1/2}(j)\sqrt{f(x)f(y)}\int K^2(z)dz. \end{aligned} \tag{3.15}$$

The inequalities (3.12)-(3.15) indicate that

$$na_n\operatorname{Cov}[f_n(x), f_n(y)]\cong\begin{cases} f(x)\int K^2(z)dz+2a_n\sum_{j=2}^{n}\left(1-\frac{j-1}{n}\right)\{f_j(x, x)-f^2(x)\} & \text{if } x=y \\ 0(a_n)+2a_n\sum_{j=2}^{n}\left(1-\frac{j-1}{n}\right)\{f_j(x, y)-f(x)f(y)\} & \text{if } x\neq y \end{cases}$$

and therefore by condition (v)

$$\lim_{n\to\infty} na_n\operatorname{Cov}[f_n(x), f_n(y)]=\begin{cases} f(x)\int K^2(z)dz & \text{if } x=y \\ 0 & \text{if } x\neq y. \end{cases}$$

From the above theorem one can state conditions under which the estimate $f_n(x)$ is consistent in quadratic means in the sense that

$$\lim_{n\to\infty} E|f_n(x)-f(x)|^2=0 \qquad \text{at all points } x.$$

This implies that $f_n(x)$ converges to $f(x)$ in probability.

LEMMA 3.7. *Assume that the conditions of Theorem* 3.6 *are satisfied. Suppose that the constants a_n satisfy* (1.3) *and the following condition*

$$\lim_{n\to\infty} na_n=\infty. \tag{3.16}$$

Then at all points x

$$\lim_{n\to\infty}|Ef_n(x)-f(x)|^2=0 \tag{3.17}$$

and

$$f_n(x) \xrightarrow{P} f(x). \tag{3.18}$$

PROOF. The mean square error can be written as

$$E|f_n(x)-f(x)|^2=\operatorname{Var} f_n(x)+|Ef_n(x)-f(x)|^2$$

which converges to 0 as $n\to\infty$ at all points x by applying Theorem 3.4, (3.10), and (3.16).

Now by (3.17) and Chebyshev's inequality, we have

$$f_n(x) \xrightarrow{P} f(x) \qquad \text{at all points } x \text{ (as } n\to\infty).$$

In the remainder of this section, we shall show that the estimate $f_n(x)$ is uniformly consistent.

We define

$$\phi(t)=\int e^{itx}f(x)dx \tag{3.19}$$

$$k(t)=\int e^{itx}K(x)dx \tag{3.20}$$

$$\phi_n(t)=\frac{1}{n}\sum_{j=1}^{n} e^{itX_j} \tag{3.21}$$

where $\phi(t)$, $k(t)$ and $\phi_n(t)$ are the characteristic function (c.f.) corresponding to the p.d.f. $f(x)$, the Fourier transform of $K(y)$, and the sample c.f., respectively.

THEOREM 3.8. *Assume $\{X_n\}$ is a sequence of stationary r.v.'s satisfying the uniform mixing condition and let $f_n(x)$ be defined as in* (1.1). *Assume that the following conditions hold:*

(*i*) *$f(x)$ is uniformly continuous*

(*ii*) *the constants a_n satisfy* (1.3), *and*

$$\lim_{n\to\infty} na_n^2=\infty \tag{3.22}$$

(*iii*) $\int|k(t)|\,dt<\infty$, *and*

(*iv*) $\sum_{j=1}^{\infty}\alpha(j)<\infty$.

Then

$$\operatorname*{Sup}_{-\infty<x<\infty}|f_n(x)-f(x)| \xrightarrow{P} 0 \quad \text{as } n\to\infty. \tag{3.23}$$

PROOF. Since $K(y)$ and $k(t)$ are absolutely integrable, we have

$$K(y)=\frac{1}{2\pi}\int e^{-ity}k(t)dt. \tag{3.24}$$

In terms of $k(t)$, we have

$$\begin{aligned} f_n(x)&=\frac{1}{2\pi}\int\left\{\frac{1}{n}\sum_{j=1}^{n}e^{itX_j}k(a_nt)\right\}e^{-itx}dt \\ &=\frac{1}{2\pi}\int e^{-itx}\phi_n(t)k(a_nt)dt. \end{aligned} \tag{3.25}$$

Hence

$$|f_n(x)-Ef_n(x)|=\left|\frac{1}{2\pi}\int\{\phi_n(t)-\phi(t)\}k(a_nt)e^{-itx}dt\right|$$

Therefore, we have

$$\operatorname*{Sup}_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|\leqq\frac{1}{2\pi}\int|\phi_n(t)-\phi(t)|\,|k(a_nt)|dt \tag{3.26}$$

using the fact that $|e^{-itx}|=1$.

It follows from (3.26), Fubini's Theorem and Schwartz's inequality that

$$\begin{aligned} E[\operatorname*{Sup}_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|]&\leqq\frac{1}{2\pi}\int\{E|\phi_n(t)-\phi(t)|^2\}^{1/2}|k(a_nt)|dt \\ &=\frac{1}{2\pi}\int\{\sigma^2[\phi_n(t)]\}^{1/2}|k(a_nt)|dt. \end{aligned} \tag{3.27}$$

By stationarity and Remark 1, we have

$$\begin{aligned} \sigma^2[\phi_n(t)]&=\frac{1}{n}E|e^{itX}-\phi(t)|^2 \\ &\quad+\frac{2}{n^2}\sum_{j=2}^{n}(n-j+1)E[e^{itX_1}-Ee^{itX_1}]\overline{[e^{itX_j}-Ee^{itX_j}]} \\ &=\frac{1}{n}[1-|\phi(t)|^2]+\frac{2}{n^2}\sum_{j=2}^{n}(n-j+1)E[e^{it(X_1-X_j)}-Ee^{itX_1}Ee^{-itX_j}] \\ &\leqq\frac{1}{n}+\frac{8}{n}\sum_{j=1}^{n}\alpha(j)\leqq\frac{1}{n}\left[1+8\sum_{j=1}^{\infty}\alpha(j)\right]. \end{aligned} \tag{3.28}$$

It follows from (3.27) and (3.28) that

$$\begin{aligned} E[\operatorname*{Sup}_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|]&\leqq\left[\frac{1}{2\pi\sqrt{n}}\int|k(a_nt)|dt\right]\left[1+8\sum_{j=1}^{\infty}\alpha(j)\Big)\right]^{1/2} \\ &=\frac{1}{2\pi\sqrt{na_n^2}}\int|k(t)|dt\left[1+8\sum_{j=1}^{\infty}\alpha(j)\right]^{1/2}. \end{aligned} \tag{3.29}$$

By applying (3.22), (iii), (iv), and (3.29), we have

$$\lim_{n\to\infty}E[\operatorname*{Sup}_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|]=0. \tag{3.30}$$

It follows from (3.30) and Markov's inequality that

$$\operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-Ef_n(x)| \xrightarrow{P} 0 \qquad (\text{as } n\to\infty). \tag{3.31}$$

In the inequality

$$\operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-f(x)| \leq \operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-Ef_n(x)| + \operatorname*{Sup}_{-\infty<x<\infty} |Ef_n(x)-f(x)|.$$

The RHS converges to 0 in probability as $n\to\infty$ by (3.31) and (3.9).

THEOREM 3.9. *Assume that* $\{X_n\}$ *is a sequence of stationary r.v.'s which satisfies the uniform mixing condition and let* $f_n(x)$ *be defined as in* (1.1). *Suppose that the following conditions hold*:

(*i*) $f(x)$ *is uniformly continuous.*

(*ii*) *The constants* a_n *satisfy* (1.3), *and*

$$\sum_{n=1}^{\infty} \frac{1}{(na_n^2)^2} < \infty,$$

(*iii*) $\int |k(t)|\,dt < \infty$, *and*

(*iv*) $\sum_{j=1}^{\infty} \alpha^{1/2}(j) < \infty.$

Then

$$\operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-f(x)| \xrightarrow{W.P.1} 0 \qquad (\text{as } n\to\infty). \tag{3.32}$$

PROOF. From (3.26), we have

$$\operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-Ef_n(x)| \leq \frac{1}{2\pi}\int |\phi_n(t)-\phi(t)|\,|k(a_nt)|\,dt. \tag{3.33}$$

It follows from (3.33), Schwartz's inequality, and Fubini's Theorem that

$$\begin{aligned} E[\operatorname*{Sup}_{-\infty<x<\infty} |f_n(x)-f(x)|^4] &\leq \frac{1}{16\pi^4} E\Big[\int |\phi_n(t)-\phi(t)|\,|k(a_nt)|\,dt\Big]^4 \\ &= \frac{1}{16\pi^4}\iiiint \prod_{i=1}^{4} |k(a_nt_i)| \Big(E\prod_{i=1}^{4} |\phi_n(t_i)-\phi(t_i)|\Big) dt_i \\ &\leq \frac{1}{16\pi^4}\iiiint \prod_{i=1}^{4} |k(a_nt_i)|\, E^{1/4}|\phi_n(t_i)-\phi(t_i)|^4 dt_i. \end{aligned} \tag{3.34}$$

Since $|e^{itX}-\phi(t)|\leq 2$, then by Lemma 4, p. 173 of Billingsley, we have

$$\begin{aligned} E|\phi_n(t)-\phi(t)|^4 &\leq \frac{1}{n^4} E\sum_{j=1}^{n} |e^{itX_j}-\phi(t)|^4 \\ &\leq \frac{(768)16}{n^2}\Big[\sum_{j=1}^{\infty} \alpha^{1/2}(j)\Big]^2. \end{aligned} \tag{3.35}$$

It follows from (3.34) and (3.35) that

$$E[\sup_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|^4]\leqq\frac{768}{n^2\pi^4}\Big(\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big)^2\Big(\int|k(a_nt)|dt\Big)^4$$

$$=\frac{768}{\pi^4n^2a_n^4}\Big(\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big)^2\Big(\int|k(t)|dt\Big)^4. \tag{3.36}$$

By Markov's inequality and (3.36), we have for $\varepsilon>0$

$$P[\sup_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|>\varepsilon]$$

$$\leqq\frac{768}{\varepsilon^4n^2a_n^4\pi^4}\Big(\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big)^2\Big(\int|k(t)|dt\Big)^4. \tag{3.37}$$

It follows from (3.37), (ii)-(iv), and Borel-Cantelli's Lemma, we have

$$\sup_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|\xrightarrow{\text{W.P.1}}0\qquad(\text{as } n\to\infty). \tag{3.38}$$

In the inequality

$$\sup_{-\infty<x<\infty}|f_n(x)-f(x)|\leqq\sup_{-\infty<x<\infty}|f_n(x)-Ef_n(x)|$$

$$+\sup_{-\infty<x<\infty}|Ef_n(x)-f(x)|.$$

The RHS converges to 0 W.P.1 as $n\to\infty$ by (3.38) and (3.9).

4. Asymptotic Normality of $f_n(x)$

In this section we will establish the asymptotic normality of $f_n(x)$ when $\{X_n\}$ is a sequence of r.v.'s satisfying the uniform mixing condition with the assumption that

$$\sum_{j=1}^{\infty}\alpha^{1/2}(j)<\infty.$$

The main idea is to present $\sqrt{na_n}[f_n(x)-Ef_n(x)]$ as a sum of big blocks separated by small blocks which will be shown to be negligible and the big blocks approximately independent. Liapounov's Theorem is then used to get the asymptotic normality.

THEOREM 4.1. *Assume that the conditions of Theorem* 3.6 *are satisfied. Suppose that the following conditions hold:*

(*i*) $\lim_{n\to\infty} na_n=\infty$

(*ii*) *for any pair of sequence* $m=m(n)$, $r=r(n)$, *such that* $m, r\to\infty$ *as* $n\to\infty$ *but* $m=o(n^{1/3}a_n^{-2/3})$, $r=o(m(n))$, *and*

$$\lim_{n\to\infty} nm^{-1}\alpha(r)=0, \tag{4.1}$$

and

(*iv*) *the joint density functions up to the fourth order are continuous and bounded.*
If $f_n(x)$ *is defined as in* (1.1), *then*

$$\sqrt{na_n}\,[f_n(x)-Ef_n(x)]\xrightarrow{D}N\Big(0,f(x)\int K^2(z)dz\Big).$$

PROOF.

$$\sqrt{na_n}[f_n(x)-Ef_n(x)]=\sum_{j=1}^{n}\frac{1}{\sqrt{na_n}}\left[K\left(\frac{x-X_j}{a_n}\right)-EK\left(\frac{x-X_j}{a_n}\right)\right]$$
$$=\sum_{q=1}^{v}A_q+\sum_{q=1}^{v+1}B_q \tag{4.2}$$

with

$$A_q=\sum_{j=(q-1)(m+r)+1}^{qm+(q-1)r}U_j,\quad q=1,\cdots,v \tag{4.3}$$

$$B_q=\sum_{j=qm+(q-1)r+1}^{q(m+r)}U_j,\quad q=1,\cdots,v \tag{4.4}$$

$$B_{v+1}=\sum_{j=v(m+r)+1}^{n}U_j \tag{4.5}$$

where

$$U_j=\frac{1}{\sqrt{na_n}}\left[K\left(\frac{x-X_j}{a_n}\right)-EK\left(\frac{x-X_j}{a_n}\right)\right] \tag{4.6}$$

and $v=v(n)=\left[\frac{n}{m+r}\right]$ is the greatest integer less than or equal to $\frac{n}{m+r}$. Notice that $v\to\infty$ as $n\to\infty$, since m, $r=o(n)$.

By stationarity,

$$E\left[\sum_{q=1}^{v+1}B_q\right]^2=vEB_1^2+EB_{v+1}^2+2\sum_{q=2}^{v}(v-q+1)\operatorname{Cov}(B_1,B_q)$$
$$+2\sum_{q=1}^{v}\operatorname{Cov}(B_q,B_{v+1}). \tag{4.7}$$

The four terms in the RHS of (4.7) are evaluated as follows:

$$vEB_1^2=\frac{vr}{na_n}\operatorname{Var}\left[K\left(\frac{x-X_1}{a_n}\right)\right]+\frac{2v}{na_n}\sum_{j=2}^{r}(r-j+1)\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right),K\left(\frac{x-X_j}{a_n}\right)\right]$$
$$\leqq\frac{vr}{na_n}\left\{\operatorname{Var}\left[K\left(\frac{x-X_1}{a_n}\right)\right]+2\sum_{j=2}^{r}\left|\operatorname{Cov}\left[K\left(\frac{x-X_1}{a_n}\right),K\left(\frac{x-X_j}{a_n}\right)\right]\right|\right\}$$
$$\leqq\frac{vr}{na_n}\left\{\operatorname{Var}\left[K\left(\frac{x-X_1}{a_n}\right)\right]+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)E\left|K\left(\frac{x-X_1}{a_n}\right)\right|^2\right\}\text{ by (2.3)}$$
$$\cong\frac{vr}{n}\left\{f(x)\int K^2(z)dz\left[1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\right]\right\}\text{ by Lemma 3.2.} \tag{4.8}$$

Similarly, the second term will be

$$EB_{v+1}^2\cong\left[\frac{n-v(m+r)}{n}\right]\left[f(x)\int K^2(z)dz\left(1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\right)\right]$$
$$\leqq\left(\frac{m+r}{n}\right)\left[f(x)\int K^2(z)dz\left(1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\right)\right] \tag{4.9}$$

because $n-v(m+r)=n-(m+r)\left[\dfrac{n}{m+r}\right]\leqq m+r$ is the number of terms in B_{v+1}. The third term is

$$2\sum_{q=2}^{v}(v-q+1)\,\mathrm{Cov}\,(B_1,\ B_q)$$

$$=2\sum_{q=2}^{v}\Big(\frac{v-q+1}{na_n}\Big)\mathrm{Cov}\Big[\sum_{j=(q-1)(m+r)+1}^{q(m+r)-m}K\Big(\frac{x-X_j}{a_n}\Big),\ \sum_{p=1}^{r}K\Big(\frac{x-X_p}{a_n}\Big)\Big]$$

$$\leqq\frac{2vr}{na_n}\sum_{q=2}^{v}\sum_{j=(q-1)(m+r)+1}^{q(m+r)-m}\mathrm{Cov}\Big[K\Big(\frac{x-X_j}{a_n}\Big),\ K\Big(\frac{x-X_r}{a_n}\Big)\Big] \tag{4.10}$$

$$\leqq\frac{4vr}{n}f(x)\int K^2(z)dz\Big[\sum_{q=2}^{v}\sum_{j=(q-1)(m+r)+1}^{q(m+r)-m}\alpha^{1/2}(j-r)\Big]$$

$$\leqq\frac{4vr}{n}f(x)\int K^2(z)dz\Big[\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big].$$

Similarly, the fourth term

$$\sum_{q=1}^{v}\mathrm{Cov}\,(B_q,\ B_{v+1})\leqq 4\Big(\frac{m+r}{n}\Big)\Big[f(x)\int K^2(z)dz\Big]\Big[\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big]. \tag{4.11}$$

Substituting (4.8)-(4.11) into (4.7), we obtain

$$E\Big[\sum_{q=1}^{v+1}B_q\Big]^2\leqq 2\Big[\frac{vr+m+r}{n}\Big]\Big[f(x)\int K^2(z)dz\Big]\Big[1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\Big]. \tag{4.12}$$

Since $\sum_{j=1}^{\infty}\alpha^{1/2}(j)<\infty$, $\dfrac{vr}{n}\cong\dfrac{r}{m}=o(1)$, m, $r=o(n)$, then

$$E\Big[\sum_{j=1}^{v+1}B_q\Big]^2=o(1). \tag{4.13}$$

It follows from Chebyshev's inequality that

$$\sum_{q=1}^{v+1}B_q\xrightarrow{P}0 \qquad \text{as } n\to\infty. \tag{4.14}$$

Next, we prove the asymptotic normality of $\sum_{q=1}^{v}A_q$.

(1) We have to show that A_q's are asymptotically independent, i.e.,

$$I_n=\Big|E\exp\Big\{\mathrm{it}\ \sum_{q=1}^{v}A_q\Big\}-\prod_{q=1}^{v}E\{\exp\mathrm{it}\ A_q\}\Big|\to 0 \qquad \text{as } n\to\infty. \tag{4.15}$$

From p. 318 of Ibragimov and Linnik [9], we have

$$I_n\leqq\sum_{s=2}^{v}\Big|E\exp\Big\{\mathrm{it}\ \sum_{q=1}^{s}A_q\Big\}-E\exp\{\mathrm{it}\ A_s\}E\exp\Big\{\mathrm{it}\ \sum_{q=1}^{s-1}A_q\Big\}\Big|. \tag{4.16}$$

Now, since $\exp\Big\{\mathrm{it}\ \sum_{q=1}^{v-1}A_q\Big\}$ is measurable with respect to $\sigma((v-1)m+(v-2)r)$ and $\exp(\mathrm{it}\ A_v)$ is measurable with respect to $\sigma((v-1)(m+r)+1)$ then by Remark 1

$$\Big|E\exp\Big\{\mathrm{it}\ \sum_{q=1}^{v}A_q\Big\}-E\exp\Big\{\mathrm{it}\ \sum_{q=1}^{v-1}A_q\Big\}E\exp\{\mathrm{it}\ A_v\}\Big|\leqq 4\alpha(r+1). \tag{4.17}$$

Similarly for $s \leqq v-1$

$$\left| E \exp\left\{ \mathrm{it} \sum_{q=1}^{s} A_q \right\} - E \exp\left\{ \mathrm{it} \sum_{q=1}^{s-1} A_q \right\} E \exp\{\mathrm{it}\ A_s\} \right| \leqq 4\alpha(r+1). \tag{4.18}$$

From (4.16)-(4.18), we have

$$I_n \leqq 4v\alpha(r) \cong 4nm^{-1}\alpha(r+1). \tag{4.19}$$

It follows from (4.1) that

$$I_n \longrightarrow 0 \qquad \text{as } n\to\infty. \tag{4.20}$$

So A_q's are asymptotically independent. Thus the normality of $\sum_{q=1}^{v} A_q$ will follow if we prove that

(2) $C_v/D_v^2 \longrightarrow 0$ as $n\to\infty$ $(v\to\infty)$

where

$$D_v = \sum_{q=1}^{v} E(A_q^2), \qquad \text{and} \tag{4.21}$$

$$C_v = \sum_{q=1}^{v} E(A_q^4). \tag{4.22}$$

By stationarity

$$\begin{aligned} D_v = vEA_1^2 = \frac{vm}{na_n}&\left\{ \mathrm{Var}\left[K\left(\frac{x-X_1}{a_n}\right)\right]\right. \\ &\left. + \frac{2}{m}\sum_{j=2}^{m}(m-j+1)\,\mathrm{Cov}\left[K\left(\frac{x-X_1}{a_n}\right),\ K\left(\frac{x-X_j}{a_n}\right)\right]\right\} \\ \leqq \frac{vm}{n}&\left\{ f(x)\int K^2(z)dz\left[1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\right]\right\} \\ \longrightarrow\ & f(x)\int K^2(z)dz\left[1+4\sum_{j=1}^{\infty}\alpha^{1/2}(j)\right]. \end{aligned} \tag{4.23}$$

The first inequality is similar to the proof of (4.8) and the limit because $\frac{vm}{n}\sim 1$.

From (4.22), we have

$$\begin{aligned} C_v = vE(A_1^4) = \frac{v}{n^2a_n^2}&\left\{ mE\left|K\left(\frac{x-X_1}{a_n}\right)-\mu\right|^4 \right. \\ &+ \sum_{i\neq j} E\left|K\left(\frac{x-X_i}{a_n}\right)-\mu\right|^2 \left|K\left(\frac{x-X_j}{a_n}\right)-\mu\right|^2 \\ &+ \sum_{i\neq j} E\left|K\left(\frac{x-X_i}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_j}{a_n}\right)-\mu\right|^3 \\ &+ \sum_{i\neq j\neq l} E\left|K\left(\frac{x-X_i}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_j}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_l}{a_n}\right)-\mu\right|^2 \\ &\left. + \sum_{i\neq j\neq l\neq w} E\left|K\left(\frac{x-X_i}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_j}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_l}{a_n}\right)-\mu\right| \left|K\left(\frac{x-X_w}{a_n}\right)-\mu\right| \right\} \end{aligned} \tag{4.24}$$

where

$$\mu = EK\Big(\frac{x-X_1}{a_n}\Big) = a_n\Big[\frac{1}{a_n}\int K\Big(\frac{z}{a_n}\Big)f(x-z)dz\Big]$$
$$\cong a_n f(x) \longrightarrow 0 \qquad \text{as } n\to\infty. \tag{4.25}$$

Then from (4.24) and (4.25) implies that

$$\begin{aligned} C_v \cong \frac{v}{n^2a_n^2}\Big\{ & mE\Big|K\Big(\frac{x-X_1}{a_n}\Big)\Big|^4 + \sum_{i\neq j}\Big[E\Big|K\Big(\frac{x-X_i}{a_n}\Big)\Big|^2\Big|K\Big(\frac{x-X_j}{a_n}\Big)\Big|^2 \\ & + E\Big|K\Big(\frac{x-X_i}{a_n}\Big)\Big|\Big|K\Big(\frac{x-X_j}{a_n}\Big)\Big|^3\Big] + \sum_{i\neq j\neq l} E\Big|K\Big(\frac{x-X_i}{a_n}\Big)K\Big(\frac{n-X_j}{a_n}\Big)\Big|\Big|K\Big(\frac{x-X_l}{a_n}\Big)\Big|^2 \\ & + \sum_{i\neq j\neq l\neq w} E\Big|K\Big(\frac{x-X_i}{a_n}\Big)K\Big(\frac{x-X_j}{a_n}\Big)K\Big(\frac{x-X_l}{a_n}\Big)K\Big(\frac{x-X_w}{a_n}\Big)\Big|\Big\}. \end{aligned} \tag{4.26}$$

But

$$E\Big|K\Big(\frac{x-X_1}{a_n}\Big)\Big|^4 = \int K^4\Big(\frac{x-z_1}{a_n}\Big)f(z_1)dz_1$$
$$\cong a_n f(z)\int K^4(w_1)dw_1 \tag{4.27}$$

$$E\Big|K\Big(\frac{x-X_i}{a_n}\Big)K^3\Big(\frac{x-X_j}{a_n}\Big)\Big| = \int\Big|K\Big(\frac{x-z_1}{a_n}\Big)K^3\Big(\frac{x-z_2}{a_n}\Big)\Big|f_{ij}(z_1, z_2)dz_1dz_2$$
$$\cong a_n^2 f_{ij}(z_1, z_2)\int|K^3(w_1)|dw_1. \tag{4.28}$$

Similarly,

$$E\Big|K^2\Big(\frac{x-X_i}{a_n}\Big)K^2\Big(\frac{x-X_j}{a_n}\Big)\Big| \cong a_n^2 f_{ij}(z_1, z_2)\Big[\int K^2(w_1)dw_1\Big]^2 \tag{4.29}$$

$$E\Big|K\Big(\frac{x-X_i}{a_n}\Big)K\Big(\frac{x-X_j}{a_n}\Big)K^2\Big(\frac{x-X_l}{a_n}\Big)\Big| \cong a_n^3 f_{ijl}(z_1, z_2, z_3)\int K^2(w_1)dw_1 \tag{4.30}$$

and

$$E\Big|K\Big(\frac{x-X_i}{a_n}\Big)K\Big(\frac{x-X_j}{a_n}\Big)K\Big(\frac{x-X_l}{a_n}\Big)K\Big(\frac{x-X_w}{a_n}\Big)\Big| \cong a_n^4 f_{ijlw}(z_1, z_2, z_3, z_4). \tag{4.31}$$

Since the joint probability density functions up to the fourth order are continuous and bounded, (4.24)-(4.31) imply that

$$C_v \leqq \frac{vm}{n^2a_n}M_1 + \frac{vm^2}{n^2}M_2 + \frac{vm^3a_n}{n^2}M_3 + \frac{vma_n^2}{n^2}M_4$$
$$\leqq \frac{M}{n^2}\Big[\frac{vm}{a_n} + vm^2 + vm^3a_n + vm^4a_n^2\Big] \tag{4.32}$$

where $M = \text{Max}\,[M_1, \cdots, M_4]$ and M_i are constants, $i=1, \cdots, 4$. But since $\frac{vm}{n} \sim 1$ and conditions (i) and (ii), we have

$$C_v \leqq M\left[\frac{1}{na_n}+\frac{m}{n}+\frac{m^2 a_n}{n}+\frac{m^3 a_n^2}{n}\right]=o(1). \tag{4.33}$$

It follows from (4.23)-(4.33), that

$$C_v/D_v^2 \longrightarrow 0 \quad \text{as } n\to\infty.$$

Acknowledgement

The authors are deeply grateful to the referee for the useful comments and suggestions.

References

[1] ABDULAL, K. I.: *On Density Estimation*, Unpublished Ph. D. Dissertation, Colorado State University, (1984).

[2] BEAN, S. J. and TSOKOS, C. P.: *Developments in Nonparametric Density Estimation*. Int. St. RVW, **48**, (1980), 267-287.

[3] BILLINGSLEY, P.: *Convergence of Probability Measures*, Wiley and Sons, New York, (1968).

[4] BRADLEY, R. C.: *Central Limit Theorems under Weak Dependence*, J. Multi. An., **11**, (1981), 1-16.

[5] BRADLEY, R. C.: *Asymptotic Normality of Some Kernel Type Estimates of Probability Density*, Stat. and Prob. Letters, **1**, (1983), 295-300.

[6] DAVIES, H. I.: *Strong Consistency of a Sequential Estimator of a.p.d.f.*, Bull. Math. Stat., **15**, (1973), 49-59.

[7] DOOB, J.: *Stochastic Processes*, Wiley and Sons, New York, (1953).

[8] IBRAGIMOV, I. A.: *Some Limit Theorems for Stationary Processes*, Theor. Probability Appl., **7**, (1962), 349-382.

[9] IBRAGIMOV, I. A. and LINNIK, YU J.: *Independent and Stationary Dependent Random Variables*, Walters-Noordhoff Publishing, Groningen, Netherlands, (1971).

[10] LEADBETTER, M. R.: *On the Nonparametric Estimation of Probability Densities*, Technical Report No. 11, Research Triangle Institute, N. C., (1963).

[11] MURTHY, V. K.: *Estimation of Probability Density*, Ann. Math. Stat., **36**, (1965), 1027-1031.

[12] NADARAYA, E. A.: *Some New Estimates for Distribution Functions*, Theor. Probability Appl., **9**, (1964), 497-500.

[13] NADARAYA, E. A.: *On Nonparametric Estimates of Density Functions and Regression Curves*, Theor. Probability Appl., **10**, (1965), 186-190.

[14] PARZEN, E.: *On the Estimation of a Probability Density Function and the Mode*, Ann. Math. Stat., **33**, (1962), 1065-1076.

[15] PRAKASA RAO, B.L.S.: *Nonparametric Functional Estimation*, Academic Press, (1983).

[16] ROSENBLATT, M.: *Density Estimates of Markov Sequences*, Nonparametric Techniques in Statistical Inference, M. L. Puri (ed.), Cambridge University Press, (1970), 199-213.

[17] ROUSSAS, G.: *Nonparametric Estimation in Markov Processes*, Ann. Inst. Statist. Math., **21**, (1969), 73-87.

[18] ROUSSAS, G.: *Nonparametric Estimation of the Transition Distribution Function of a Markov Process*, Ann. Math. Stat., **40**, (1969), 1386-1399.

[19] TAPIA, R. A. and THOMPSON, J. R.: *Nonparametric Probability Density Estimation*, The Johns Hopkins University Press, Baltimore, (1978).

[20] YAMATO, H.: *Sequential Estimation of a Continuous Probability Density Function and Mode,* Bull. Math. Stat., **14**, (1971), 1-12.

Communicated by S. Kano
Received November 2, 1984
Revised February 14, 1985