RATE OF CONVERGENCE FOR NON PARAMETRIC DENSITY ESTIMATION IN LINEAR PROCESS

Basu, A. K.
University of Calcutta

Sahoo, D. K.
R. K. M. College

https://doi.org/10.5109/13407

バージョン：
権利関係：
RATE OF CONVERGENCE FOR NON PARAMETRIC DENSITY ESTIMATION IN LINEAR PROCESS

By

A. K. BASU* and D. K. SAHOO**

Abstract

Rate of convergence to normality for the density estimators of Kernel type is obtained when the observations are from a stationary linear process. At first, the case of estimating the density at a fixed point is considered and latter on, it is extended for estimating joint density. Also the problem of estimating the density at several points is considered.

1. Introduction

Problem of estimating the unknown density function of a population have been considered by several authors. An excellent survey of the results is given in Rosenblatt [9] when the sample observations are independent. Attempts also have been made to extend the results to other than independent observations. Rosenblatt [8] considered the case where the sample observations are dependent in the sense that they are sampled from a stationary Markov sequence and also derived some interesting results about the Kernel type density estimators.

In this paper our aim is to extend these results for the random sequence \( \{X_t\}_{t=-\infty}^{\infty} \) when they form a stationary linear process. Most of the important stochastic process models, such as, Auto regressive schemes, Moving average schemes etc. are linear processes. In an unpublished paper, Kamal Chanda considered the problem of estimating the probability density function for the linear process and try to investigate the asymptotic properties like almost sure convergence and asymptotic normality. But unfortunately there are some mistakes in his calculations and also he does not consider the rate of convergence to normality. We exploit the techniques of Blume and Wittwer [2] and N. Kersten [5] and use some results of Ibragimov [4] to find the rate of convergence to normality of the estimated non parametric density function. In section 2 we consider the case of estimating the density at a fixed point and in section 3 we extend it for estimating joint density. Also in section 4 we consider the problem of estimating probability density at several points. Using different methods uniform rates of convergence to normality for non parametric density estimates have been obtained by Wertz [10], Prakasha Rao [6, 7], Basu and Sahoo [1] and their rates where at best of the order of \( n^{-1/3+\gamma} \), \( 0<\gamma<1/3 \). In this paper assuming quadratic

* University of Calcutta, India.
** R.K.M. College, Narendrapur, India.
Kernels and existence of higher moments we are able to get rate of convergence of the order of $n^{-1/5} (\log n)^\alpha$, $\alpha > 0$. Our method can be used for polynomial Kernels and Kernels which can well approximated by polynomial Kernels.

2. Probability Density Estimate and Its Rate of Convergence to Normality

Let $\{X_t\} \in \mathbb{R}$ be a stationary linear process defined by

$$X_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = 1$ and $\{\alpha_j\}_{j=0}^{\infty}$ is a sequence of real numbers with $\alpha_0 \geq 0$ and $\sum_{j=0}^{\infty} |\alpha_j| < \infty$.

Suppose the process is observed at the points $t=1, 2, \ldots, N$ and $X_i$ has the pdf $f(x)$. We formally define the estimator $f_N(x)$ of $f(x)$ as,

$$f_N(x) = \frac{1}{N} \sum_{t=1}^{N} \frac{1}{h} K \left[ \frac{x-X_t}{h} \right]$$

where $\{h = h_N\}$ is a sequence of real numbers such that $h_N \to 0$ but $Nh_N \to \infty$ as $N \to \infty$ (e.g. $h_N \sim N^{-1/5}$) and $K$ is a Kernel function satisfying:

$$\begin{align*}
(\text{i}) & \quad \sup_{y} K(y) < \infty \\
(\text{ii}) & \quad \int_{-\infty}^{\infty} K(y) dy < \infty, \quad \lim_{y \to \pm \infty} y K(y) = 0
\end{align*}$$

We make the following additional assumptions:

(iii) if $\phi_x$ denote the characteristic function of $\varepsilon_t$ then $\int_{-\infty}^{\infty} |u \phi_x(u)| du < \infty$

(iv) $E|\varepsilon_t| < \infty$

(v) $\sum_{j=N}^{\infty} |\alpha_j| \leq C \cdot 2^{-\alpha N}/N^{1/5}$, where $C$ is a generic symbol which denotes a positive finite constant, independent of $N$ and $\alpha$ is a positive constant.

We set

$$f_t = \frac{1}{\sqrt{h}} \left[ K((x-X_t)/h) - EK((x-X_t)/h) \right]$$

and

$$T_N = \sqrt{Nh} \left[ f_N - \tilde{f}_N \right], \quad \text{where} \quad \tilde{f}_N = (1/h) EK((x-X_t)/h)$$

where

$$\begin{align*}
(\text{2.5}) & \quad \frac{1}{N} \sum_{t=1}^{N} \frac{1}{h} K((x-X_t)/h) - EK((x-X_t)/h) \\
(\text{2.5}) & \quad \frac{1}{N} \sum_{t=1}^{N} \frac{1}{h} K((x-X_t)/h) - EK((x-X_t)/h) \\
(\text{2.5}) & \quad \frac{1}{N} \sum_{t=1}^{N} \frac{1}{h} K((x-X_t)/h) - EK((x-X_t)/h)
\end{align*}$$
In case the strict stationarity of \( \{X_t\} \) is valid,
\[
P\left[ \frac{1}{\sqrt{N}} \sum_{t=1}^{N} f_t < \varepsilon \right] = P\left[ \frac{1}{(1/N)^{1/2}} \sum_{t=1}^{N} f_t < \varepsilon \right].
\]
We take a second degree polynomial Kernel \( K(y) = a_0 + a_1 y + a_2 y^2 \). Such type of Kernels are widely used. In fact if we take \( a_0 = 3/4 \sqrt{5} \), \( a_1 = 0 \), \( a_2 = -3/20 \sqrt{5} \) then we will get Epanechnikov [3] Kernel which is optimum in the sense that it minimizes integrated mean square error. So
\[
K((x - X_t)/h) = A_0 + A_1 X_t + A_2 X_t^2,
\]
where \( A_0 = a_0 + a_1 x/h + a_2 x^2/h^2 \)
\[
A_1 = -(a_1/h + 2a_2 x/h^2),
\]
\[
A_2 = a_2/h^2.
\]
Therefore,
\[
f_t = (1/h)^{1/2}[A_0 + A_1 X_t + A_2 X_t^2 - A_3],
\]
where \( A_3 = E K((x - X_t)/h) \)
\[
= (1/h)^{1/2}[A_1 X_t + A_2 X_t^2 - A_4],
\]
where \( A_4 = A_3 - A_0 \)
\[
= (1/h)^{1/2}[A_1 \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} + A_2 \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}^2 - A_4].
\]
(2.6)

Under the above representation we have the following theorem:

**Theorem 2.1.** For a stationary linear process \( \{X_t\} \) with \( E \varepsilon_t^4 < \infty \) let \( \alpha \) be a positive constant such that,
\[
\sum_{j=-N}^{N} | \alpha_j | \leq C(2^{-\alpha N})/\sqrt{N}
\]
(2.7)
and \( h_N \propto N^{-1/2} \) then,
\[
\left| P\left( \frac{\sqrt{Nh}}{\sigma} \right) \left( \frac{f_n - \hat{f}_n}{\sigma} \right) < \varepsilon \right) - \Phi(z) \right| \leq C \cdot (\log N/N)^{1/2}
\]
(2.8)
where
\[
\sigma^2 = Ef_t^2 + 2 \sum_{j=1}^{\infty} Ef_{t-j}^2
\]
(2.9)
and \( Ef_t^2 = f(x) \int K^2(u) du > 0 \) as \( N \to \infty \).

To prove the above theorem we require the following result.

**Result 2.1.** Suppose the conditions of theorem 2.1 hold. Then
\[
\sum_{j=-N}^{N} | Ef_{t-j} | \leq C \cdot h^{1/2}
\]
(2.10)

**Proof.** Let \( f_t \) be the distribution function of \( Y = X_{t-2} \) and \( F_t \) be the joint distribution function of \( X_t \) and
\[
Y_t = X_{t-2}, \quad \text{where } X_{t-2} = \sum_{j=0}^{t-2} \alpha_j \varepsilon_{t-j}
\]
\[
X_{t-2}^* = X_t - X_{t-2},
\]
Define
\[ I = E \left[ \frac{1}{h^d} K \left( \frac{x - X_1}{h} \right) K \left( \frac{x - X_t}{h} \right) \right] \]
\[ = \frac{1}{h^d} K \left( \frac{x - X_1}{h} \right) K \left( \frac{x - Y_2 - y}{h} \right) dF_t(Y_1, Y_2) dJ_t(y). \quad (2.11) \]

Since \( \varepsilon_t \)'s are iid and \( \phi_\varepsilon \) is the characteristic function of \( \varepsilon_t \), the characteristic function of \( X_{1-t} \) is
\[ E \left( e^{it^2 \sum_{j=0}^{t-1} \varepsilon_j} \right) = \prod_{j=0}^{t-2} E \left( e^{it^2 \varepsilon_j} \right) = \prod_{j=0}^{t-2} \phi_\varepsilon(\alpha_j \alpha) \]
and, therefore by (iii)
\[ J_t'(y) \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\phi_\varepsilon(u)| du < c \quad (2.12) \]
and
\[ |J_t''(y)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |u \phi_\varepsilon(u)| du < c \quad (2.13) \]

Now for some constant \( \delta > 0 \), we define \( A = \{ |Y_2| \leq \delta \} \). Hence we have
\[ I = I_1 + I_2, \quad (2.14) \]
where
\[ I_1 = \int_A \frac{1}{h^d} K \left( \frac{x - Y_1}{h} \right) dF_t(X_1, Y_2) \int_{-\infty}^{\infty} K(y) J_t'(x - y_2 - hy) dy \]
and
\[ I_2 = \int_{A^C} \frac{1}{h^d} K \left( \frac{x - Y_1}{h} \right) dF_t(X_1, Y_2) \int_{-\infty}^{\infty} K(y) J_t'(x - y_2 - hy) dy \]

Again in \( A \), by (2.13) we have
\[ m_N \leq \int K(y) J_t'(x - y_2 - hy) dy \leq M_N \quad (2.15) \]
where
\[ m_N = \min \left( C, \int_{-\infty}^{\infty} K(y) J_t'(x - hy) + C \delta \right) dy \]
\[ M_N = \int_{-\infty}^{\infty} K(y) \max \{0, J_t'(x - hy) - C \delta\} dy. \]

Obviously \( m_N \geq 0, M_N \geq 0 \) and
\[ M_N - m_N \leq \int K(y) \{ J_t'(x - hy) + C \delta \} dy + 2C \delta \int_{B^c} K(y) dy \]
where \( B = \{ y, J_t'(x - hy) \leq C \delta \} \). Therefore
\[ 0 < M_N - m_N \leq C \delta \quad (2.16) \]
and
\[ I_1 \leq m_N \frac{1}{h^d} K \left( \frac{x - Y_1}{h} \right) dF_t(Y_1, Y_2) = M_N \bar{f}_N(x). \quad (2.17) \]
Defining $G_t = P(|X_{t-2,t}^*| > \delta)$ we have

$$I = I_1 + I_2 \leq M_N \bar{f}_N(x) + (C/h) G_t .$$

(2.19)

Again

$$I \geq m_N \int_{\mathbb{R}} 1/h K((x - y)/h) dF_t(y) \geq m_N \bar{f}_N(x) = (C/h) G_t .$$

(2.20)

Therefore, by (2.19) and (2.20), we have,

$$m_N \bar{f}_N(x) = (C/h) G_t \leq I \leq M_N \bar{f}_N(x) + (C/h) G_t .$$

(2.21)

Denote $V_t$ as the distribution function of $X_{t-2,t}^*$, then

$$\bar{f}_N(x) = \int_A + \int_{A^c} \frac{1}{h} K((x - y)/h) dV_t(y) = I_3 + I_4 ,$$

say

(2.22)

where

$$I_3 = \int_A dV_t(y_2) \int K((y_1 - y)/h) dV_t(y) \leq M_N \int_A dV_t(y) ,$$

by (2.15)

$$\leq C \cdot M_N$$

and

$$I_4 \leq C \cdot \int dV_t(y_2) \leq C \cdot G_t .$$

Therefore

$$\bar{f}_N(x) = I_3 + I_4 \leq C \cdot (M_N + G_t) .$$

(2.23)

Also

$$\bar{f}_N(x) \geq I_3 \geq m_N \int_A dV_t(y_2) = m_N (1 - G_t) .$$

(2.24)

Combining (2.23) and (2.24) we have,

$$m_N (1 - G_t) \leq \bar{f}_N(x) \leq C (M_N + G_t) .$$

This implies

$$m_N \bar{f}_N(x) (1 - G_t) \leq \bar{f}_N(x) \leq C \bar{f}_N(x) (M_N + G_t) .$$

(2.25)

Hence from (2.21) and (2.25) we have,

$$m_N \bar{f}_N(x) - (C/h) G_t \leq \bar{f}_N(x) (M_N + G_t) \leq I - \bar{f}_N(x) \leq M_N \bar{f}_N(x) + (C/h) G_t .$$

(2.26)
Since \( M_N, m_N, \tilde{f}_N(x) < C \) and \( h = h_N \to 0 \) as \( N \to \infty \) for all sufficiently large \( N \), we have

\[
|E(f_1f_2)| = \frac{1}{h} \left[ E \{K((x-X_1)/h) \right.

\[ - EK((x-X_1)/h) \} \{K((x-X_2)/h) - EK((x-X_2)/h)\} \left. \} \right]
\]

\[
\frac{1}{h} \left[ E \{K((x-X_1)/h) \right.

\[ - EK((x-X_1)/h) \} \{K((x-X_2)/h) - EK((x-X_2)/h)\} \left. \} \right],
\]

because of stationarity.

\[
= h |I - \tilde{f}_N(x)|
\]

\[
\leq C h \left[ f_N(M_N - m_N) + G_t/h \right]
\]

\[
\leq C \cdot (h \delta + G_t).
\]  

(2.27)

Also, it is easy to prove that

\[
G_t = P( |X^*_{t-1} | > \delta)
\]

\[
\leq E |X^*_{t-1} | / \delta
\]

\[
\leq C \cdot \sum_{j=t-1}^{\infty} |\alpha_j| / \delta, \quad \text{by A2}
\]

\[
= C \cdot \gamma_t / \delta \quad \text{where} \quad \gamma_t = \sum_{j=t-1}^{\infty} |\alpha_j|.
\]

Hence from (2.27), we have

\[
|E(f_1f_2)| \leq C \cdot (h \delta + \gamma_t / \delta)
\]

\[
= C \cdot h^{1/2} (\gamma_t)^{1/2}, \quad \text{selecting} \quad \delta = h^{-1/2} (\gamma_t)^{1/2}.
\]  

(2.28)

Now (2.10) is true because of the fact that

\[
\sum_{t=2}^{\infty} (\gamma_t)^{1/2} = \sum_{t=2}^{\infty} \left( \sum_{j=t-1}^{\infty} |\alpha_j| \right)^{1/2}
\]

\[
\leq C \cdot \sum_{t=1}^{\infty} (2^{-\alpha}(t-1)/(t-1)^{1/2})^{1/2}
\]

since by (v),

\[
\sum_{j=N}^{\infty} |\alpha_j| \leq C \cdot 2^{-\alpha N} / (N)^{1/2}, \quad \alpha > 0
\]

\[
= C \sum_{n=1}^{\infty} (2^{n \alpha /2} \cdot n^{1/4})^{-1}
\]

\[
< \infty.
\]
Rate of convergence for non parametric density estimation in linear process

PROOF OF THEOREM 2.1. From (2.6) we have,

\[ f_1 = \frac{1}{\sqrt{h}} \left[ A_1 \sum_{j=0}^{\infty} \alpha_j \varepsilon_{-1-j} + A_2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \alpha_i \alpha_j \varepsilon_{-1-i} \varepsilon_{-1-j} - A_4 \right] \]  

(2.29)

and

\[
E[ f_1 / \varepsilon_{-1}, \varepsilon_{-2}, \ldots, \varepsilon_{-N}] 
= E \left[ (A_1 / \sqrt{h}) \cdot \sum_{j=0}^{N-1} \alpha_j \varepsilon_{-1-j} + (A_2 / \sqrt{h}) \cdot \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_i \alpha_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right] 
\]

\[
+ (A_1 / \sqrt{h}) \cdot \sum_{j=0}^{N-1} \alpha_j \varepsilon_{-1-j} + (A_2 / \sqrt{h}) \cdot \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_i \alpha_j \varepsilon_{-1-i} \varepsilon_{-1-j} + (A_2 / \sqrt{h}) \cdot \sum_{i=0}^{N-1} \alpha_i^2 - A_4 .
\]

We set \( S_{ij} = \varepsilon_{-1-i} \varepsilon_{-1-j} - E(\varepsilon_{-1-i} \varepsilon_{-1-j}) \). Following Blume and Wittwer [2] and using the inequality

\[
E^{1/3} \left[ X + Y + Z \right] \leq E^{1/3} |X|^3 + E^{1/3} |Y|^3 + E^{1/3} |Z|^3 .
\]

We get,

\[
E^{1/3} \left[ f_1 - E \left[ f_1 / \varepsilon_{-1}, \ldots, \varepsilon_{-N} \right] \right] \leq E^{1/3} \left( \frac{A_1}{\sqrt{h}} \sum_{i=N}^{\infty} \alpha_i \right) \leq E^{1/3} \left( \frac{A_2}{\sqrt{h}} \sum_{i=N}^{\infty} \sum_{j=N}^{\infty} \alpha_i \alpha_j S_{ij} \right)^{3/2} ,
\]

(2.30)

Further using the inequality \( E |X Y Z| \leq [E |X|^3 |Y|^3 |Z|^3]^{1/3} \) we see that, the first summand is not greater than

\[
\frac{A_1}{\sqrt{h}} E^{1/3} \left( \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right)^{1/3} \]

\[
\leq \frac{A_1}{\sqrt{h}} \left( \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j \alpha_{i+j} \alpha_{i+j} \varepsilon_{-1-i} \varepsilon_{-1-j} \right)^{1/3} \]

(2.31)

and also

\[
E^{1/3} |\varepsilon_{-1-i} | \leq E^{1/3} |\varepsilon_{-1} |^{1/3} .
\]

(2.32)

Similarly, the second summand is not greater than

\[
\frac{A_2}{\sqrt{h}} E^{1/3} \left( \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j \alpha_{i+j} \alpha_{i+j} \varepsilon_{-1-i} \varepsilon_{-1-j} \right)^{1/3} \]

\[
\leq \frac{A_2}{\sqrt{h}} \left( \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j \alpha_{i+j} \alpha_{i+j} \right) E \left( \sum_{i=0}^{N} \sum_{j=0}^{N} S_{i+j} S_{i+j} \right)^{1/3} .
\]

(2.33)

An analogous inequality is also valid for the 3rd summand. Further we obtain,

\[
E^{1/3} |S_{ij}| \leq E^{1/3} |\varepsilon_{-1-i} \varepsilon_{-1-j}| + 1 \leq E^{1/3} (\varepsilon_0^3) + 1 .
\]

(2.34)

Therefore, combining (2.31)-(2.34) we get, from (2.30), the following:
$$E^{1/3}(|f_1 - E(f_1 | \varepsilon_{-1}, \ldots, \varepsilon_{-N})|^3)$$

\[ \leq \frac{A_1}{\sqrt{h}} \sum_{i=0}^{N} |\alpha_i| + \frac{A_2}{\sqrt{h}} (\sum_{i=0}^{N} |\alpha_i| + \sum_{i=0}^{N} |\alpha_i| + \sum_{i=0}^{N} |\alpha_i|) (E^{1/3}(\varepsilon_i^0) + 1) \]

\[ \leq C \cdot \frac{A_1}{\sqrt{h}} \frac{2^{-nN}}{\sqrt{N}} E^{1/3}(\varepsilon_i^0) + \frac{A_2}{\sqrt{h}} (E^{1/3}(\varepsilon_i^0) + 1) \left( C \cdot \frac{2^{-nN}}{\sqrt{N}} + C \cdot \frac{2^{-nN}}{\sqrt{N}} \right) \sum_{i=0}^{N} |\alpha_i| , \]

since by (vi), \( \sum_{i=0}^{N} |\alpha_i| \leq C \cdot 2^{-nN} \).

\[ \leq \frac{C}{h^{N/2}} \left[ E^{1/3}(\varepsilon_i^0) + 2(E^{1/3}(\varepsilon_i^0) + 1) \sum_{i=0}^{N} |\alpha_i| \right] \cdot 2^{-aN} \frac{2^{-nN}}{\sqrt{N}} \]

since \( A_1 \leq C/h^2 \) and \( A_2 \leq C/h^2 \)

\[ \leq C \cdot 2^{-aN} , \]

since \( h \geq N^{-1/5} \) and \( E(\varepsilon_i^0) < \infty, E(\varepsilon_i^0) < \infty, \sum_{i=0}^{N} |\alpha_i| < \infty. \)

Hence the condition (1.4) of Ibragimov [4] Theorem 2 is fulfilled for \( \delta = 1. \)

Also

\[ Ef_1^2 = \frac{1}{h} E K^2(x - X_1)/h - h \vec{f}_N^2(x) , \quad \text{where} \quad \vec{f}_N^2(x) = (1/h) E K((x - X_1)/h) \]

\[ = f(x) \int K^2(u) du - h \vec{f}_N^2(x) \]

\[ - f(x) \int K^2(u) du > 0 , \quad \text{as} \quad N \to \infty . \]

Therefore using (2.10) of Result 2.1 we have

\[ \sigma^2 = Ef_1^2 + 2 \sum_{j=2}^{N} Ef_1 f_j \neq 0 \quad \text{and also} \quad \sigma^3 < \infty . \]

Hence the condition 1.5 of Ibragimov [4], Theorem 2 is also fulfilled. So by Ibragimov's Theorem 2 we get, for \( \delta = 1, \)

\[ P \left( \frac{1}{\sigma \sqrt{N}} \sum_{i=1}^{N} f_i < z \right) - \Phi(z) \leq C \cdot \left( \frac{\log N}{N} \right)^{1/2} . \]

This implies

\[ P \left( \frac{\sqrt{Nh}}{\sigma} \left| f_N - \vec{f}_N \right| < z \right) - \Phi(z) \leq C \cdot \left( \frac{\log N}{N} \right)^{1/2} , \]

which is (2.8). Hence the theorem is proved.

3. Estimation of Joint Density and Its Rate of Convergence to Normality

We consider the covariance of estimates defined by (2.2) at two distinct points \( x \) and \( y \) as follows:
\[\text{Cov}(\hat{f}_N(x), \hat{f}_N(y)) = \frac{1}{N h^2} \text{Cov}(K( (x - X)/h), K( (y - X)/h)) \]

\[= \frac{1}{N h^2} \left[ \int K( (x - X)/h)K( (y - X)/h)f(X) dX - \left( \int K( (x - X)/h)f(X) dX \right)^2 \right] \]

\[= \frac{1}{N h} \left[ \int K(u)K( (y - x)/h + u)f(x - hu) du - h \left( \int K(u)f(x - hu) du \right)^2 \right]. \]

From this it follows that \(\sqrt{Nh}[\hat{f}_N(x) - Ef_N(x)]\) and \(\sqrt{Nh}[\hat{f}_N(y) - Ef_N(y)]\) are asymptotically independent.

We assume that all joint distributions of a finite number of distinct \(X_t\)'s (where \(X_t\)'s are as defined in (2.1)) are absolutely continuous with uniform bounded continuous density functions. Let \(f(x, y)\) be the joint density function of \(X_t, X_{t+1}.\) A natural estimate of \(f(x, y)\), used by Rosenblatt [8] is given by

\[\hat{f}_N(x, y) = \frac{1}{N h^2} \sum_{t=1}^{N-1} K( (x - X_t)/h, (y - X_{t+1})/h) \] (3.1)

where \(K(x, y)\) is a bounded continuous density function. We set

\[f_t = \frac{1}{h} \left[ K( (x - X_t)/h, (y - X_{t+1})/h) - EK( (x - X_t)/h, (y - X_t)/h) \right] = f_1 - f_2 - f_3 - \ldots\] (3.2)

and

\[T_N = (Nh^2)^{-1/2}[\hat{f}_N - \hat{f}_N], \quad \text{where } \hat{f}_N = \frac{1}{N h^2} E \left[ K( (x - X_t)/h, (y - X_t)/h) \right] \]

\[= (Nh^2)^{-1/2} \sum_{t=1}^{N-1} \left[ K( (x - X_t)/h, (y - X_{t+1})/h) - \frac{N}{N-1} EK( (x - X_t)/h, (y - X_t)/h) \right] \]

\[\sim (Nh^2)^{-1/2} \sum_{t=1}^{N} \left[ K( (x - X_t)/h, (y - X_{t+1})/h) - EK( (x - X_t)/h, (y - X_t)/h) \right] \]

\[= (Nh^2)^{-1/2} \sum_{t=1}^{N-1} \left[ K( (x - X_t)/h, (y - X_{t+1})/h) - EK( (x - X_t)/h, (y - X_t)/h) \right] \]

\[= N^{-1/2} \sum_{t=1}^{N-1} f_t. \] (3.3)

In case the strict stationarity of \(\{X_t\}\) is valid then

\[P \left[ N^{-1/2} \sum_{t=N}^{N} f_t \leq z \right] = P \left[ N^{-1/2} \sum_{t=1}^{N} f_t \leq z \right]. \]

Like in section 2 we take a second degree polynomial in two variables as Kernel \(K(x, y) = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy.\) Such bivariate Kernels are again optimal in the sense that it minimizes integrated mean square error and widely been used in the literature. Therefore
\[ K_t(x_i - x_{i+1})/h, (y_i - x_{i+1})/h) \\
= A_0 + A_1 x_i + A_2 x_{i+1} + A_3 x_i^2 + A_4 x_{i+1} + A_3 x_i x_{i+1} \]
where \( A_i \leq C/h^2 \) for \( i = 0, 1, \ldots, 5 \). We write \( A_i = EK((x_i - x_{i+1})/h, (y_i - x_{i+1})/h) \) and \( A_7 = A_6 - A_5 \).

Therefore
\[
\begin{align*}
\frac{1}{h} \left[ A_1 x_i + A_2 x_{i+1} + A_3 x_i^2 + A_4 x_{i+1} + A_3 x_i x_{i+1} - A_7 \right] \\
= \frac{1}{h} \left[ A_1 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} + A_2 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} + A_3 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} \right] \\
+ A_3 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} - A_7 \right]. \quad (3.4)
\end{align*}
\]

Now under this representation we have the following theorem:

**Theorem 3.1.** If the conditions of Theorem 2.1 together with the assumption
\[
\max \left[ \sum_{i=N-1}^{N-2} |a_i|, \sqrt[3]{\sum_{i=N-1}^{N-2} a_i^3}, \sqrt[3]{\sum_{i=N-1}^{N-2} a_i^3} \right] \leq C \cdot 2^{-a \sqrt{N}} \quad (3.5)
\]
holds, then for normalised deviation \( \frac{\sqrt{N}h^2}{\sigma} [f_N(x, y) - f(x, y)] \)
\[
\left| P \left( \frac{\sqrt{N}h^2 [f_N(x, y) - f(x, y)]}{\sigma} < z \right) - \Phi(z) \right| \leq C \left( \frac{\log N}{N} \right)^{1/2}
\]
holds (3.5),

where
\[
\sigma^2 = E[f_1^2] + 2 \sum_{j=2}^{N} E(f_1 f_j) \quad \text{and} \quad E[f_1^2] = E(K(u, v)) du dv \quad \text{as} \quad N \to \infty. \quad (3.6)
\]

**Proof.** Proceeding exactly similar way as in the proof of Theorem 2.1 we shall verify the conditions of Ibragimov’s Theorem. We note that,
\[
f_i = \frac{1}{h} \left[ A_1 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} + A_2 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} + A_3 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} \right] \\
+ A_3 \sum_{j=0}^{\infty} a_j \varepsilon_{-j} - A_7 \right].
\]

Therefore, for \( N \geq 2 \),
\[
E[f_i / \varepsilon_{-1}, \varepsilon_{-2}, \ldots, \varepsilon_{-(N-1)}] \]
\[
= E \left[ \frac{1}{h} \sum_{j=0}^{N-3} a_j \varepsilon_{-j} + \frac{A_2}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} \right] \\
+ \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} \right] \\
+ \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} \right] \\
+ \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} \right] \\
+ \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-j} \right] - A_7 \right]
\]
\[
\begin{aligned}
&= \left( \frac{A_1}{h} \sum_{j=0}^{N-2} a_j \varepsilon_{-1-j} + \frac{A_2}{h} \sum_{j=0}^{N-1} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{j=0}^{N-2} \sum_{t=0}^{N-2} a_i a_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right) \\
&\quad + \left( \frac{A_1}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} \alpha_i a_j \varepsilon_{-i} \varepsilon_{-j} + \frac{A_3}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} a_i a_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right) \\
&\quad + \left( \frac{A_2}{h} \sum_{j=0}^{N-1} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i a_{i+1} \right) - A_2.
\end{aligned}
\]

Therefore,
\[
\begin{aligned}
f_1 - E[f_1/\varepsilon_0, \varepsilon_{-1}, \ldots, \varepsilon_{-(N-1)}] \\
&= \left( \frac{A_1}{h} \sum_{j=0}^{N-1} a_j \varepsilon_{-1-j} + \frac{A_2}{h} \sum_{j=0}^{N-1} a_j \varepsilon_{-j} + \frac{A_3}{h} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_i a_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right) \\
&\quad + \left( \frac{A_1}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} \alpha_i a_j \varepsilon_{-i} \varepsilon_{-j} + \frac{A_3}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} a_i a_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right) \\
&\quad + \left( \frac{A_2}{h} \sum_{j=0}^{N-1} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i a_{i+1} \right) - A_2.
\end{aligned}
\]

Now following Blume and Wittwer [2] and Kersten [5] and by using the inequality
\[
E^{1/3}[|X+Y+Z+\cdots|^3] \leq E^{1/3}[|X|^3 + E^{1/3}[|Y|^3 + E^{1/3}[|Z|^3 + \cdots]]
\]
we get,
\[
\begin{aligned}
E^{1/3}(|f_1 - E[f_1/\varepsilon_0, \varepsilon_{-1}, \ldots, \varepsilon_{-(N-1)}]|) \\
&\leq E^{1/3}(\left( \frac{A_1}{h} \sum_{j=0}^{N-1} a_j \varepsilon_{-1-j} \right)^3) + E^{1/3}(\left( \frac{A_2}{h} \sum_{j=0}^{N-1} a_j \varepsilon_{-j} \right)^3) \\
&\quad + E^{1/3}(\left( \frac{A_3}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} a_i a_j \varepsilon_{-1-i} \varepsilon_{-1-j} \right)^3) \\
&\quad + E^{1/3}(\left( \frac{A_1}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} \alpha_i a_j \varepsilon_{-i} \varepsilon_{-j} \right)^3) \\
&\quad + E^{1/3}(\left( \frac{A_2}{h} \sum_{i=0}^{N-2} \sum_{j=0}^{N-1} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i^2 + \frac{A_3}{h} \sum_{i=0}^{N-2} \alpha_i a_{i+1} \right)^3)
\end{aligned}
\]
where
\[
T_{ij} = \varepsilon_{-1-i} \varepsilon_{-j} = e_1 + e_2 + e_3 + e_4 + e_5 + e_6 + e_7, \text{ say.}
\]
Now
\[
\begin{aligned}
e_1 &\leq \frac{A_1}{h} E^{1/3}(\sum_{j_1,j_2,j_3=0}^{N-1} |a_{j_1} a_{j_2} a_{j_3} \varepsilon_{-1-j_1} \varepsilon_{-1-j_2} \varepsilon_{-1-j_3}|) \\
&\leq \frac{A_1}{h} \left( \sum_{j_1,j_2,j_3=0}^{N-1} |a_{j_1} a_{j_2} a_{j_3}| \left| E[\varepsilon_{-1-j_1}] E[\varepsilon_{-1-j_2}] E[\varepsilon_{-1-j_3}] \right|^{1/2} \right)^{1/2}
\end{aligned}
\]
using the inequality
\[ E |XYZ| \leq |E| |X||E| |Y||E| |Z||E| \leq \frac{A_1}{h} \sum_{j=N-1}^{\infty} |\alpha_j| E^{1/3} |\varepsilon|^3. \]

Similarly
\[ e_3 \leq (A_3/h) \sum_{j=N}^{\infty} |\alpha_j| E^{1/3} (\varepsilon_0^3). \]

Also
\[ E \left( \sum_{i=N-1}^{\infty} \sum_{j=0}^{N-1} \left| \alpha_i \alpha_j (\varepsilon_{i,j}) \right| \right) \]
\[ \leq \frac{A_3}{h} \left( \sum_{i=0}^{\infty} \sum_{j=0}^{N-1} \left| \alpha_i \alpha_j (\varepsilon_{i,j}) \right| \right) \]
\[ \leq \frac{A_3}{h} \left( \sum_{i=0}^{\infty} \sum_{j=0}^{N-1} \left| \alpha_i \alpha_j (\varepsilon_{i,j}) \right| \right) \]
\[ = \frac{A_3}{h} \left( \sum_{i=0}^{\infty} \sum_{j=0}^{N-1} |\alpha_i| |\alpha_j| \right) E^{1/3} (\varepsilon_0^3). \]

Similarly
\[ e_4 \leq \left( \frac{A_4}{h} \sum_{i=0}^{\infty} |\alpha_i| \right) E^{1/3} (\varepsilon_0^3) \]
\[ e_5 \leq \left( \frac{A_5}{h} \sum_{i=0}^{\infty} |\alpha_i| \right) E^{1/3} (\varepsilon_0^3) \]

and
\[ e_6 \leq \left( \frac{A_6}{h} \sum_{i=0}^{\infty} |\alpha_i| \right) E^{1/3} (\varepsilon_0^3). \]

Also following Kersten [5] we get
\[ e_4 \leq \frac{A_3}{h} \left[ E (\varepsilon_0^3) \sum_{i=0}^{\infty} \alpha_i^4 + 4 \left( \sum_{i=0}^{\infty} \alpha_i \right) \left( \sum_{i=0}^{\infty} \alpha_i^4 \right) \right] \]
\[ \leq \frac{A_3}{h} \left( E (\varepsilon_0^3) + 4 \sum_{i=0}^{\infty} \alpha_i^4 \right) \cdot M \]

where \( M = \max \left\{ \sum_{i=N-1}^{\infty} \alpha_i^4, \sum_{i=N-1}^{\infty} \alpha_i^4 \right\} \). This implies
\[ e_4 \leq C A_3 \frac{h}{A_5} M^{1/3}, \text{ because } \sum_{i=0}^{\infty} |\alpha_i| < \infty \implies \sum_{i=0}^{\infty} \alpha_i^4 < \infty. \]

Hence
\[ E^{1/3} (|f_1 - E[f_1/\varepsilon_0, \ldots, \varepsilon_{(N-1)}]|^3) \]
\[ \leq \frac{A_1}{h} \sum_{i=N-1}^{\infty} |\alpha_i| E^{1/3} (\varepsilon_0^3) + \frac{A_2}{h} \sum_{j=N}^{\infty} |\alpha_j| E^{1/3} (\varepsilon_0^3) \]
\[ + \frac{A_3}{h} \left( \sum_{i=0}^{\infty} |\alpha_i| \right) \left( \sum_{j=0}^{\infty} |\alpha_j| \right) E^{1/3} (\varepsilon_0^3) \]
\[ + \frac{A_4}{h} \left( \sum_{i=0}^{\infty} |\alpha_i| \right) \left( \sum_{j=0}^{\infty} |\alpha_j| \right) E^{1/3} (\varepsilon_0^3) \]
So the condition (1.4) of Ibragimov [4] Theorem 2 is fulfilled for δ = 1. Proceeding exactly similar way as in Result 2.1 we get \( \sum_{i=1}^{N} E(f_1 f_j) \to 0 \) as \( N \to \infty \). Also \( E|f|^2 \to f(x, y)\int K(u, v)du dv > 0 \). Hence \( \sigma^2 \neq 0 \). So the condition (1.5) of Ibragimov [4] Theorem 2 is also fulfilled and hence the theorem follows immediately.

### 4. Estimation of Probability Density at Several Points and Its Rate of Convergence to Normality

We consider the rate of convergence of a density estimate at several points \( x_1, x_2, \cdots, x_m \). Assume that \( K \) is a continuous density function with \( K(u) = o(\frac{1}{|u|}) \) as \( |u| \to \infty \). The density ‘\( f \)’ is taken to be bounded and continuous. Rosenblatt [9] considers the asymptotic distribution of

\[
\sqrt{Nh}\left[ f^*(x_p) - E f^*(x_p) \right], \quad p=1, 2, \cdots, m
\]

for any finite \( m \)-tuple of points \( x_p \). He shows that these random variables are asymptotically jointly normal and independent with means zero and variances given by (4.1) below:

\[
\lim_{h \to 0} N h \text{Cov} \left[ f^*_N(x_p), f^*_N(x_q) \right] = \begin{cases} f(x) \int K^2(u)du & \text{if } p=q \\ 0 & \text{if } p \neq q. \end{cases}
\]

(4.1)

Our object is to consider the rate of convergence of the normalized deviations

\[
\sqrt{Nh}\left[ f_N(x_p) - \tilde{f}_N(x_p) \right], \quad p=1, 2, \cdots, m.
\]

where

\[
\tilde{f}_N(x_p) = \frac{1}{h} E K \left( \frac{x_p - X_1}{h} \right).
\]

To do this we require the lemma-1 of Kersten [5], (due to Gabasov (1977)) which is stated below:
Let \( \{s_t\}_{t=0,1,...} \) be a stationary sequence of random variables and let \( f_p(s_t, s_{t+1}, ..., s_N) \), \( p=1, 2, ..., m \) be random variables with \( Ef_p=0 \). The random vector \( f_t \) is defined by

\[
\begin{align*}
&f_t = (f_{st}, f_{st+1}, ..., f_{st+m})^T \\
&f_{st} = f_p(s_t, s_{t+1}, ..., s_N),
\end{align*}
\]

By \( \Phi \) we denote the distribution function of a \( m \)-dimensional centered Gaussian distribution \( Z \) with the covariance matrix \( V \). With these denotations we have the following:

**LEMMA.** Let the following conditions be fulfilled with the positive constant \( C_1 \) and \( C_2 \) and any positive and constant number \( u \)

(i) \( E|f_t|^u < \infty \)

(ii) \( E|f_{st}-E(f_{st}/s_t, ..., s_N)|^2 \leq C_1 N^{-2u} \) for \( p=1, 2, ..., m \) and \( N=1, 2, ... \)

(iii) The sequence \( \{s_t\} \) fulfills the uniformly strongly mixing condition with the coefficients \( \phi_j \leq C_2 j^{-u} \), for \( j=0, 1, 2, ... \)

(iv) \( \det V \neq 0 \) where \( V=(v_{ij})_{j=1, ..., m} \) with

\[
v_{ij} = \lim_{N \to \infty} \frac{1}{N} E(\sum_{t=1}^{N} f_{st} f_{st+j}).
\]

Then the inequality

\[
\sup_{B} P\left( \frac{1}{\sqrt{N}} \sum_{t=1}^{N} f_t \in B \right) \leq C_3 N^{-1/2}
\]

is valid where \( B \) is a measurable convex set of the \( m \)-dimensional Euclidean space \( R^m \), \( \Phi_1[B] = P(Z \in B) \) and \( \varepsilon = 3u^{-1/2} \). The constant \( C_3 \) depends on the dimension \( m \), on the constants \( C_1, C_2 \), on \( u \), and on the smallest eigenvalue \( \lambda_{\text{min}} \) of \( V \). We note that

\[
f_{st}(x_p) = \frac{1}{N h} \sum_{t=1}^{N} K\left(\frac{x_p - x_{s_t}}{h}\right), \ p=1, 2, ..., m.
\]

As in section 2 we set,

\[
f_{st} = \frac{1}{\sqrt{N}} \left[ K\left(\frac{x_p - x_{s_t}}{h}\right) - E K\left(\frac{x_p - x_{s_t}}{h}\right) \right] = f_p(s_{t-1}, s_{t-1}, ...)
\]

and

\[
T_{st} = \frac{1}{h} E K\left(\frac{x_p - x_{s_t}}{h}\right) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} f_{st}.
\]

Also

\[
\sqrt{Nh}(f_{N} - f_{st}) = \sqrt{Nh}\left[ (f_{st}(x_1) - f_{N}(x_1)), ..., (f_{st}(x_p) - f_{N}(x_p)) \right]^T
\]

where \( f_t = (f_{st}, ..., f_{st+m})^T \).

It may be noted that for the strict stationarity of \( \{X_t\} \),

\[
P\left( \frac{1}{N} \sum_{t=1}^{N} f_{st} < z \right) = P\left[ \left( \frac{1}{N} \sum_{t=1}^{N} f_{st} \right)^2 < \frac{1}{N} \sum_{t=1}^{N} f_{st} < z \right].
\]

As in section 2 taking the Kernel to be a polynomial of degree two one can write

\[
f_{st} = \frac{1}{\sqrt{h}} \left[ A_0 + A_1 X_{s_t} + A_2 X_{s_t}^2 - A_3 \right]
\]
where
\[ A_0 = a_0 + a_1 x_0 / h + a_2 x_0^2 / h^2, \quad A_1 = -(a_1 / h + 2a_2 x_0 / h^2), \]
\[ A_2 = a_2 / h^2, \quad A_3 = EK ((x_0 - X_1) / h). \]

Therefore,
\[ f_{pt} = \frac{1}{\sqrt{h}} \left[ A_1 \sum_{j=0}^{\infty} \alpha_j \delta_{t-j} + A_2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_k \alpha_j \delta_{t-j-k} - A_4 \right] \]

where \( A_4 = A_3 - A_2 \). Under this representation we have the following:

**Theorem 4.1.** For a stationary linear process \( \{X_t\} \) satisfying

(i) \( E |\epsilon_t|^u < \infty \) for a fixed number \( u \geq 2 \)

(ii) \( \sum_{i=1}^{\infty} |\alpha_i| \leq K_1 N^{-\left(\frac{u+1}{2}\right)} \), \( h = O(N^{-\frac{1}{2}}) \)

(iii) \( \det V \neq 0 \) where \( V = (v_{ij})_{i=1, \ldots, m} \) with
\[ v_{ij} = \lim_{N \to \infty} Nh \text{Cov}(f_N(x_i), f_N(x_j)) \]

then
\[ \sup_B \left| P \{ (Nh)^{1/2} (f_N - \bar{f}_N) \in B \} - \Phi_B(B) \right| \leq K_2 N^{-1/2} \epsilon \]

where \( B \) and \( \epsilon \) are defined as above, and the constant \( K_2 \) depends on \( U, u, K_1, \text{Var } X_t \) and the smallest eigen value \( \lambda_{\min} \) of \( V \).

**Proof.** Proceeding exactly similar way as in section 2 we see that
\[ E[f_{pt} | \epsilon_{-1}, \epsilon_{-2}, \ldots, \epsilon_{-N}] = \frac{A_1}{\sqrt{h}} \sum_{j=0}^{\infty} \alpha_j \delta_{t-j} \]
\[ + \frac{A_2}{\sqrt{h}} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_k \alpha_j \delta_{t-j-k} - A_4 \]

and
\[ E^{1/2} \left( |f_{pt} - E(f_{pt} | \epsilon_{-1}, \epsilon_{-2}, \ldots, \epsilon_{-N})|^2 \right) \]
\[ \leq \frac{A_1}{\sqrt{h}} \sum_{j=0}^{\infty} |\alpha_j| + 2C \frac{A_2}{\sqrt{h}} \sum_{j=0}^{\infty} |\alpha_j| (E^{1/2} \epsilon)^1 + 1 \]
\[ \leq C \cdot N^{-u} \text{ for } u \geq 2 \text{ by (i) and (ii)}. \]

Therefore
\[ E(|f_{pt} - E(f_{pt} | \epsilon_{-1}, \epsilon_{-2}, \ldots, \epsilon_{-N})|^2) \leq C_1 N^{-2u}. \]

So condition (ii) of Lemma is satisfied. Condition (i) is also holds because of the assumption that \( E |\epsilon_t|^u < \infty \). Condition (iii) is also fulfilled with any \( u \) and \( C_2 \) because of the independence of \( \epsilon_t \). Now
\[ v_{ij} = \lim_{N \to \infty} Nh \text{Cov}(f_N(x_i), f_N(x_j)), \quad i \neq j \]
\[ = 0, \]

by (4.1). Also
\[ v_{ii} = \lim_{N \to \infty} Nh \text{Var}(f_N(x_i)) = f(x) \int K^2(u) du, \quad \text{for } i = 1, 2, \ldots, m. \]

Therefore, \( \det V \neq 0 \). So condition (iv) of Lemma is also satisfied and hence the theorem
follows by applying the Lemma.

References


Received June 15, 1988
Communicated by T. Kitagawa