Suppose that $X_1, \ldots, X_n$ are samples drawn from a population with density function $f$, and $f_n(x) = f(x; X_1, \ldots, X_n)$ is an estimate of $f(x)$. Denote by $m_{nr} = \int |f_n(x) - f(x)|^r dx$ and $M_{nr} = \mathrm{E}(m_{nr})$ the Integrated $r$-th Order Error and Mean Integrated $r$-th Order Error of $f$, for some $r > 1$. When $r = 2$, they are the familiar...
and widely studied ISE and MISE). In this paper the same necessary and sufficient condition for \( \lim_{n \to \infty} M_n = 0 \) and \( \lim_{n \to \infty} m_n = 0 \) a.s. is obtained when \( f(x) \) is the ordinary histogram estimator.
NECESSARY AND SUFFICIENT CONDITIONS FOR
THE CONVERGENCE OF INTEGRATED AND
MEAN-INTEGRATED r-th ORDER ERROR OF
HISTOGRAM DENSITY ESTIMATES*

X.R. Chen and L.C. Zhao
Center for Multivariate Analysis
University of Pittsburgh
NECESSARY AND SUFFICIENT CONDITIONS FOR
THE CONVERGENCE OF INTEGRATED AND
MEAN-INTEGRATED r-th ORDER ERROR OF
HISTOGRAM DENSITY ESTIMATES*

X.R. Chen and L.C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

April 1987

Technical Report No. 87-07

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* Research sponsored by the Air Force Office of Scientific Research under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
ABSTRACT

Suppose that \( X_1, \ldots, X_n \) are samples drawn from a population with density function \( f \), and \( f_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{X_k \leq x} \) is an estimate of \( f(x) \). Denote by \( m_{nr} = \int |f_n(x) - f(x)|^r \, dx \) the integrated r-th Order Error and Mean Integrated r-th Order Error of \( f_n \), for some \( r \geq 1 \) (when \( r = 2 \), they are the familiar and widely studied ISE and MISE). In this paper the same necessary and sufficient condition for \( \lim_{n \to \infty} m_{nr} = 0 \) and \( \lim_{n \to \infty} m_{nr} = 0 \) a.s. is obtained when \( f_n(x) \) is the ordinary histogram estimator.

AMS 1980 Subject Classification : Primary 62G05.

Key Words and Phrases : Consistency, Density Estimation, Histogram.
Suppose that \( X_1, \ldots, X_n \) are iid samples drawn from a \( d \)-dimensional population with density function \( f \). Let \( \hat{f}_n(x) = \hat{f}_n(x; X_1, \ldots, X_n) \) be an estimator of \( f(x) \). The Integrated Square Error (ISE) and Mean Integrated Square Error (MISE) of \( \hat{f}_n \) are defined by

\[
\text{ISE} = \int |\hat{f}_n(x) - f(x)|^2 dx,
\]
\[
\text{MISE} = E\int |\hat{f}_n(x) - f(x)|^2 dx
\]
respectively. They are important and widely used criteria in evaluating the performance of an estimator \( \hat{f}_n \). Quite a lot of works appeared in the statistical literature dealing with the asymptotic properties of them, for various types of estimators, such as kernel estimator, orthogonal series estimator, nearest neighbor estimator etc. For example, a much discussed problem is to find the conditions under which the assertions

\[
\lim_{n \to \infty} \text{MISE} = 0, \quad \lim_{n \to \infty} \text{ISE} = 0, \text{a.s.}
\]

are true. Various conditions have been proposed in the literature which are far from necessary.

In this paper we obtain the necessary and sufficient conditions of (1) for the histogram estimator. In fact, we have done more. For given constant \( r \geq 1 \), define the integrated \( r \)-th order error \( m_{nr} \) and mean integrated \( r \)-th order error \( M_{nr} \) by

\[
m_{nr} = \int |\hat{f}_n(x) - f(x)|^r dx, \quad M_{nr} = E(m_{nr})
\]

We obtain the necessary and sufficient conditions of \( M_{nr} \to 0 \) and \( m_{nr} \to 0 \), a.s. In the case of \( r=1 \), the problem has been considered by Abou-Jaoude[1], [7] (see also [4], pp 19-23).

A \( d \)-dimensional histogram is defined as follows. Choose \( a_n = (a_{n1}, \ldots, a_{nd}) \in \mathbb{R}^d \) with \( h_n > 0 \), \( i=1, \ldots, d \). Denote by \( k \) the number of those \( X_1, \ldots, X_n \) falling into the set \( A = \{ x=(x_{1i}, \ldots, x_{di}) \mid c_i h_n < x_{1i} < a_{ni} + (c_i + 1)h_n, i=1, \ldots, d \} \).
Define
\[ f_n(x) = k/(nh)_{i=1}^{d} h_{i}, \text{ when } x \in A_{n}(c_{1}, \ldots, c_{d}), \]
\[ c_{1}, \ldots, c_{d} = 0, \pm 1, \pm 2, \ldots \]
(2)

which is a histogram estimator of \( f(x) \).

**Theorem.** Suppose that
\[ \int f^{r}(x)dx < \infty, \text{ for some } r \geq 1 \]
(3)

\[ \lim_{n \to \infty} \max_{1 \leq i \leq d} h_{i} = 0 \]
(4)

\[ \lim_{n \to \infty} nh_{i} = \infty \]
(5)

Then for the histogram \( f_{n} \) defined by (2) we have
\[ \lim_{n \to \infty} n^{-1} m = 0, \text{ a.s.} \]
(6)

\[ \lim_{n \to \infty} M_{n} = 0, \]
(7)

Conversely, if (6) or (7) is true, then (3), (5) are true. Further, if \( f(x) \) does not have a form
\[ f(x) = \sum_{i_{1}, \ldots, i_{d} = -\infty}^{\infty} b_{i_{1}, \ldots, i_{d}} A(i_{1}, \ldots, i_{d}) (x) \]

Then (4) is also true. Here \( b_{i_{1}, \ldots, i_{d}} \) is a constant, \( A(i_{1}, \ldots, i_{d}) = \{ x = (x_{1}, \ldots, x_{d}) \}

\[ \ldots, x_{j} : a_{j0} + ih_{j0} \leq x_{j} < a_{j0} + (i + 1)h_{j0}, \text{ for } j = 1, \ldots, d \}, \text{ and } I_{B}(x) \]
denotes the indicator of \( B \).

**Proof.** For simplicity of writing, we shall give only the proof for the case of \( d = 1 \), as the general case involves no essential change. For \( d = 1 \) and writing \( h_{n} \) for \( h_{n} \), the conditions (4) and (5) reduce to
\[ \lim_{n \to \infty} n^{-1} m = 0 \]
(6)
\[ \lim_{n \to \infty} n h = \infty \]  
\hfill (5)

In the sequel we shall often write \( h \) for \( h_n \), \( a \) for \( a_n \). Note that we may assume \( |a| \leq h \) without loss of generality. Write
\[
\Delta_n = \Delta_n^h = [a + n h, a + (n + 1) h),
\]
\[
\mathcal{I}_n(a) = \mathcal{I}_n(a^h),
\]
\[
p_\Delta = \int f(x) \, dx
\]
Then \( E_n(x) = p_\Delta / h \), when \( x \in \Delta_n^h \). The symbol \( "C" \) will be used to denote any constant not depending on \( n \) (but may depend on \( r \)).

The proof will be divided into four parts.

(A) Sufficiency of (3)-(5) for (7)

First consider the case of \( r > 1 \). In order to prove (7), we need only to show that
\[
\int f_n(x) - Ef_n(x) \, dx = n^{-r} h^{-r+1} \sum_{i=1}^{n} \left( \mathcal{I}_n(x_i) - p_\Delta \right) \to 0
\]
\hfill (8)
\[
\int f(x) - Ef_n(x) \, dx \to 0
\]
\hfill (9)
as \( n \to \infty \) (9) is easy to prove: Suppose first that \( f \) is continuous on \((-\infty, \infty)\), then for any fixed constant \( A > 0 \),
\[
\int_{-A}^{A} f(x) - Ef_n(x) \, dx \to 0, \text{ as } n \to \infty.
\]
On the other hand, by Hölder inequality, for any fixed constant \( B > 0 \) we have
\[
\sum_{\Delta_n^h} \{ \int (Ef_n(x))^r \, dx : \Delta_n^h \subset [-B, B]^c \}
\]
\[
= \sum_{\Delta_n^h} \{ h^{-r} + \frac{1}{r} \left( \int f(x) \, dx \right)^r : \Delta_n^h \subset [-B, B]^c \}
\]
\[
\leq \sum_{\Delta_n^h} \{ \int f^r(x) \, dx : \Delta_n^h \subset [-B, B]^c \}
\].
Hence

\[
\sum \left\{ \int_{\Delta_{\lambda}} \left| f(x) - Ef_n(x) \right|^r dx : \Delta_{\lambda} \subseteq [-B,B]^c \right\} \leq 2^r \left\{ \int_{[-B,B]^c} f^r(x) dx + \sum_{\lambda} \left( \int_{\Delta_{\lambda}} \left( Ef_n(x) \right)^r dx : \Delta_{\lambda} \subseteq [-B,B]^c \right) \right\} \leq 2^r \int_{[-B,B]^c} f^r(x) dx
\]

Summing up the above arguments and noticing (3), we obtain (9). For the general case, choose a function \(g\) such that \(\int \left| f(x) - g(x) \right|^r dx \leq \text{some given } \epsilon > 0\). Define \(g_n(x) = h^{-1} \int_{\Delta_{\lambda}} g(x) dx\) for \(x \in \Delta_{\lambda}, \lambda = 0, \pm 1, \ldots\), an argument similar to that leading to (10) gives

\[
\int \left| Ef_n(x) - g_n(x) \right|^r dx \leq \int \left| f(x) - g(x) \right|^r dx \leq \epsilon
\]

From this and the fact that (9) is true when \(f\) is continuous, (9) follows easily.

For a proof of (8), put \(Y_{i\lambda} = 1_{\{X_i = \lambda\}} - p\lambda\), \(S_{\lambda} = \sum_{i=1}^j (1_{\{X_i = \lambda\}} - p\lambda)\), \(T_n = (\sum_{\lambda} |S_{\lambda}|^r)^{1/r}\). If \(1 < r \leq 2\), then from the inequality \(|1 + a|^r \leq 1 + ra + C|a|^r\) (a: real), we have

\[
|S_{\lambda}|^r \leq \left| S_{\lambda, n-1} \right|^r + C \left| Y_{\lambda} \right|^r + r \left| S_{\lambda, n-1} \right|^r \left| S_{\lambda, n-1} \right|^r
\]

Therefore,

\[
E \left| S_{\lambda} \right|^r \leq E \left| S_{\lambda, n-1} \right|^r + CE \left| Y_{\lambda} \right|^r \leq C \sum_{i=1}^n E \left| Y_{i\lambda} \right|^r \leq 2Cn p_{\lambda}
\]

which in turn implies

\[
n^{-r} \sup_{\lambda} \sum_{j=1}^n \left| f_j \right|^{r-1} \leq n^{-r} h^{-r+1} \sum_{\lambda} 2Cn p_{\lambda} = 2C (nh)^{-r+1} \to 0
\]

Suppose that (11) is true for \(r \in (1, m]\). We proceed to show that it is true for \(r \in [m, m+1]\). Since when \(r > 2\) we have the inequality

\[
|1 + a|^r \leq 1 + ra + Ca^2 (1 + |a|^{r-2}) (a: \text{ real})
\]

It follows that
\[ |S_{n\ell}|^r \leq |S_{n-1,\ell}|^r + r|S_{n-1,\ell}|^{-2}S_{n-1,\ell} \gamma \eta \ell \\
+ C|S_{n-1,\ell}|^{-2}\gamma \eta \ell^2 + C|\eta \ell|^r \]  \hspace{1cm} (12)

which in turn implies

\[ E|S_{n\ell}|^r \leq E|S_{n-1,\ell}|^r + CE|S_{n-1,\ell}|^{-2}EY_{n\ell}^2 + CE|\eta \ell|^r \]

Noticing that \( EY_{n\ell}^2 \leq p_{\ell} E|Y_{n\ell}|^r \leq p_{\ell} \) we have

\[ E|S_{n\ell}|^r \leq E|S_{n-1,\ell}|^r + C_\ell (E|S_{n-1,\ell}|^{-2}+1), \; \tilde{n} = n,n-1, \ldots \]

Therefore

\[ E|S_{n\ell}|^r \leq C_\ell \sum_{j=1}^{n-1} E|S_{j\ell}|^r - 2 + np_\ell C \] \hspace{1cm} (13)

\[ n^{-r} \sup_{\ell} \max \{ \sum_{j=1}^{\tilde{n}} E|S_{j\ell}|^r \} \leq (nh)^{-r+1} \max \sum_{\ell \leq n} p_{\ell} E|S_{j\ell}|^r - 2 + C(nh)^{-r+1} \] \hspace{1cm} (14)

Since \( f \) is a probability density and \( \int f^r(x)dx < \infty, \; r > 2 \), we have \( \int f^r(x)dx = C < \infty \)

Hence when \( r > 2 \)

\[ \int f^{r-1}(x)dx = C < \infty \] \hspace{1cm} (15)

By induction hypothesis

\[ n^{-r} h^{-r+1} \max \{ \sum_{j=1}^{\tilde{n}} E|S_{j\ell}|^r \} \xrightarrow{n \to \infty} 0, \; \text{as } n \to \infty \] \hspace{1cm} (16)

From (14)-(16), we have

\[ n^{-r} h^{-r+1} \max \{ \sum_{j=1}^{\tilde{n}} E|S_{j\ell}|^r \} \leq C(nh)^{-r+1} \left( \sum_{j=1}^{\tilde{n}} p_{\ell} E^{r-1} (\max_1^{\tilde{n}} E|S_{j\ell}|^r) / (r-1) \right) ^{(r-2)/(r-1)} \]

\[ + C(nh)^{-r+1} \]

\[ \leq C(nh)^{-r+1} \left( h (r-2)/(r-1) \right) \left( n^{-r} h (r-2)^2/(r-1) \right) \]

\[ (n^{-r} h^{-r+2} \max \{ \sum_{j=1}^{\tilde{n}} E|S_{j\ell}|^r \} / (r-1) + C(nh)^{-r+1} \]

\[ = C(nh)^{-r+1} \left( n^{-r} h^{-r+2} \max \{ \sum_{j=1}^{\tilde{n}} E|S_{j\ell}|^r \} / (r-1) + C(nh)^{-r+1} \right) \xrightarrow{n \to \infty} 0 \]

as \( n \to \infty \). This shows that (11) is true for \( r \in (m,m+1] \), concluding the proof of (11).
hence (7), for \( r > 1 \). When \( r = 1 \), (7) is a consequence of (6) for \( r = 1 \), which we are now going to prove.

(B) Sufficiency of (3)-(5) for (6)

Again first consider the case of \( r > 1 \). Define \( T_n \) as before, then

\[
\lim_{n \to \infty} T_n = 0, \quad \text{a.s.}
\]

Define

\[
\xi_{j,l} = \gamma_{j,l}, \quad \eta_{j,l} = |S_{j,l}|^{r-1} \text{sign}(S_{j,l}),
\]

\[
\xi_{j,l} = \gamma_{j,l}^{2} - \rho_{j,l}(1-\rho_{j,l}), \quad \eta_{j,l} = |S_{j,l}|^{r-1}
\]

and proceed to show that for any given \( \epsilon > 0 \)

\[
\omega \leq \frac{P(n^{-r}h^{-r+1} \sum_{2 \leq k \leq n} \xi_{j,l} \eta_{j,l} \geq \epsilon)}{\epsilon} \leq Cn^{-2}, \quad i = 1, 2
\]

In the following we shall write \( \xi_{j,l} \), \( \eta_{j,l} \) for \( \xi_{j,l} \), \( \eta_{j,l} \). Since \( \sum_{j=2}^{k} \xi_{j,l} \eta_{j-1,l} \), \( k = 2, 3, \ldots \) is a martingale sequence, we have

\[
J \leq \epsilon - 4n^{-4r}h^{-4r+4} C \left( \sum_{j=2}^{n} \left( \sum_{k} \xi_{j,l} \eta_{j-1,l}, k \right) \right)^{4}
\]

From an inequality of Rosenthal (see [5], p. 23),

\[
E \left( \sum_{j=2}^{n} \left( \sum_{k} \xi_{j,l} \eta_{j-1,l}, k \right)^{4} \right)
\]

\[
\leq C \left( \sum_{j=2}^{n} E \left( \left( \sum_{k} \xi_{j,l} \eta_{j-1,l}, k \right)^{2} \right) \right)^{2}
\]

\[
+ \sum_{j=2}^{n} E \left( \left( \sum_{k} \xi_{j,l} \eta_{j-1,l}, k \right)^{4} \right)
\]

(20)

Here \( F \) is the \( \sigma \)-field generated by \( X_1, \ldots, X_j \). Since \( \sum_{k} |\xi_{j,l}| \leq 3 \), by Jensen's inequality we have
It can be shown easily by using (13) that
\[
\mathbb{E}_{\eta_{j-1}, \mathbb{L}} \leq C (1 + (n \mathbb{P}_\mathbb{L})^{2r-2}) \leq C (1 + (n \mathbb{P}_\mathbb{L})^{2r-1})
\]
Noticing that \( \sum_{\mathbb{L}} \mathbb{P}_\mathbb{L}^{-n} \leq \int |f(x)|^n \, dx < \infty \) and \( nn \to \infty \), we have
\[
\sum_{\mathbb{L}} E \left( \sum_{j=2}^{n} \mathbb{P}_{\mathbb{L}} - n \mathbb{P}_{\mathbb{L}} \right)^2 \leq \int |f(x)|^n \, dx < \infty
\]
Therefore, from
\[
\mathbb{E}_{\eta_{j-1}, \mathbb{L}} \leq C (1 + (n \mathbb{P}_\mathbb{L})^{2r-2}) \leq C (1 + (n \mathbb{P}_\mathbb{L})^{2r-1})
\]
\[
E \left( \sum_{j=2}^{n} E \left[ \left( \sum_{k \in \mathcal{E}} (\xi_{j,\eta_{j-1},2})^2 \right)^2 \right] \right)^{1/2} \\
\leq C n \sum_{j=2}^{n} \sum_{\mathcal{E}} p_{\mathcal{E}} E \eta_{j-1,\mathcal{E}}^4 + C n \sum_{j=2}^{n} \sum_{\mathcal{E}} p_{\mathcal{E}} p_{\mathcal{E}'} (E \eta_{j-1,\mathcal{E}}^4 E \eta_{j-1,\mathcal{E}'}^4)^{1/2} \\
\leq C n \sum_{j=2}^{n} \sum_{\mathcal{E}} p_{\mathcal{E}} \left( 1 + (n p_{\mathcal{E}})^{2r-2} \right) + C n \sum_{j=2}^{n} \left( \sum_{\mathcal{E}} p_{\mathcal{E}} \left( 1 + (n p_{\mathcal{E}})^{r-1} \right)^2 \right)^{1/2} \\
\leq C n^2 \left( 1 + (n h)^{2r-2} \right) + C n^2 \left( 1 + (n h)^{-r} \right) \sum_{j=2}^{n} \sum_{\mathcal{E}} p_{\mathcal{E}} \eta_{j-1,\mathcal{E}}^{r+r+1} \\
\leq C n^2 (n h)^{2r-2} 
\]  

(23)

From (19)-(21) and (23), (18) follows.

Now to prove that for \( r > 1 \) and given \( \varepsilon > 0 \), we have

\[
P \left( \max_{1 \leq k \leq n} T_{kr} \geq \varepsilon \right) \leq C/n^2 
\]  

(24)

First suppose that \( 1 < r \leq 2 \). We have

\[
|S_{n,\mathcal{E}}|^r \leq |S_{n-1,\mathcal{E}}|^r + C |Y_{n,\mathcal{E}}|^r + r |S_{n-1,\mathcal{E}}|^r |\text{sign}(S_{n-1,\mathcal{E}}) Y_{n,\mathcal{E}}|^r 
\]

Hence

\[
|S_{n,\mathcal{E}}|^r \leq C \sum_{j=1}^{n} |Y_{j,\mathcal{E}}|^r + r \sum_{j=2}^{n} \sum_{j' \in \mathcal{E}} |Y_{j,\mathcal{E}}| Y_{j',\mathcal{E}}^{-1} 
\]

and

\[
n^{-r} h^{-r+1} \max_{1 \leq k \leq n} T_{kr} \\
\leq C n^{-r} h^{-r+1} \sum_{j=1}^{n} \sum_{\mathcal{E}} |Y_{j,\mathcal{E}}|^r \\
+ n^{-r} h^{-r+1} \max_{2 \leq k \leq n} \left[ \sum_{j=2}^{n} \sum_{\mathcal{E}} |Y_{j,\mathcal{E}}| Y_{j',\mathcal{E}}^{-1} \right]. 
\]  

(25)

Since

\[
n^{-r} h^{-r+1} \sum_{j=1}^{n} \sum_{\mathcal{E}} |Y_{j,\mathcal{E}}|^r \leq 2 (n h)^{-r+1} \to 0 
\]

From (18) and (25), (24) follows.

Suppose that (24) is true for \( r \leq m \), and proceed to show that it is true for \( r \in (m, m+1) \). Since \( r > 2 \), we have
\[
|s_{n-1טען} |^r \leq |s_{n-1,ט} |^r + r|s_{n-1,ט} |^{-2} s_{n-1,ט} \gamma_n \n亙 + C|s_{n-1,ט} |^{-2} \gamma_n ^2 + C|\gamma_n |^r
\]

Hence
\[
n^{-r}_{\text{max}} \sum_{k=1}^{n \text{ max}} T_{k,r} \leq Cn^{-r}_{\text{max}} \sum_{k=1}^{n \text{ max}} |\gamma_j |^r
\]
\[
+ r n^{-r}_{\text{max}} \sum_{k=1}^{n \text{ max}} |\gamma_j |^r
\]
\[
+ n^{-r}_{\text{max}} \sum_{k=1}^{n \text{ max}} |\gamma_j |^r
\]
\[
+ C(n \text{ h})^{-r \text{ max}} \sum_{k=1}^{n \text{ max}} p_k |s_{j-1,ט} |^{-2} \frac{4}{\sum_{i=1}^{H_i}}
\]

(26)

where \( \gamma_j \text{-ט} = |s_{j-1,ט} |^{-2} \). Observe that \( H_2 /r \) is the expression in (18). \( nhH_3 /C \) is also of the form in (18) with \( r \) replaced by \( r - 1 \). Further, since \( r > 2 \) and \( \int f(x)dx < \infty \), we have \( \int f^{-1}(x)dx < \infty \). Therefore, the inequality (18) can be applied to both \( H_2 \) and \( nhH_3 \). Further, \( nh \rightarrow \infty \), so we obtain
\[
P(H_2 + H_3 \geq \varepsilon) \leq C/n^2
\]

(27)

for arbitrarily given \( \varepsilon > 0 \). Also,
\[
n^{-r}_{\text{max}} \sum_{j=1}^{n \text{ max}} |\gamma_j |^r \leq 2(nh)^{-r \text{ max}} \rightarrow 0
\]

(28)

Since \( \int f^{-1}(x)dx < \infty \), by induction hypothesis, we have
\[
P(n^{-r \text{ max}} h^{-r \text{ max}} \sum_{k=1}^{n \text{ max}} T_{k,r-1} \geq \varepsilon) \leq C/n^2
\]

(29)

By Hölder inequality and the fact \( \sum_{k \text{-ט}} p_k ^{-1} \leq C h^{-2} \),
\[
(nh)^{-r \text{ max}} \sum_{k=1}^{n \text{ max}} p_k |s_{j-1,ט} |^{-2}
\]
\[
\leq C(nh)^{-1} (n^{-r \text{ max}} h^{-r \text{ max}} \sum_{k=1}^{n \text{ max}} T_{k,r-1}) ^{(r-2) / (r-1)}
\]

(30)

(29), (30) together give
\[ P(H_n \geq \varepsilon) \leq C/n^2 \]  

(31)

for arbitrarily given \( \varepsilon > 0 \). Summing up (26)-(28) and (31) we see that (24) is true for \( r \in [m, m + 1] \), concluding the proof of (24) hence (17). Thus we have proved (6) for the case of \( r > 1 \).

Now consider the case of \( r = 1 \). Since \( f(x) \) as a function of \( x \) in \( (-\infty, \infty) \) is a probability density, in order to prove (6), it suffices to show that for any fixed positive integer \( N \), it is true that

\[ \lim_{n \to \infty} \int |f_n(x) - f(x)| \, dx = 0, \quad a.s. \]  

(32)

where \( l = \Delta \), \( H = \lfloor h^{-1} \rfloor \), the integer part of \( h^{-1} \).

To prove (32), it suffices to verify the following two assertions:

\[ \lim_{n \to \infty} \int |f_n(x) - Ef_n(x)| \, dx = 0 \]  

(33)

\[ \lim_{n \to \infty} \int |f_n(x) - Ef_n(x)| \, dx = 0, \quad a.s. \]  

(34)

The second assertion follows directly from Lemma 3 of Devroye [3] if we note the fact that \( NH/n \leq N/h \to 0 \). The first assertion can be verified by using a continuous function \( g \) on \([a-N-1,a+N+1]\) such that the integral

\[ \int_{a-N-1}^{a+N+1} |f(x) - g(x)| \, dx \]

does not exceed given \( \varepsilon > 0 \). Trivial details are omitted.

(C) Necessity of (3), (4)

Since

\[ \int f_n^r(x) \, dx = n^{-r}h^{-r+1} \sum_{k} \sum_{i=1}^{n} (1_{k}(x_i) - p_k) \leq 2^r/h^{r-1} < \infty \]

Therefore, if (6) or (7) is true, then (3) is true.
Suppose that \( f(x) \) does not have a form
\[
\sum_{\ell = -\infty}^{\infty} C_{\ell} \chi_{\left[a_0 + \ell h_0, a_0 + (\ell + 1)h_0 \right)}(x)
\] (35)
for some \( a_0, h_0 > 0 \) and \( \{C_{\ell}\} \). We want to prove that if (6) or (7) is true, then \( h_n \to 0 \). Suppose in the contrary that \( h_n \nrightarrow 0 \). Then there exists subsequence \( \{h_{n_i}\} \) such that \( h_{n_i} \to h_0 > 0 \) as \( i \to \infty \). \( h_0 \) must be finite, otherwise we would have \( f_n(x) \to 0 \) uniformly in \( \{X_1, X_2, \ldots\} \) and \( X \), and this contradicts obviously with any one of (6) or (7). Since \( \{|a_i|\} \leq h_n \), without losing generality we may assume that \( a_i \to a_0 \) also finite. From these facts it follows by the law of large numbers that if we define
\[
g(x) = h_0 \int_{a_0 + \ell h_0}^{a_0 + (\ell + 1)h_0} f(t) \, dt,
\]
then we have
\[
\lim_{n \to \infty} \int_I \left| f_n(x) - g(x) \right| \, dx = 0, \quad \text{a. s.}
\] (36)
for any bounded interval \( I \). Since at least one of (6) and (7) is true, (36) implies that \( f = g \) almost everywhere on \( (-\infty, \infty) \) (Lebesgue measure). Hence \( f \) can be expressed in the form (35), contradicting the assumption.

(D). Necessity of (5).

Suppose in the contrary that (5) is not true, then there exists subsequence \( \{h_i\} \) such that \( h_i \to 0 \), \( n h_i \to u < \infty \). We restrict ourselves to the discussion of the subsequence. For simplicity of writing and without losing generality, we may assume that \( h_n \to 0 \) and
\[
\lim_{n \to \infty} n h_n = u < \infty
\] (37)
Since \( h \to 0 \), we have
\[
\lim_{n \to \infty} \int_I \left| f(x) - Ef_n(x) \right| \, dx = 0
\] (38)
In fact, as pointed out earlier, (33) is true for any bounded interval I. Since f(x) and \( Ef_n(x) \) are probability density functions on \((-\infty, \infty)\), this fact implies (38). Now it is easily seen that if at least one of (6) and (7) is true, then

\[
\lim_{n \to \infty} E \int_I |f_n(x) - Ef_n(x)| dx = 0
\]

(39)

In fact, if (7) is true, then for any bounded interval I we have by Hölder inequality that

\[
E \int_I |f_n(x) - f(x)| dx \to 0, \quad \text{as } n \to \infty
\]

(40)

Since f(x), f_n(x) are probability density functions, it is easily seen that

\[
\int_I |f_n(x) - f(x)| dx \leq 2 \int_I |f_n(x) - f(x)| dx + 2 \left[ 1 - \int_I f(x) dx \right]
\]

(41)

From (40), (41), it follows that

\[
\lim_{n \to \infty} E \int_I |f_n(x) - f(x)| dx = 0
\]

(42)

Now (39) follows from (38) and (42). If (6) is true, then by Hölder inequality we have

\[
\lim_{n \to \infty} \int_I |f(x) - f_n(x)| dx = 0, \quad \text{a.s.}
\]

(43)

for any bounded interval I. From (43) and the dominated convergence theorem, we again have (40) and hence (39).

First we assume that \( v = 0 \). Write \( A_n = \bigcup_{i=1}^n [X_i - h, X_i + h] \). By the definition, \( f_n(x) = 0 \) for \( x \in A \). So we have

\[
E \int_{A_n} f_n(x) dx = 1.
\]

Denote by \( \lambda(A_n) \) the Lebesgue measure of \( A_n \). Then

\[
\lambda(A_n) \leq 2nh \to 0, \quad \text{as } n \to \infty,
\]

which implies that

\[
\lim_{n \to \infty} E \int_{A_n} f(x) dx = 0.
\]

Thus

\[
E \int_{A_n} |f_n(x) - f(x)| dx \geq E \int_{A_n} |f_n(x) - f(x)| dx
\]

\[
E \int_{A_n} |f_n(x) - f(x)| dx \geq E \int_{A_n} |f_n(x) - f(x)| dx
\]
\[ \int_{A_n} f_n(x) \, dx - \int_{A_n} f(x) \, dx \to 1, \text{ as } n \to \infty. \]

From this and (38), it follows that

\[ \lim_{n \to \infty} \int f_n(x) - Ef_n(x) \, dx = 1, \tag{44} \]

which contradicts (39).

Now we assume \( 0 < u < \infty \). By (39), there exist a sequence of positive constants, say \( C_n \), such that \( C_n \to \infty \), \( C_n/n \to 0 \) and

\[ C_n \int |f_n(x) - Ef_n(x)| \, dx \to 0. \tag{45} \]

Write \( k = \lceil n/C \rceil \). Then

\[ C_n \int |f_n(x) - Ef_n(x)| \, dx \]
\[ = C_n (nh)^{-1} \sum_{k} \mathbb{E} \left| \sum_{i=1}^{n} \left( I_k(x_i) - p_k \right) \right| \]
\[ \geq (kh)^{-1} \sum_{k} \mathbb{E} \left| \sum_{i=1}^{k} \left( I_k(x_i) - p_k \right) \right| x_1, \ldots, x_k \right| \]
\[ = (kh)^{-1} \sum_{k} \mathbb{E} \left| \sum_{i=1}^{k} \left( I_k(x_i) - p_k \right) \right|. \tag{46} \]

Since \( h \to 0 \) and \( kh \leq nh/C \to 0 \), by (44) we have

\[ \lim_{k \to \infty} (kh)^{-1} \sum_{k} \mathbb{E} \left| \sum_{i=1}^{k} \left( I_k(x_i) - p_k \right) \right| = 1. \]

On the other hand, (45) and (46) together imply

\[ (kh)^{-1} \sum_{k} \mathbb{E} \left| \sum_{i=1}^{k} \left( I_k(x_i) - p_k \right) \right| \to 0, \text{ as } n \to \infty. \]

Thus a contradiction is reached, and the proof of the Theorem is completed.

Remark. The method of proof in (D) can easily be adopted to the case of the kernel estimate.
REFERENCES


END
DATE
FILMED
DEC.
1987