Nihonkai Math. J.
Vol.8 (1997), 195-207

MEAN INTEGRATED SQUARED ERROR AND DEFICIENCY OF NONPARAMETRIC RECURSIVE KERNEL ESTIMATORS OF SMOOTH DISTRIBUTION FUNCTIONS

Keiichi HIROSE*

Abstract

In this paper we derive asymptotic expressions of the mean integrated squared error (MISE) as a global measure for a nonparametric recursive kernel estimator of a distribution function. We also investigate whether the recursive kernel estimator has asymptotically better performance than the empirical distribution function. It is proved that the relative deficiency of the empirical distribution function with respect to an appropriately chosen recursive kernel estimator quickly tends to infinity as the sample size increases.

1. Introduction

Let $X_1, X_2, \ldots$ be independent and identically distributed random variables having a common cumulative distribution function (c.d.f) F. Throughout this paper, the c.d.f F is assumed to be absolutely continuous with respect to Lebesgue measure, and let a probability density function (p.d.f) of F be denoted by f.

In this paper, we consider the problem of estimating the c.d.f F. Traditionally, as an estimator of F, the empirical distribution function

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \le x) \tag{1.1}$$

has been chosen, where $I(\cdot)$ denotes the indicator function. It is well known that the estimator is strongly uniformly consistent. But this estimator does not take into account the smoothness of F, that is, the existence of the density f. On the other hand, taking into account this point, alternative estimators of the c.d.f F have been considered by (to mention a few) Nadaraya [7], Yamato [17], Winter [15, 16], Hill [3], Puri and Ralescu [8], Yukitch [18] and Sadra [11]. Especially, for kernel-type estimators, Singh et al. [13] investigated their asymptotic properties including strong uniform consistency and asymptotic normality, and also derived rates of convergence.

*Department of Mathematical Science, Graduate School of Science and Technology, Niigata University, Niigata, 950-21, Japan [hirose@dmis.gs.niigata-u.ac.jp]

Under a certain condition on the c.d.f F, Read [9] showed that the empirical distribution function is inadmissible with respect to the mean integrated squared error (MISE). Reiss [10] and Falk [2] showed that the kernel-type estimators of a distribution function at a single point have an asymptotically better performance on the level of deficiency than the empirical distribution function for appropriately chosen kernels and sufficiently smooth F. Swanepoel [14] derived asymptotic expressions of the MISE as a global measure of a kernel-type estimator of F. He also considered the problem of finding optimal kernels and asserted that the kernel-type estimator is asymptotically more efficient than the empirical distribution function. Shirahata and Chu [12] treated the integrated squared error (ISE) of kernel-type estimator and compared the error with that of the empirical distribution function.

The kernel-type estimators mentioned above are not recursive in nature, i.e. when the sample size increases, the estimators must be computed from the beginning. Besides, we are required to store extensive data in order to calculate them. With these things in mind, Isogai and Hirose [5] proposed a class of recursive kernel estimators $\left\{\hat{F}_n\right\}$ and studied their asymptotic properties, including strong uniform consistency. Rates of convergence of the mean squared error were also investigated and asymptotic normality of the estimators was shown. By 'recursive', we mean that the estimator $\hat{F}_n$ based on the first n observations is a function of $\hat{F}_{n-1}$ and the n-th observation. Of course, the empirical distribution function is recursive. Then, we should expect that the recursive kernel estimator $\hat{F}_n$ has better performance than the empirical distribution function F_n in (1.1) with respect to a reasonable measure of performance. We shall derive asymptotic expressions of the MISE for the recursive kernel estimator and prove that the relative deficiency of the empirical distribution function with respect to the recursive kernel estimator quickly tends to infinity as the sample size increases.

In Section 2 we shall define the class of recursive kernel estimators of the c.d.f F proposed by Isogai and Hirose [5]. Section 3 gives the asymptotic expansions of the MISE of the recursive kernel estimators and the relative efficiency and deficiency of the empirical distribution function with respect to them. Proofs of the results in Section 3 will be postponed until the final section.

2. Recursive kernel estimators

In this section, we shall define the class of recursive kernel estimators proposed by Isogai and Hirose [5].

Let k (called a kernel) be a real-valued Borel measurable function on the real line R such that

$$\int |k(t)|dt < \infty \quad \text{and} \quad \int k(t)dt = 1,$$

where all integrals are taken over R, unless otherwise specified. Let $\{h_n\}$ (called bandwidths) be a sequence of positive numbers converging to zero. For each $n \geq 1$, set

$$k_n(t) = \frac{1}{h_n} k\left(\frac{t}{h_n}\right), \quad K_n(x) = \int_{-\infty}^{x} k_n(t)dt, \quad K(x) = \int_{-\infty}^{x} k(t)dt$$

and

$$\hat{f}_n(x) = \frac{1}{n} \sum_{j=1}^{n} k_j(x - X_j). \tag{2.1}$$

Now we define the recursive kernel estimators $\{\hat{F}_n\}$ of the c.d.f F as

$$\hat{F}_n(x) = \int_{-\infty}^{x} \hat{f}_n(y)dy = \frac{1}{n} \sum_{j=1}^{n} K_j(x - X_j) \quad \text{for } n \geq 1. \tag{2.2}$$

These estimators can be recursively computed as follows:

$$\hat{F}_n(x) = \left(1 - \frac{1}{n}\right) \hat{F}_{n-1}(x) + \frac{1}{n} K_n(x - X_n) \quad \text{for } n \geq 1,$$

where $\hat{F}_0(x) = K(x)$.

3. Mean integrated squared error and deficiency

In this section, under the assumption that for some integer $m \geq 1$ the $(m+1)th$ derivative $F^{(m+1)} = f^{(m)}$ of F exists, we shall give the asymptotic expansions of the MISE of the estimator $\hat{F}_n$ of F in (2.2). Furthermore, we shall show that the estimators have an asymptotically better performance on the level of deficiency than the empirical distribution function for appropriately chosen kernels and bandwidths and sufficiently smooth F.

As a global measure we use the mean integrated squared error given by

$$MISE\left(\hat{F}_n\right) = E\left\{\int \left(\hat{F}_n(x) - F(x)\right)^2 f(x)dx\right\}.$$

For some $m \geq 1$, let $\mathcal{K}_m$ be the class of all kernels k such that

$$\int k(t)dt = 1, \quad \int t^i k(t)dt = 0, \; i = 1, \ldots, m$$

and

$$M_{m+1} \equiv \frac{1}{(m+1)!} \int \left|t^{m+1} k(t)\right| dt < \infty.$$

We state here assumptions on f and $\{h_n\}$ imposed in this section.

ASSUMPTION A:

(A1) $f^{(m)}$ is bounded and continuous almost everywhere with respect to Lebesgue measure (a.e. continuous).

(A2) f^2 is integrable and f' is bounded.

(A3) In addition to (A2), f' is a.e. continuous.

ASSUMPTION H:

(H1) $$\sum_{j=1}^{n} h_j^{m+1} \to \infty \quad \text{as } n \to \infty.$$

(H2) $$\sum_{j=1}^{n} h_j \to \infty \quad \text{as } n \to \infty.$$

(H3) $$n\left(\frac{1}{n}\sum_{j=1}^{n} h_j^2\right)^2 \to 0 \quad \text{as } n \to \infty.$$

(H4) $$\sum_{j=1}^{n} h_j^2 \to \infty \quad \text{as } n \to \infty.$$

The following lemmas provide the exact asymptotic behavior for the integrated squared value (ISV) of the bias and the variance of $\hat{F}_n$.

LEMMA 3.1. For some integer $m \geq 1$, assume (A1), (H1) and $k \in \mathcal{K}_m$. Then, as $n \to \infty$,

$$\begin{aligned}&\int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx \\ &= \left(\frac{1}{n}\sum_{j=1}^{n} h_j^{m+1}\right)^2 \mu_{m+1}^2 \int \left(f^{(m)}(x)\right)^2 f(x)dx + o\left(\left(\frac{1}{n}\sum_{j=1}^{n} h_j^{m+1}\right)^2\right),\end{aligned}$$

where $\mu_{m+1} = \dfrac{1}{(m+1)!}\displaystyle\int t^{m+1}k(t)dt.$

LEMMA 3.2. Assume (A2), (H2) and $k \in \mathcal{K}_1$. Then, as $n \to \infty$,

$$\begin{aligned}&\int Var\left(\hat{F}_n(x)\right) f(x)dx \\ &= \frac{1}{6n} - \frac{1}{n^2}\sum_{j=1}^{n} h_j \int tb(t)dt \int f^2(x)dx + o\left(\frac{1}{n^2}\sum_{j=1}^{n} h_j\right)\end{aligned}$$

where $b(t) \equiv 2k(t)K(t)$.

Since (H1) implies (H2) and $\mathcal{K}_1$ includes $\mathcal{K}_m$, from Lemmas 3.1 and 3.2 we get the asymptotic expansion of the MISE of $\hat{F}_n$.

THEOREM 3.1. For some integer $m \geq 1$, assume (A1), (A2), (H1) and $k \in \mathcal{K}_m$. Then, as $n \to \infty$,

$$\begin{aligned} MISE\left(\hat{F}_n\right) &= \frac{1}{6n} + \left(\frac{1}{n}\sum_{j=1}^{n} h_j^{m+1}\right)^2 \mu_{m+1}^2 \int \left(f^{(m)}(x)\right)^2 f(x)dx \\ &\quad -\frac{1}{n^2}\sum_{j=1}^{n} h_j \int tb(t)dt \int f^2(x)dx + o\left(\left(\frac{1}{n}\sum_{j=1}^{n} h_j^{m+1}\right)^2 + \left(\frac{1}{n^2}\sum_{j=1}^{n} h_j\right)\right). \end{aligned}$$

Remark 3.1. We note that for $m \geq 2$ any fixed k in $\mathcal{K}_m$ can be negative, thus leading to the negative estimator $\hat{F}_n$ for the c.d.f F. But, Lemma 3.1 shows the reduction of the ISV of the bias of $\hat{F}_n$. Hence, one can improve the rates of convergence of the ISV of the bias to zero by using the kernel k in $\mathcal{K}_m$. On the other hand, in case of $m = 1$ we can choose a kernel k in $\mathcal{K}_1$ as a p.d.f. Then $\hat{f}_n(x)$ in (2.1) and $\hat{F}_n(x)$ in (2.2) are a p.d.f and a c.d.f, respectively.

We shall introduce a criterion under which we compare $\hat{F}_n$ and F_n given by

$$j(n) \equiv \min\left\{k \geq 1 : MISE(F_k) \leq MISE\left(\hat{F}_n\right)\right\},$$

that is, $j(n)$ denotes the sample size which is needed so that the empirical distribution function has the same or a smaller MISE as $\hat{F}_n$. Now, a comparison of the two estimators $\hat{F}_n$ and F_n is made by comparing $j(n)$ with n. This can be carried out by considering the ratio $j(n)/n$ or the difference $j(n) - n$. The ratio $j(n)/n$ is usually called relative or first order efficiency while the difference $j(n) - n$ is known as second order efficiency of deficiency. The limit values $\lim_{n\to\infty} j(n)/n$ and $\lim_{n\to\infty}(j(n) - n)$ are called, if they exist, asymptotic relative efficiency and asymptotic deficiency, respectively. The concept of deficiency was introduced by Hodges and Lehmann [4]. Akahira [1] discussed the asymptotic deficiencies of various estimators.

We shall give a behavior of the ratio of the MISE of the two estimators $\hat{F}_n$ and F_n.

LEMMA 3.3. Assume (A2), (H2) and $k \in \mathcal{K}_1$. Then, as $n \to \infty$,

$$\begin{aligned} &\frac{MISE\left(\hat{F}_n\right)}{MISE(F_n)} \\ &= 1 - \frac{6}{n}\sum_{j=1}^{n} h_j \int tb(t)dt \int f^2(x)dx + 6n\int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx + o\left(\frac{1}{n}\sum_{j=1}^{n} h_j\right). \end{aligned}$$

The following proposition states that the recursive kernel estimator of the c.d.f F with fixed kernel does not have an asymptotically better performance on the level of efficiency than the empirical distribution function.

PROPOSITION 3.1. Assume (A3), (H2) and $k \in \mathcal{K}_1$. Then

$$\text{(i)} \qquad \limsup_{n\to\infty} \frac{j(n)}{n} \le 1$$

and

$$\text{(ii)} \qquad \lim_{n\to\infty} \frac{j(n)}{n} = 1 \quad \text{if and only if} \quad \lim_{n\to\infty} n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx = 0.$$

From Lemma 3.1, we can always ensure that under (A1) with $m = 1$, (H3) and (H4)

$$\lim_{n\to\infty} n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx = 0.$$

Proposition 3.1 implies $\lim_{n\to\infty} j(n)/n = 1$ for any $k \in \mathcal{K}_1$. Hence, in this case, we cannot distinguish between the asymptotic performance of the estimators with different kernels when compared to the empirical distribution function. Therefore, we have to consider the level of relative deficiency.

Example 3.1. Let

$$h_n = cn^{-\alpha} \quad \left(\frac{1}{2} \ge \alpha > \frac{1}{4}\right)$$

where $c > 0$ is a constant. Then (H3) and (H4) are satisfied.

We shall give the asymptotic expression of the relative deficiency of the empirical distribution function with respect to the recursive kernel estimator.

THEOREM 3.2. Assume (A3), (H3), (H4) and $k \in \mathcal{K}_1$. Then,

$$\text{(i)} \qquad \frac{j(n) - n}{\sum_{j=1}^n h_j} = \theta - \frac{1}{n^{-1}\sum_{j=1}^n h_j} \cdot \frac{6n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx}{1 + 6n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx} + o(1)$$

where $\theta \equiv 6 \int tb(t)dt \int f^2(x)dx$

and

$$\text{(ii)} \qquad \lim_{n\to\infty} \frac{j(n) - n}{\sum_{j=1}^n h_j} = \theta$$

if and only if $n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx = o\left(\frac{1}{n}\sum_{j=1}^n h_j\right)$.

Example 3.2. Let

$$h_n = cn^{-\alpha} \quad \left(\frac{1}{2} \geq \alpha > \frac{1}{3}\right)$$

where $c > 0$ is a constant, then $n \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx = o\left(\frac{1}{n}\sum_{j=1}^{n} h_j\right)$.

From Theorem 3.2, if $\int tb(t)dt$ is negative, then the empirical distribution function is better than the recursive kernel estimator and if it is positive, then the recursive kernel estimator with $\{h_n\}$ in Example 3.2 is prefered to the empirical distribution function. Since we can choose a p.d.f. $k \in \mathcal{K}_1$ satisfying $\int tb(t)dt > 0$, the recursive kernel estimator has asymptotically better performance than the empirical distribution function.

4. The proofs

In this section all the proofs of the results in Section 3 are given.

PROOF OF LEMMA 3.1. From (2.1) and $\int k(t)dt = 1$ we have

$$\lim_{x \to -\infty} K(x) = 0, \quad \lim_{x \to \infty} K(x) = 1 \quad \text{and} \quad \int b(t)dt = 1, \tag{4.1}$$

where $b(\cdot)$ is the function given in Lemma 3.2. Thus, using integration by parts we get

$$E\left[\hat{F}_n(x)\right] = \frac{1}{n}\sum_{j=1}^{n} \int \frac{1}{h_j} k\left(\frac{x-y}{h_j}\right) F(y)dy. \tag{4.2}$$

Set

$$A_n = \sum_{j=1}^{n} h_j^{m+1} \quad \text{and} \quad a_j(x) = \frac{(-1)^{m+1}}{m!} \int t^{m+1} k(t) \left(\int_0^1 (1-y)^m f^{(m)}(x - th_j y)dy\right) dt.$$

It follows from (4.2), the Taylor expansion and $k \in \mathcal{K}_m$, we have

$$\begin{aligned} & \left(\frac{n}{A_n}\right)^2 \int \left(E\left[\hat{F}_n(x)\right] - F(x)\right)^2 f(x)dx \\ = & \left(\frac{n}{A_n}\right)^2 \int \left\{\frac{1}{n}\sum_{j=1}^{n} \int k(t)(F(x - th_j) - F(x))dt\right\}^2 f(x)dx \\ = & \int \left(\frac{1}{A_n}\sum_{j=1}^{n} h_j^{m+1} a_j(x)\right)^2 f(x)dx. \end{aligned} \tag{4.3}$$

Since by (A1)

$$\lim_{j\to\infty}\int_0^1 (1-y)^m f^{(m)}(x-th_jy)dy = \frac{1}{m+1}f^{(m)}(x) \quad \text{for a.e. } x,$$

using $t^{m+1}k(t) \in L_1$ and the dominated convergence theorem, we have

$$\lim_{j\to\infty} a_j(x) = (-1)^{m+1}\mu_{m+1}f^{(m)}(x) \quad \text{a.e. } x,$$

which, together with (H1) and Toeplitz's lemma (see Loève [6, p.250]), yields

$$\lim_{n\to\infty}\frac{1}{A_n}\sum_{j=1}^{n} h_j^{m+1}a_j(x) = (-1)^{m+1}\mu_{m+1}f^{(m)}(x) \quad \text{for a.e. } x.$$

Hence, from (A1), (4.3) and the dominated convergence theorem, we get the result of the lemma. Therefore the proof of Lemma 3.1 is complete.

Before proving the result of Lemma 3.2 we give two lemmas.

LEMMA 4.1. For some integer $m \geq 1$, let $f^{(m)}$ be bounded. Assume $k \in \mathcal{K}_m$ and (H2). Then, as $n \to \infty$,

$$\text{(i)} \quad \frac{1}{n}\sum_{j=1}^{n}\int\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}F(x)f(x)dx = o\left(\frac{1}{n}\sum_{j=1}^{n}h_j\right)$$

and

$$\text{(ii)} \quad \frac{1}{n}\sum_{j=1}^{n}\int\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}^2 f(x)dx = o\left(\frac{1}{n}\sum_{j=1}^{n}h_j\right).$$

PROOF. First we shall show (i). Set $A_n \equiv \sum_{j=1}^{n} h_j$ and

$$a_j \equiv h_j^m\frac{(-1)^{m+1}}{m!}\int\left\{\int t^{m+1}k(t)\left(\int_0^1(1-y)^m f^{(m)}(x-th_jy)dy\right)dt\right\}F(x)f(x)dx.$$

Then by the Taylor expansion and $k \in \mathcal{K}_m$, we have

$$\begin{aligned}&\frac{n}{A_n}\left[\frac{1}{n}\sum_{j=1}^{n}\int\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}F(x)f(x)dx\right]\\ =\ &\frac{1}{A_n}\sum_{j=1}^{n}h_ja_j. \qquad (4.4)\end{aligned}$$

By virtue of the fact that $f^{(m)}$ be bounded and $t^{m+1}k(t) \in L_1$, we get $a_j = O\left(h_j^m\right)$, which, together with (H2) and Toepliz's lemma, implies

$$\lim_{n\to\infty} \frac{1}{A_n} \sum_{j=1}^{n} h_j a_j = 0.$$

Thus, from (4.4) we obtain (i). By the same argument as (i) we can show (ii). This completes the proof.

LEMMA 4.2. Assume (A2), (H2) and $k \in \mathcal{K}_1$. Then, as $n \to \infty$,

$$\begin{aligned} &\frac{1}{n}\sum_{j=1}^{n} \int \left\{ \int b(t)(F(x-th_j) - F(x))dt \right\} f(x)dx \\ = \ & -\frac{1}{n}\sum_{j=1}^{n} h_j \int tb(t)dt \int f^2(x)dx + o\left(\frac{1}{n}\sum_{j=1}^{n} h_j\right). \end{aligned}$$

PROOF. By the Taylor expansion, we have

$$\begin{aligned} &\int \left\{ \int b(t)(F(x-th_j) - F(x))dt \right\} f(x)dx \\ = \ & -h_j \int tb(t)dt \int f^2(x)dx \\ &+h_j^2 \int \left\{ \int t^2 b(t) \left(\int_0^1 (1-y) f'(x - th_j y)dy \right) dt \right\} f(x)dx. \end{aligned} \tag{4.5}$$

It follows from the assumption that the second term of the right-hand side of (4.5) is $O\left(h_j^2\right)$, which, together with (H2), Toeplitz's lemma and (4.5), yields

$$\begin{aligned} &\frac{n}{\sum_{j=1}^{n} h_j} \cdot \frac{1}{n}\sum_{j=1}^{n} \int \left\{ \int b(t)(F(x-th_j) - F(x))dt \right\} f(x)dx \\ \to \ & -\int tb(t)dt \int f^2(x)dx \quad \text{as } n \to \infty. \end{aligned}$$

This completes the proof.

We are now in a position to prove the rest of the results in Section 3.

PROOF OF LEMMA 3.2. From (4.1) and integration by parts, we get

$$\begin{aligned} &Var\left(\hat{F}_n(x)\right) \\ = \ & \frac{1}{n^2}\sum_{j=1}^{n} \int \frac{1}{h_j} b\left(\frac{x-y}{h_j}\right) F(y)dy - \frac{1}{n^2}\sum_{j=1}^{n} \left(\int \frac{1}{h_j} k\left(\frac{x-y}{h_j}\right) F(y)dy \right)^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{j=1}^{n}\int b(t)F(x-th_j)dt - \frac{1}{n^2}\sum_{j=1}^{n}\left(\int b(t)dt\right)F(x) \\
&\quad -\frac{1}{n^2}\sum_{j=1}^{n}\left(\int k(t)F(x-th_j)dt\right)^2 + \frac{F(x)}{n} \\
&= \frac{F(x)(1-F(x))}{n} - \frac{2}{n^2}\sum_{j=1}^{n}\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}F(x) \\
&\quad -\frac{1}{n^2}\sum_{j=1}^{n}\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}^2 \\
&\quad +\frac{1}{n^2}\sum_{j=1}^{n}\int b(t)(F(x-th_j)-F(x))dt,
\end{aligned}
$$

which gives

$$
\begin{aligned}
&\int Var\left(\hat{F}_n(x)\right)f(x)dx \\
&= \frac{1}{6n} - \frac{2}{n^2}\sum_{j=1}^{n}\int\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}F(x)f(x)dx \\
&\quad -\frac{1}{n^2}\sum_{j=1}^{n}\int\left\{\int k(t)(F(x-th_j)-F(x))dt\right\}^2 f(x)dx \\
&\quad +\frac{1}{n^2}\sum_{j=1}^{n}\int\left\{\int b(t)(F(x-th_j)-F(x))dt\right\}f(x)dx.
\end{aligned}
$$

Thus Lemmas 4.1 and 4.2 conclude the result of the lemma. This completes the proof.

PROOF OF LEMMA 3.3. Since $MISE(F_n) = 1/6n$, from Lemma 3.2 we have

$$
\begin{aligned}
\frac{MISE\left(\hat{F}_n\right)}{MISE(F_n)} &= 6n\int Var\left(\hat{F}_n(x)\right)f(x)dx + 6n\int\left(E\left[\hat{F}_n(x)\right]-F(x)\right)^2 f(x)dx \\
&= 1 - \frac{6}{n}\sum_{j=1}^{n}h_j\int tb(t)dt\int f^2(x)dx \\
&\quad +6n\int\left(E\left[\hat{F}_n(x)\right]-F(x)\right)^2 f(x)dx + o\left(\frac{1}{n}\sum_{j=1}^{n}h_j\right).
\end{aligned}
$$

Therefore the proof of Lemma 3.3 is complete.

PROOF OF PROPOSITION 3.1. First, we shall show (i). Since $MISE(F_n) = 1/6n$, it follows from the definition of $j(n)$ that

$$
\frac{j(n)-1}{n} < \frac{MISE(F_n)}{MISE\left(\hat{F}_n\right)} \le \frac{j(n)}{n}. \tag{4.6}
$$

Thus, from Lemma 3.3 we have

$$\begin{aligned}\limsup_{n\to\infty}\frac{j(n)-1}{n} &\le \limsup_{n\to\infty}\left\{1-\frac{6}{n}\sum_{j=1}^{n}h_j\int tb(t)dt\int f^2(x)dx+o\left(\frac{1}{n}\sum_{j=1}^{n}h_j\right)\right\}^{-1}\\ &= 1.\end{aligned}$$

This concludes (i). We shall next show (ii). If $\lim_{n\to\infty} n\int\left(E\left[\hat{F}_n(x)\right]-F(x)\right)^2 f(x)dx=0$ then from Lemma 3.3, we have

$$\lim_{n\to\infty} MISE\left(\hat{F}_n\right)/MISE(F_n)=1,$$

which, together with (4.6), yields $\liminf_{n\to\infty}\frac{j(n)}{n}\ge 1$. Thus, from (i) we get $\lim_{n\to\infty}\frac{j(n)}{n}=1$. Conversely, if $\lim_{n\to\infty}\frac{j(n)}{n}=1$, then from (4.6) we have

$$\lim_{n\to\infty} MISE(F_n)/MISE\left(\hat{F}_n\right)=1.$$

Hence, it follows from Lemma 3.3 that

$$\begin{aligned}&\lim_{n\to\infty}\left\{6n\int\left(E\left[\hat{F}_n(x)\right]-F(x)\right)^2 f(x)dx\right\}\\ =&\lim_{n\to\infty}\left\{\frac{MISE\left(\hat{F}_n\right)}{MISE(F_n)}-1+\frac{6}{n}\sum_{j=1}^{n}h_j\int tb(t)dt\int f^2(x)dx+o\left(\frac{1}{n}\sum_{j=1}^{n}h_j\right)\right\}=0.\end{aligned}$$

Thus, we conclude (ii). Therefore the proof of Proposition 3.1 is complete.

PROOF OF THEOREM 3.2. Set

$$B_n\equiv 6n\int\left(E\left[\hat{F}_n(x)\right]-F(x)\right)^2 f(x)dx \quad\text{and}\quad H_n\equiv\frac{1}{n}\sum_{j=1}^{n}h_j,$$

and let

$$\theta_n\equiv\frac{n\left\{\theta H_n-B_n+o\left(H_n\right)\right\}}{1+B_n-\theta H_n+o\left(H_n\right)}.$$

Then, from (4.6) and Lemma 3.3 we have

$$\frac{\theta_n}{nH_n}\le\frac{j(n)-n}{nH_n}<\frac{\theta_n+1}{nH_n}. \tag{4.7}$$

Furthermore, by (H3) and Lemma 3.1 we get

$$\begin{aligned}&\left|\frac{\theta_n}{nH_n}-\theta+\frac{B_n}{(1+B_n)H_n}\right|\\ =&\left|\frac{-2\theta B_n-\theta B_n^2+\theta^2 H_n+\theta^2 B_nH_n+o(1)}{\left\{1+B_n-\theta H_n+o(H_n)\right\}(1+B_n)}\right|\\ \to&\ 0\quad\text{as } n\to\infty,\end{aligned}$$

which, together with (4.7) and (H4), gives (i). (ii) is an immediate conclusion of (i). Therefore the proof of Theorem 3.2 is complete.

Acknowledgment.

I am grateful to Professor Eiichi Isogai for his valuable comments and I would also like to thank the referee for a careful reading of my manuscript.

References

[1] Akahira, M. : The Structure of Asymptotic Deficiency of Estimators, *Queen's Papers in Pure and Applied Mathematics*, **No.75**, Queen's University, 1986.

[2] Falk, M. : Relative efficiency and deficiency of kernel type estimators of smooth distribution functions, *Statistica Neerlandica*, **37** (1983), 73–83.

[3] Hill, P. D. : Kernel estimation of a distribution function, *Comm. Statist.-Theory Meth.*, **14** (1985), 605–620.

[4] Hodges, J. L. and Lehmann E. L. : Deficiency, *Ann. Math. Statist.*, **41** (1970), 783–801.

[5] Isogai, E. and Hirose, K. : Nonparametric recursive kernel estimators of distribution functions, *Bull. Inform. Cybernetics*, **26** (1994), 87–99.

[6] Loève, M. : *Probability Theory, Vol. I*, 4th ed., Springer, New york, 1977.

[7] Nadaraya, E. A. : Some new estimates for distribution functions, *Theory Probab. Appl.*, **9** (1964), 497–500.

[8] Puri, M. L. and Ralescu, S. S. : Central limit theorem for perturbed empirical distribution functions evaluated at a random point, *J. Multivariate Anal.*, **19** (1986), 273–279.

[9] Read, R. D. : The asymptotic inadmissibility of the sample distribution function, *Ann. Math. Statist.*, **43** (1972), 98–95.

[10] Reiss, R. D. : Nonparametric estimation of smooth distribution functions, *Scand. J. Statist.*, **8** (1981), 116–119.

[11] Sadra, P. : Smoothing parameter selection for smooth distribution function, *J. Statist. Plann. Inference*, **35** (1993), 65–75.

[12] Shirahata, S. and Chu, I. S. : Integrated squared error of kernel-type estimator of distribution function, *Ann. Inst. Statist. Math.*, **44** (1992), 579–591.

[13] Singh, R. S., Gasser, T. and Prasad, B. : Nonparametric estimates of distribution functions, *Comm. Statist.-Theory Meth.*, **12** (1983), 2095–2108.

[14] Swanepoel, J. W. H. : Mean integrated squared error properties and optimal kernels when estimating a distribution function, *Comm. Statist.-Theory Meth.*, **17** (1988), 3785–3799.

[15] Winter, B. B. : Strong uniform consistency of integrals of density estimators, *Canad. J. Statist.*, **1** (1973), 247–253.

[16] Winter, B. B. : Convergence rates of perturbed empirical distribution functions, *J. Appl. Probab.*, **16** (1979), 163–173.

[17] Yamato, H. : Uniform convergence of an estimator of a distribution function, *Bull. Math. Statist.*, **15** (1973), 69–78.

[18] Yukitch, J. E. : A note on limit theorems for perturbed empirical processes, *Stochastic Process. Appl.*, **33** (1989), 163–173.

Received June 2, 1997