Journal of the Korean Statistical Society (2001), **30**: *1, pp 77–87*

On Copas' Local Likelihood Density Estimator

W. C. Kim, B. U. Park[1] and Y. G. Kim[2]

ABSTRACT

Some asymptotic results on the local likelihood density estimator of Copas (1995) are derived when the locally parametric model has several parameters. It turns out that it has the same asymptotic mean squared error as that of Hjort and Jones (1996).

Keywords: local likelihood ; kernel weight ; density estimator ; asymptotic bias ; asymptotic variance

1. INTRODUCTION

Copas (1995), Hjort and Jones (1996) and Loader (1996) have developed local likelihood procedures for density estimation. The idea is to fit a parametric model $f(\cdot;\theta)$ locally at the point x of interest by

$$\widehat{f}(x) = f(x, \widehat{\theta})$$

where $\widehat{\theta} = \widehat{\theta}(x)$ is a maximizer of the local likelihood. For a suitably chosen local likelihood, the resulting estimator can enjoy the efficiency advantages of parametric inference as well as the adaptivity of nonparametric models. For example, consider the normal parametric model $f(t,\theta) = (2\pi\,\theta_2)^{-1/2} \times \exp\{-(t-\theta_1)^2/2\,\theta_2\}$. For a given point of interest x, the local likelihood procedure assumes that this model is true only in a neighborhood of the point x. With a wide neighborhood it is close to the parametric approach based on the normal model while with a narrow one it is nearly nonparametric.

[1]Department of Statistics, Seoul National University, Seoul, 151-742, KOREA

[2]Department of Mathematics, Humboldt State University, Arcata, CA95521, U.S.A.

Let $X_1, \ldots, X_n$ be a univariate random sample from the distribution with unknown density $f(\cdot)$. Hjort and Jones (1996) and Loader (1996) have proposed the local likelihood

$$\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)\, \log f(X_i;\theta) - \int K_h(t-x) f(t;\theta)\, dt \tag{1.1}$$

where $K_h(u) = K(u/h)/h$, K is a kernel function and the scale parameter h is the bandwidth controlling the amount of smoothing. With the normal parametric model and the standard normal kernel function $K(u) = (2\pi)^{-1/2}\exp(-u^2/2)$ it equals, being multiplied by $\sqrt{2\pi}\, h$,

$$\frac{1}{n}\sum_{i=1}^{n} \exp\left\{-\frac{(X_i - x)^2}{2h^2}\right\}\left\{-\frac{(X_i-\theta_1)^2}{2\theta_2} - \frac{1}{2}\log(2\pi\theta_2)\right\}$$
$$-h\,(\theta_2 + h^2)^{-1/2}\exp\left\{-\frac{(x-\theta_1)^2}{2(\theta_2+h^2)}\right\}.$$

When $h \to \infty$, the local likelihood is identical, in the limit, to the parametric global likelihood. On the other hand, when h is small, it produces an estimator which is very close to the fully nonparametric density estimator $\tilde{f}(x) = n^{-1}\sum_{i=1}^{n} K_h(X_i - x)$. This is because when $h \simeq 0$ only those X_i's which are close to x can have significant contribution to the likelihood so that one can replace, in the limiting situation, both of $(X_i - \theta_1)^2/2\theta_2$ and $(x-\theta_1)^2/2(\theta_2+h^2)$ by $(x-\theta_1)^2/2\theta_2$.

The local likelihood (1.1) was motivated by a locally weighted version of the Kullback-Leibler distance from f to $f(\cdot;\theta)$,

$$\int f(t)\log\left\{\frac{f(t)}{f(t;\theta)}\right\} dt = \int \left[f(t)\log\left\{\frac{f(t)}{f(t;\theta)}\right\} - \{f(t) - f(t;\theta)\}\right] dt.$$

Loader (1996) concentrates mainly on approximating $\log f(x;\theta)$ by polynomials, and Hjort and Jones (1996) considers more general local models. Hjort and Jones (1996) have shown that the rate of convergence of the resulting estimator depends on the number of parameters and not on the particular form of the locally parametric model.

On the other hand, Copas (1995) worked with the local likelihood

$$L_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} w_h(X_i - x)\,\log f(X_i;\theta) + \left\{1 - \frac{1}{n}\sum_{i=1}^{n} w_h(X_i - x)\right\}\log B_h(\theta) \tag{1.2}$$

where $w_h(u) = K_h(u)/K_h(0)$ and $B_h(\theta) = 1 - \int w_h(t-x)f(t;\theta)\,dt$. With this local likelihood, Copas obtained the asymptotic mean squared error of the resulting estimator when θ is a 1-dimensional parameter. Copas' local likelihood (1.2) was motivated by artificial censoring models in which X_i is only observed with probability $w(X_i)$, otherwise censored. It is interesting to note that, under such censoring mechanism, the Kullback-Leibler distance between models with densities f and $f(\cdot;\theta)$ is given by

$$\int w_h(t)f(t)\log\left\{\frac{w(t)f(t)}{w(t)f(t;\theta)}\right\}dt + \{1 - \int w(t)f(t)\,dt\}\log\left\{\frac{1-\int w(t)f(t)\,dt}{1-\int w(t)f(t;\theta)\,dt}\right\},$$

an empirical version of which motivates the local likelihood (1.2).

Recently, Eguchi and Copas (1998) produced an interesting study of these two local likelihood methods for large h. In fact, they set the two methods in a slightly wider context by giving a more general form of local likelihood density estimation with the two proposals as special cases. The remaining question is what happens for small h. In particular, it is intriguing to know whether the two resulting local likelihood density estimators based on (1.1) and (1.2) have the same small h asymptotic properties. The purpose of this paper is to answer this question when the locally parametric model has several parameters, and the result shows that the answer is affirmative.

2. ASYMPTOTIC PROPERTIES

In this section, asymptotic properties for Copas' (1995) local likelihood density estimator will be derived when the locally parametric model $f(\cdot;\theta) =$

$f(\cdot;\theta_1,\cdots,\theta_p)$ has several parameters. The results will be obtained under the following assumptions :

(A.1) K is a symmetric probability density with compact support ;

(A.2) $f(t;\theta)$ and $f(t)$ have bounded continuous (partial) derivatives up to order 6 ;

(A.3) The components $u_{1\theta}(t),\cdots,u_{p\theta}(t)$ of $u_\theta(t) = \partial \log f(t;\theta)/\partial\theta$ are functionally independent for any θ.

Copas' local likelihood density estimator is given by

$$\widehat{f}_n(x) = f(x;\widehat{\theta}_n) \tag{2.1}$$

where $\widehat{\theta}_n = \widehat{\theta}_n(x)$ is a maximizer of the local likelihood $L_n(\theta)$ in (1.2) or equivalently a solution to the equation

$$\frac{\partial}{\partial\theta}L_n(\theta) = 0. \tag{2.2}$$

At this point it should be mentioned that, as in $L_n(\theta)$, the dependence on x, the point of interest, will be suppressed whenever there is no confusion.

Remark 1. Since $\widehat{\theta}_n = \widehat{\theta}_n(x)$ depends on x, the estimator defined at (2.1) may not integrate to 1. To obtain a *bona fide* density we may divide it by its integral over the whole real line. It is easy to see that the asymptotic variance of the scaled density estimator is the same as the unscaled one. The asymptotic biases are different in constant factor but are same in order of convergence.

The solution $\widehat{\theta}_n$ to the equation (2.2) is expected to get closer to a solution $\theta_h = \theta_h(x)$ to the equation

$$\frac{\partial}{\partial\theta}\mathrm{E}L_n(\theta) = 0, \tag{2.3}$$

as n grows. In fact, arguments analogous to the maximum likelihood estimation can be applied to get an expansion of $\widehat{\theta}_n$ under the following assumptions :

(A.4) Expectations and differentiations of $L_n(\theta)$ with respect to θ can be interchanged ;

(A.5) The solutions $\widehat{\theta}_n$ and θ_h to the equations (2.2) and (2.3), respectively, exist uniquely, and the third order derivatives of $L_n(\theta)$ are bounded in a neighborhood of θ_h.

These assumptions are stronger than necessary, and some explicit models satisfying these are given in Hjort and Jones (1996). Under these assumptions, it is not difficult to observe that

$$\widehat{\theta}_n - \theta_h = -\{\mathrm{E}\,\ddot{L}_n(\theta_h)\}^{-1}\dot{L}_n(\theta_h) + O_p((nh)^{-1}) \tag{2.4}$$

as $n \to \infty$, $h \to 0$ and $nh \to \infty$ where $\cdot$ denotes the differentiation with respect to θ. Thus the next result follows from (2.1) and (2.4) :

Theorem 1. *Let x be an interior point of the support of f. Then under the assumptions (A.1)~(A.5), we have*

$$\widehat{f}_n(x) = f(x;\theta_h) + (nh)^{-1/2}\sigma(x)Z_n + O((nh)^{-1})$$

as $n \to \infty$, $h \to 0$ and $nh \to \infty$ where Z_n asymptotically obeys a standard normal distribution and

$$\sigma^2(x) = \begin{cases} f(x)\int K^2(z)\,dz & \text{if } p = 1 \text{ or } 2, \\ f(x)\int(\mu_2 z^2 - \mu_4)^2 K^2(z)\,dz/(\mu_4 - \mu_2^2)^2 & \text{if } p = 3 \text{ or } 4, \end{cases}$$

with μ_r denoting the r-th moment of K, i.e. $\mu_r = \int z^r K(z)\,dz$ $(r = 1, 2, \cdots)$.

The asymptotic variance in the above result is exactly the same as that of Hjort and Jones' (1996) local likelihood density estimator. In fact, even the asymptotic bias coincides with the Hjort and Jones' estimator as in the next result :

Theorem 2. *Let x be an interior point of f. Then the following identities hold for $b_h(x) = f(x;\theta_h) - f(x)$ as $h \to 0$:*

(i) $p = 1$;

$$b_h(x) = \frac{1}{2}\mu_2\Big\{(f - f_0)^{(2)}(x) + 2\frac{u_0^{(1)}(x)}{u_0(x)}(f - f_0)^{(1)}(x)\Big\}h^2 + o(h^2),$$

(ii) $p = 2$;

$$b_h(x) = \frac{1}{2}\mu_2(f - f_0)^{(2)}(x)h^2 + o(h^2),$$

(iii) $p = 3$;

$$\begin{aligned} b_h(x) &= \frac{1}{2}\frac{\mu_2\mu_6 - \mu_4^2}{\mu_4 - \mu_2^2}\Big\{\frac{A}{9}(f - f_0)^{(3)}(x) - \frac{1}{12}(f - f_0)^{(4)}(x)\Big\}h^4 \\ &\quad + o(h^4), \end{aligned}$$

(iv) $p = 4$;

$$b_h(x) = -\frac{1}{24}\frac{\mu_2\mu_6 - \mu_4^2}{\mu_4 - \mu_2^2}(f - f_0)^{(4)}(x)h^4 + o(h^4),$$

where, for a function g, $g^{(k)}$ denotes the k-th derivative of the function, A is some constant to be specified in Section 3, and

$$f_0(t) = \lim_{h\to 0} f(t;\theta_h), \quad u_0(t) = \lim_{h\to 0} u_{\theta_h}(t).$$

Remark 2. It should be pointed out, as in Hjort and Jones (1996), that the convergence rate depends on the number of parameters in the same manner as that of the local polynomial regression estimator. In general, one may show, sufficient differentiability of $f(t,\theta)$ and $f(t)$ permitting, that the order of leading bias is h^{p+1} when p is odd and h^p when p is even. Deriving a general formula for the constant factor seems out of reach.

3. PROOFS

Proof of Theorem 1. Throughout the proof, it is assumed that $h \to 0$, $nh \to \infty$ as $n \to \infty$, and the convolution is denoted by $*$ with all the convolutions being evaluated at x, the point of interest. Write

$$\begin{aligned} B &= (u_0, u_0^{(1)}, \cdots, u_0^{(p-1)}/(p-1)!) \\ V_h(y) &= (1, hy, \cdots, h^{p-1}y^{p-1})^t \\ e_1 &= (1, 0, \cdots, 0)^t. \end{aligned}$$

Note that B is a $p \times p$ nonsingular matrix by (A.3), and

$$f(x)e_1^t\Big\{\int K(y)V_h(y)V_h(y)^t\,dy\Big\}^{-1}\Big\{\int K^2(y)V_h(y)V_h(y)^t\,dy\Big\} \times\Big\{\int K(y)V_h(y)V_h(y)^t\,dy\Big\}^{-1}e_1$$

is the asymptotic variance $\sigma^2(x)$ in Theorem 1.

It follows from (2.1) and (2.4) that

$$\widehat{f}_n(x) = f(x,\theta_h) + \dot{f}(x,\theta_h)^t\Big\{-E\ddot{L}_n(\theta_h)\Big\}^{-1}\dot{L}_n(\theta_h) + O_p((nh)^{-1}).$$

Note that $\dot{L}_n(\theta)$ can be written as

$$\dot{L}_n(\theta) = \frac{h}{a}\Big\{\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)u_\theta(X_i) - C_h(\theta)\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)\Big\} + C_h(\theta) \quad (3.1)$$

where $a = K(0)$, $C_h(\theta) = \frac{\partial}{\partial\theta}\log B_h(\theta)$ and $X_1, \cdots, X_n$ are independent random variables with density f. Since θ_h is chosen so that $\dot{L}_n(\theta_h)$ has mean 0 by (2.3), $\dot{L}_n(\theta_h)$ asymptotically obeys a normal distribution with mean 0 under the assumptions (A.1)$\sim$(A.5). Note further that

$$C_h(\theta) = -\frac{h}{a}K_h * \{u_\theta f(\cdot;\theta)\}\Big\{1 - \frac{h}{a}K_h * f(\cdot;\theta)\Big\}^{-1} = O(h), \quad (3.2)$$

which implies

$$\begin{aligned} nh\,\mathrm{var}\dot{L}_n(\theta_h) &= \frac{h^2}{a^2}\Big[nh\,\mathrm{var}\Big\{\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)u_{\theta_h}(X_i)\Big\} + O(h)\Big] \\ &= \frac{h^2}{a^2}\Big\{K_h^2 * (u_0u_0^t)f(x) + O(h)\Big\} \\ &= \frac{h^2}{a^2}\Big\{B\int K^2(y)V_h(y)V_h^t(y)\,dy\; B^tf(x) + O(h)\Big\}. \end{aligned}$$

Thus the result will follow if we prove that

$$\begin{aligned} \dot{f}(x,\theta_h) &= f(x)e_1^tB^t + O(h) & (3.3) \\ -\mathrm{E}\ddot{L}_n(\theta_h) &= \frac{h}{a}\Big\{B\int K(y)V_h(y)V_h^t(y)\,dy\; B^tf(x) + O(h)\Big\}. & (3.4) \end{aligned}$$

To prove (3.3), we need

$$f(x;\theta_h) = f(x) + O(h^2)$$

which will be proved in the proof of Theorem 2. This implies

$$\begin{aligned}\dot{f}(x;\theta_h)^t &= f(x;\theta_h)u^t_{\theta_h}(x) \\ &= f(x)u^t_0(x) + O(h) \\ &= f(x)e^t_1 B^t + O(h).\end{aligned}$$

Now to prove (3.4), note that (3.1) and (3.2) imply

$$\begin{aligned}\dot{L}_n(\theta) &= \frac{h}{a}\left[\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)u_\theta(X_i) - K_h * \{u_\theta f(\cdot;\theta)\}\right. \\ &\left. + C_h(\theta)\Big\{K_h * f(\cdot;\theta) - \frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)\Big\}\right].\end{aligned}$$

Thus we have

$$\begin{aligned}-\mathrm{E}\ddot{L}_n(\theta) &= \frac{h}{a}\Big[K_h * \{u_\theta u^t_\theta f(\cdot;\theta)\} + K_h * \{\dot{u}_\theta(f(\cdot;\theta) - f)\} \\ &\quad -\dot{C}_h(\theta)K_h * \{f(\cdot;\theta) - f\} - C_h(\theta)K_h * \{u^t_\theta f(\cdot;\theta)\}\Big]\end{aligned}$$

which yields

$$\begin{aligned}-\mathrm{E}\ddot{L}_n(\theta_h) &= \frac{h}{a}\{K_h * (u_0 u^t_0)f(x) + O(h)\} \\ &= \frac{h}{a}\Big\{B\int K(y)V_h(y)V^t_h(y)\,dy\ B^t f(x) + O(h)\Big\}.\end{aligned}$$

Proof of Theorem 2. It follows from (3.1) and (3.2) that the equation (2.3) can be written as

$$K_h * (u_\theta v_\theta) + \frac{h}{a}\{K_h * (u_\theta f)K_h * v_\theta - K_h * (u_\theta v_\theta)K_h * f\} = 0 \qquad (3.5)$$

where $v_\theta(t) = f(t;\theta) - f(t)$. The left hand side of the equation (3.5) admits a Taylor expansion as a function of h under the assumptions (A.1) and (A.2). Thus θ_h, being the unique solution to the equation (3.5), admits a Taylor expansion by the implicit function theorem. Hence we may write

$$\begin{aligned}u_{\theta_h} &= u_0 + u_1 h + u_2 h^2 + \cdots + u_6 h^6 + o(h^6) \\ v_{\theta_h} &= v_0 + v_1 h + v_2 h^2 + \cdots + v_6 h^6 + o(h^6),\end{aligned}$$

and we can expand the left hand side of (3.5) with θ_h replacing θ. For this purpose, let C_r denote the coefficient of h^r $(r = 0, 1, \cdots, 6)$ in such an expansion and write

$$(u \times v)_r = \sum_{k=0}^{r} u_k v_{r-k}.$$

A little algebra shows that

$$\begin{aligned}
C_0 &= u_0 v_0, \quad C_1 = u_0 v_1 + v_0 u_1 \\
C_2 &= (u \times v)_2 + \frac{1}{2}\mu_2 (u \times v)_0^{(2)} \\
C_3 &= (u \times v)_3 + \frac{1}{2}\mu_2 (u \times v)_1^{(2)} + \mu_2 u_0^{(1)} (v_0 f^{(1)} - v_0^{(1)} f) \\
C_4 &= (u \times v)_4 + \frac{1}{2}\mu_2 (u \times v)_2^{(2)} + \frac{1}{24}\mu_4 (u \times v)_0^{(4)} \\
&\quad + \mu_2 \{ u_0^{(1)} (v_1 f^{(1)} - v_1^{(1)} f) + u_1^{(1)} (v_0 f^{(1)} - v_0^{(1)} f) \} \\
C_5 &= (u \times v)_5 + \frac{1}{2}\mu_2 (u \times v)_3^{(2)} + \frac{1}{24}\mu_4 (u \times v)_1^{(4)} \\
&\quad + \mu_2 \{ u_0^{(1)} (v_2 f^{(1)} - v_2^{(1)} f) + u_1^{(1)} (v_1 f^{(1)} - v_1^{(1)} f) \\
&\qquad + u_2^{(2)} (v_0 f^{(1)} - v_0^{(1)} f) \} \\
&\quad + \frac{1}{4}\mu_2^2 \{ 2u_0^{(1)} (v_0^{(2)} f^{(1)} - v_0^{(1)} f^{(2)}) + u_0^{(2)} (v_0^{(2)} f - v_0 f^{(2)}) \} \\
&\quad + \frac{1}{24}\mu_4 \{ 4u_0^{(1)} (v_0 f^{(3)} - v_0^{(3)} f) + 6u_0^{(2)} (v_0 f^{(2)} - v_0^{(2)} f) \\
&\qquad + 4u_0^{(3)} (v_0 f^{(1)} - v_0^{(1)} f) \} \\
C_6 &= (u \times v)_6 + \frac{1}{2}\mu_2 (u \times v)_4^{(2)} + \frac{1}{24}\mu_4 (u \times v)_2^{(4)} + \frac{1}{720}\mu_6 (u \times v)_0^{(6)} \\
&\quad + \mu_2 \{ (u^{(1)} \times v)_3 f^{(1)} - (u^{(1)} \times v^{(1)})_3 f \} \\
&\quad + \frac{1}{4}\mu_2^2 \{ 2(u^{(1)} \times v^{(1)})_1 f^{(1)} - 2(u^{(1)} \times v^{(1)})_1 f^{(2)} \\
&\qquad + (u^{(2)} \times v^{(2)})_1 f - (u^{(2)} \times v)_1 f^{(2)} \} \\
&\quad + \frac{1}{24}\mu_4 \{ 4(u^{(1)} \times v)_1 f^{(3)} + 6(u^{(2)} \times v)_1 f^{(2)} + 4(u^{(3)} \times v)_1 f^{(1)} \\
&\qquad - 4(u^{(1)} \times v^{(3)})_1 f - 6(u^{(2)} \times v^{(2)})_1 f - 4(u^{(3)} \times v^{(1)})_1 f \}.
\end{aligned}$$

Setting these equal to 0 and using (A.3) that $u_0, u_0^{(1)}, \cdots, u_0^{(p-1)}$ are linearly independent, we obtain the following :

(i) $p = 1$;

$$v_0 = v_1 = 0, \quad v_2 = -\frac{1}{2}\mu_2(v_0^{(2)} + 2\frac{u_0^{(1)}}{u_0}v_0^{(1)}),$$

which yield

$$f(x;\theta_h) - f(x) = \frac{1}{2}\mu_2\Big\{(f-f_0)^{(2)}(x) + 2\frac{u_0^{(1)}(x)}{u_0(x)}(f-f_0)^{(1)}(x)\Big\}h^2 + o(h^2),$$

(ii) $p = 2$; In addition to the above,

$$v_0^{(1)} = v_1^{(1)} = 0, \quad v_2 = -\frac{1}{2}\mu_2 v_0^{(2)}, \quad v_3 = -\frac{1}{2}\mu_2 v_1^{(2)},$$

which yield

$$f(x;\theta_h) - f(x) = \frac{1}{2}\mu_2(f-f_0)^{(2)}(x)h^2 + o(h^2).$$

(iii) $p = 3$; In addition to the above,

$$\begin{aligned} &v_2 = v_3 = 0, \quad v_0^{(2)} = 0, \quad v_1^{(2)} = 0, \\ &v_4 = -\frac{1}{2}\mu_2 v_2^{(2)} - \frac{1}{24}\mu_4 v_0^{(4)}, \quad v_2^{(1)} = -\frac{1}{6}\frac{\mu_4}{\mu_2}v_0^{(3)}, \end{aligned}$$

$$\begin{aligned} &u_0\Big(v_6 + \frac{1}{2}\mu_2 v_4^{(2)} + \frac{1}{24}\mu_4 v_2^{(4)} + \frac{1}{720}\mu_6 v_0^{(6)}\Big) \\ &+u_0^{(1)}\Big(\mu_2 v_4^{(1)} + \frac{1}{6}\mu_4 v_2^{(3)} + \frac{1}{120}\mu_6 v_0^{(5)}\Big) \\ &+u_0^{(2)}\Big(\frac{\mu_4 - \mu_2^2}{4}v_2^{(2)} + \frac{\mu_6 - \mu_2\mu_4}{48}v_0^{(4)}\Big) \\ &+u_0^{(3)}\frac{\mu_2\mu_6 - \mu_4^2}{36\mu_2}v_0^{(3)} \\ &= 0, \end{aligned}$$

which yield

$$f(x;\theta_h) - f(x) = \frac{1}{2}\frac{\mu_2\mu_6 - \mu_4^2}{\mu_4 - \mu_2^2}\Big\{\frac{A}{9}(f-f_0)^{(3)}(x) - \frac{1}{12}(f-f_0)^{(4)}(x)\Big\}h^4 + o(h^4),$$

where $A = \left\{ \frac{\mu_4 - \mu_2^2}{4} v_2^{(2)} + \frac{\mu_6 - \mu_2\mu_4}{48} v_0^{(4)} \right\} / \frac{\mu_2\mu_6 - \mu_4^2}{36\mu_2} v_0^{(3)}$.

(iv) $p = 4$; In addition to the above,

$$v_0^{(3)} = 0, \quad v_0^{(1)} = 0, \quad v_6 + \frac{1}{2}\mu_2 v_4^{(2)} + \frac{1}{24}\mu_4 v_2^{(4)} + \frac{1}{720}\mu_6 v_0^{(6)} = 0$$

$$\mu_2 v_4^{(1)} + \frac{1}{6}\mu_4 v_2^{(3)} + \frac{1}{120}\mu_6 v_0^{(5)} = 0$$

$$v_4 = \frac{1}{24}\frac{\mu_6\mu_2 - \mu_4^2}{\mu_4 - \mu_2^2} v_0^{(4)}, \quad v_2^{(2)} = -\frac{1}{12}\frac{\mu_6 - \mu_4\mu_2}{\mu_4 - \mu_2^2} v_0^{(4)},$$

which yield

$$f(x;\theta_h) - f(x) = -\frac{1}{24}\frac{\mu_6\mu_2 - \mu_4^2}{\mu_4 - \mu_2^2}(f - f_0)^{(4)}(x)h^4 + o(h^4).$$

ACKNOWLEDGEMENT

The authors wish to acknowledge the support of the Korean Research Foundation (1998-015-D00047) made in the program year of 1998. The work is supported also by the Brain Korea 21 Project.

REFERENCES

Copas, J. B. (1995). Local likelihood based on kernel censoring. *Journal of Royal Statistical Society*, Series B **57**, 221-235.

Eguchi, S. and Copas, J. B. (1998). A class of local likelihood methods and near-parametric asymptotics. *Journal of Royal Statistical Society*, Series B **60**, 709-724.

Hjort, N. L. and Jones, M. C. (1996). Locally parametric nonparametric density estimation. *The Annals of Statistics* **24**, 1619-1647.

Loader, C. R. (1996). Local likelihood density estimation. *The Annals of Statistics* **24**, 1602-1618.