Research Article

The Local Linear M-Estimation with Missing Response Data

Shuanghua Luo,[1,2] Cheng-Yi Zhang,[1] and Fengmin Xu[3]

[1] *School of Science, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China*

[2] *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*

[3] *Institute of Information and System Science and School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*

Correspondence should be addressed to Fengmin Xu; fengminxu@mail.xjtu.edu.cn

This paper studies the nonparametric regressive function with missing response data. Three local linear M-estimators with the robustness of local linear regression smoothers are presented such that they have the same asymptotic normality and consistency. Then finite-sample performance is examined via simulation studies. Simulations demonstrate that the complete-case data M-estimator is not superior to the other two local linear M-estimators.

1. Introduction

Local polynomial regression methods, which have advantages over popular kernel methods in terms of the ability of design adaption and high asymptotic efficiency, have been demonstrated as effective nonparametric smoothers. In addition, the local polynomial regression smoothers can adapt to almost all regression settings and cope very well with the edge effects. For details, see [1] and references therein. However, a drawback of these local regression estimators is lack of robustness. It is well known that M-type of regression estimators have many desirable robustness properties. As a result, M-type of regression estimators is natural candidates.

In fact, some methods such as kernel, local regression, spline, and orthogonal series methods can estimate nonparametric functions. For an introduction to this subject area, see [1–3]. Local linear smoother, as an intuitively appealing method, has become popular in recent years because of its attractive statistical properties. As is shown in [4–7], local regression provides many advantages over modified kernel methods. Consequently, it is reasonable to expect that the local regression based M-type estimators carry over those advantages.

In the present paper, local M-type regression estimators are applied to propose three M-estimators of $m(x)$ with missing response data, including the complete-case data M-estimator, the weighted M-estimator, and the estimated weighted M-estimator, such that these estimators have the same asymptotic normality and consistency. Finite sample simulations show that the complete-case data M-estimator is not superior to the other two local linear M-estimators.

In the regression analysis setup, the basic inference begins by considering the random sample

$$(X_i, Y_i, \delta_i), \quad i = 1, 2, \ldots, n, \tag{1}$$

where the design point X_i is observed and

$$\delta_i = \begin{cases} 0, & Y_i \text{ is missing}, \\ 1, & \text{otherwise}. \end{cases} \tag{2}$$

Theoretically, this is actually a missing response problem. Recently, considerable interest in the work on nonparametric regression analysis with missing data and many methods has been developed. Cheng (see [8]) employed a kernel regression imputation approach to define the estimator of the mean of Y. Hirano et al. (see [9]) defined a weighted estimator for the response mean when response variable is missing. Then, Wang et al. (see [10]) developed estimation theory for semiparametric regression analysis in the presence of missing response. In the last year, Liang [11] and Wang and Sun [12] discussed the generalized partially linear models with

missing covariate data and the partially linear models with missing responses at random, respectively.

To study the missing data (1), the MAR assumption would require that there exists a chance mechanism, denoted by $p(X)$, such that

$$P\left(\delta = 1 \mid X, Y\right) = P\left(\delta = 1 \mid X\right) = p\left(X\right) \qquad (3)$$

holds almost surely. In practice, (3) might be justifiable by the nature of the experiment when it is legitimate to assume that the missing of Y mainly depends on X.

The paper is organized as follows. Some notations and preliminary results about three M-estimators of $m(x)$ with missing response data including the local M-estimator with the complete-case data, the weighted M-estimator, and the estimated weighted M-estimator are given in Section 2. The asymptotic normality and consistency of three M-estimators are then presented in Section 3. In Section 4, simulation studies give some comparison results of the proposed estimators. Sketches of the proofs are given in Section 5.

2. Model and Estimators

2.1. The Models. The nonparametric regression model that we will consider for the incomplete data (1) is given by

$$Y_i = m\left(X_i\right) + \varepsilon_i \qquad (4)$$

for $i = 1, 2, \ldots, n$. Here $m(\cdot)$ is the regression function, X_i is design point, $\varepsilon_i = Y_i - m(X_i)$ is regression error, and Y_i is a response variable. The regression error ε_i is conveniently assumed to be independent and identically distributed (i.i.d.) random variable with mean 0 and variance $\sigma^2 \in (0, \infty)$. Furthermore, the variance of ε_i will be assumed to depend on X_i, denoted by $\sigma^2(X_i)$, allowing for heteroscedasticity. To simplify our preliminary discussion the covariate X_i will be assumed to be real-valued.

2.2. The Local M-Estimator with the Complete-Case Data. The local M-estimator with the complete-case data is defined as the solution of the following problems; that is, find a and b to minimize

$$\sum_{i=1}^{n} \rho\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \delta_i, \qquad (5)$$

or to satisfy the local estimation system of equations

$$\Psi_{1cc}\left(a, b\right) = \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \delta_i = 0,$$

$$\Psi_{2cc}\left(a, b\right)$$

$$= \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) \frac{X_i - x}{h_n} K\left(\frac{X_i - x}{h_n}\right) \delta_i = 0,$$

$$(6)$$

where $\rho(\cdot)$ is a given outlier-resistant function, $\psi(\cdot)$ is the derivative of $\rho(\cdot)$, $K(\cdot)$ is a kernel function, and h_n is a sequence of positive numbers tending to zero.

The M-estimations of $m(x)$ and $m'(x)$ are defined as $\widehat{a}$ and $\widehat{b}$, which are the solution to the system of (6). We denote them by $\widehat{m}_{Mcc}(x)$ and $\widehat{m}'_{Mcc}(x)$, respectively.

2.3. The Weighted M-Estimator. If we find a and b to minimize

$$\sum_{i=1}^{n} \rho\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{p\left(X_i\right)}, \qquad (7)$$

or to satisfy the local estimation system of equations

$$\Psi_{1cp}\left(a, b\right)$$

$$= \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{p\left(X_i\right)} = 0,$$

$$\Psi_{2cp}\left(a, b\right)$$

$$= \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) \frac{X_i - x}{h_n} K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{p\left(X_i\right)} = 0,$$

$$(8)$$

then the solutions a and b are called the weighted M-estimator. Furthermore, the solutions of the system of (8) will be denoted by $\widehat{m}_{Mcp}(x)$ and $\widehat{m}'_{Mcp}(x)$, respectively.

2.4. The Estimated Weighted M-Estimator. In practice, the selection probabilities function $p(x)$ is usually unknown. To estimate the selection probabilities, we apply the local linear smoother to find a and b such that

$$\sum_{i=1}^{n} \left[\delta_i - a - b\left(X_i - x\right)\right]^2 K\left(\frac{X_i - x}{h_n}\right) \qquad (9)$$

is minimized. A straightforward calculation yields

$$\widehat{p}\left(x\right) = \frac{\sum_{i=1}^{n} v_i \delta_i}{\sum_{j=1}^{n} v_j}, \qquad (10)$$

where

$$v_j = K\left(\frac{X_j - x}{h_n}\right)\left[m_{n,2} - \left(X_j - x\right) m_{n,1}\right],$$

$$m_{n,l} = \sum_{j=1}^{n} K\left(\frac{X_j - x}{h_n}\right)\left(X_j - x\right)^l, \quad l = 0, 1, 2.$$

$$(11)$$

With the estimator $\widehat{p}(x)$ defined in (10), we can give the estimated weighted M-estimator by finding a and b to minimize

$$\sum_{i=1}^{n} \rho\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{\widehat{p}\left(X_i\right)}, \qquad (12)$$

or to satisfy the local estimation system of equations

$$\Psi_{1wp}(a, b)$$

$$= \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{\widehat{p}(X_i)} = 0,$$

$$\Psi_{2wp}(a, b)$$

$$= \sum_{i=1}^{n} \psi\left(Y_i - a - b\left(X_i - x\right)\right) \frac{X_i - x}{h_n} K\left(\frac{X_i - x}{h_n}\right) \frac{\delta_i}{\widehat{p}(X_i)}$$

$$= 0.$$

$$(13)$$

By $\widehat{m}_M^*(x)$ and $\widehat{m}_M^{*\prime}(x)$ denote the solutions a and b of the system of (13), respectively.

3. Main Results

In this section, we will present the main results of this paper to explore the asymptotic distribution and consistencies of $\widehat{m}_{Mcc}(x)$, $\widehat{m}_{Mcp}(x)$, and $\widehat{m}_M^*(x)$. The following conditions will be used in the rest of this section.

(1) The regression function $m(\cdot)$ has a continuous second derivative at the given point x and $m''(x)$ is continuous and bounded on support field D.

(2) The sequence of bandwidths h_n tends to zero such that $nh_n \to +\infty$.

(3) The design density $f(\cdot)$ is continuous at the given point x and $f(x) > 0$.

(4) The kernel function $K(\cdot)$ is a continuous probability density function with bounded support field $[-1, 1]$.

(5) $E[\psi(\varepsilon) \mid X = x] = 0$ with $\varepsilon = Y - m(X)$.

(6) The function $\psi(\cdot)$ is continuous and has a derivative $\psi'(\cdot)$. Further, assume that $\psi_\varepsilon(x) = E[\psi'(\varepsilon) \mid X = x]$ and $\psi_\varepsilon^2(x) = E[\psi^2(x) \mid X = x]$ are positive and continuous at the given point x and there exists $r > 0$ such that $E[|\psi^{2+r}(\varepsilon)| \mid X = x]$ is bounded in a neighborhood of x.

(7) The function $\psi'(\cdot)$ satisfies that

$$E\left[\sup_{|z| \le \delta}\left|\psi'(\varepsilon + z) - \psi'(\varepsilon)\right| \mid X = x\right] = o(1),$$

$$E\left[\sup_{|z| \le \delta}\left|\psi(\varepsilon + z) - \psi(\varepsilon) - \psi'(\varepsilon) z\right| \mid X = x\right] = o(\delta),$$

$$(14)$$

as $\delta \to 0$, uniformly in a neighborhood of x.

In addition, for the convenience of representation and proof, $\widehat{m}_M(x)$ will denote three estimators $\widehat{m}_{Mcc}(x), \widehat{m}_{Mcp}(x)$, and $\widehat{m}_M^*(x)$, and $\widehat{m}_M'(x)$ will denote the estimators $\widehat{m}_{Mcc}'(x)$, $\widehat{m}_{Mcp}'(x)$, and $\widehat{m}_M^{*\prime}(x)$. Lastly, assume $s_l = \int_{-\infty}^{+\infty} u^l K(u)du < \infty$ for $l = 0, 1, 2$.

The following will present the main theorems such that the three M-estimators mentioned above have the same consistency and asymptotic normality.

Theorem 1. *Under conditions (1)–(7), $f(\cdot)$, $p''(\cdot)$ are bounded; furthermore, $f(\cdot)$ also satisfies Lipchitz's condition; that is, $|f(x) - f(y)| \le c|x - y|$ for $c > 0$. If $h_n = cn^{-\beta}$, $0 < \beta < 1$, then*

$$\widehat{m}_M(x) - m(x) \xrightarrow{\mathscr{P}} 0, \quad n \longrightarrow \infty,$$

$$h_n\left(\widehat{m}_M'(x) - m'(x)\right) \xrightarrow{\mathscr{P}} 0, \quad n \longrightarrow \infty.$$

$$(15)$$

Theorem 2. *Under conditions (1)–(7), $f(\cdot)$, $p''(\cdot)$ are bounded and $f(\cdot)$ also satisfies Lipchitz's condition; that is, $|f(x) - f(y)| \le c|x - y|$ for $c > 0$. If $h_n = cn^{-\beta}, 0 < \beta < 1$, then*

$$\sqrt{nh_n}\left(\widehat{m}_M(x) - m(x) - C_n\right) \xrightarrow{\mathscr{L}} N(0, \pi(x)), \quad (16)$$

where $C_n = (1/2)m''(x)h_n^2 s_2(1 + o_p(1))$ and $\pi(x) = ((\psi_\varepsilon^2(x)/[\psi_\varepsilon(x)]^2)(1/f(x)p(x))) \int_{-\infty}^{+\infty} K^2(u)du$.

4. Simulation Studies

In this section, we conducted some simulations to better understand the finite-sample performance of the present three M-estimators. Then, we will compare the biases, the sample mean square errors (MSE), and the sample mean average square errors (MASE) of three M-estimators. The MASE of $\widehat{m}(x)$ is defined by

$$\text{MASE}(h) = E\left[N^{-1}\sum_{i=1}^{N}(\widehat{m}(x_i) - m(x_i))^2 W(x_i)\right], \quad (17)$$

where $W(x_i) = I_{[h, 1-h]}(x_i)$ and $\{x_i, i = 1, \ldots, N\}$ is grid points.

The sample size was set $n = 500$ and the regression model $Y_i = X_i^3 + \varepsilon_i$ was considered, where X_i is a uniform $(0, 1)$ and ε_i $(i = 1, 2, \ldots, n)$ is a random sample from $N(0, 0.1)$ and is independent of X_i. The kernel function was taken as the Epanechnikov kernel: $K(u) = 0.75(1 - u^2)_+$ and $\psi(x) = x$. To generate the indicators δ_i, the function $p(x)$ was chosen as $p(x) = 0.8$, for all $x \in [0, 1]$, and the pseudo i.i.d. uniform random variables $U(0, 1)$. As a result, each $U_i \in U(0, 1)$ was generated. If $U_i \le 0.8$, then $\delta_i = 1$; otherwise, $\delta_i = 0$. By these, 500 independent sets of data were generated.

For $\widehat{m}_M^*(x)$, $\widehat{m}_{Mcp}(x)$, and $\widehat{m}_{Mcc}(x)$, the optimal bandwidth h_{opt} was obtained by 500 simulations. Let h_{opt}^i be the optimal bandwidth of the $\text{MAES}(h_n)$ obtained by the ith simulation and then compute $h_{\text{opt}} = (1/500) \sum_{i=1}^{500} h_{\text{opt}}^i$. When $h_{\text{opt}} = 0.2$, we compare the bias, MSE, and MAES of the present three M-estimators. The comparison results are shown in Figures 1–3.

Figure 1 shows the biases of $\widehat{m}_M^*(x)$, $\widehat{m}_{Mcp}(x)$, and $\widehat{m}_{Mcc}(x)$. It is easy to see that $\widehat{m}_{Mcc}(x)$ has considerable bias while $\widehat{m}_{Mcp}(x)$ and $\widehat{m}_M^*(x)$ are very close in most points.
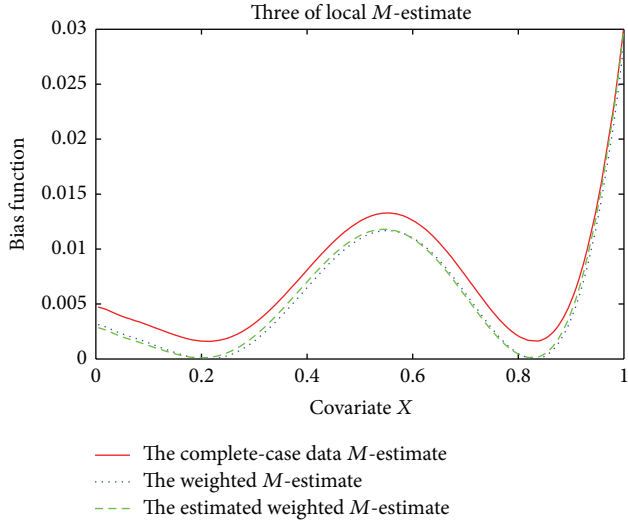
FIGURE 1: Comparison on the biases of three M-estimators, where the solid curve, the dashed curve, and the dot curve denote the biases of $\widehat{m}_{Mcc}(x)$, $\widehat{m}_M^*(x)$, and $\widehat{m}_{Mcp}^*(x)$, respectively.

Figure 2 demonstrates the MSE of $\widehat{m}_M^*(x)$, $\widehat{m}_{Mcp}(x)$, and $\widehat{m}_{Mcc}(x)$. It follows from Figure 2 that $\widehat{m}_{Mcc}(x)$ has considerable MSE while $\widehat{m}_{Mcp}(x)$ and $\widehat{m}_M^*(x)$ are very close in most points.

Figure 3 shows the MASE of $\widehat{m}_M^*(x)$, $\widehat{m}_{Mcp}(x)$, and $\widehat{m}_{Mcc}(x)$. We can see that $\widehat{m}_{Mcc}(x)$ has considerable MMSE while $\widehat{m}_{Mcp}(x)$ and $\widehat{m}_M^*(x)$ are very close in most points.

The comparison results on the biases, the MSE, and the MASE of three M-estimators obtained by Figures 1, 2, and 3 show that the weighted M-estimator and the estimated weighted M-estimator are obviously superior to the complete-case data M-estimator while there is no appreciable distinction on the superiority between weighted M-estimator and the estimated weighted M-estimator.

5. The Proofs of Theorems

The proofs of Theorems 1 and 2 will be given in this section, respectively. The following lemmas will be needed for our technical proofs.

Lemma 3. *Under conditions (1)–(7) and for any random sequence $\{\eta_j\}_{j=1}^n$, if $\max_{1 \le j \le n}|\eta_j| = o_p(1)$ and $K_j = K((X_j - x)/h_n)$, then*

$$\frac{1}{n}\sum_{j=1}^n \psi'\left(\varepsilon_j + \eta_j\right)\delta_j K_j\left(X_j - x\right)^l$$

$$= \psi_\varepsilon(x) h_n^{l+1} f(x) p(x) s_l\left(1 + o_p(1)\right),$$

$$\frac{1}{n}\sum_{j=1}^n \psi'\left(\varepsilon_j + \eta_j\right) R\left(X_j\right)\delta_j K_j\left(X_j - x\right)^l$$

$$= \frac{1}{2}\psi_\varepsilon(x) h_n^{l+3} m''(x) f(x) p(x) s_{l+2}\left(1 + o_p(1)\right),$$

where $R(X_j) = m(X_j) - m(x) - m'(x)(X_j - x)$.

(18)

FIGURE 2: Comparison on the MSE of three M-estimators, where the solid curve, the dashed curve, and the dot curve denote the MSE of $\widehat{m}_{Mcc}(x)$, $\widehat{m}_M^*(x)$, and $\widehat{m}_{Mcp}^*(x)$, respectively.
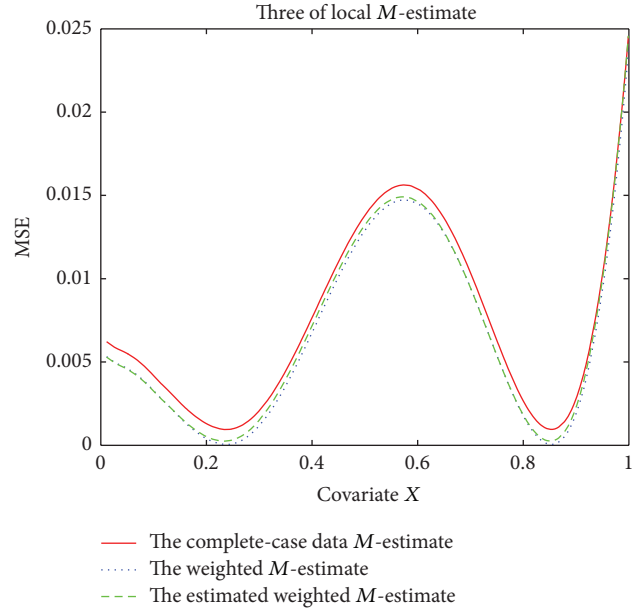

FIGURE 3: Comparison on the MASE of three M-estimators, where the solid curve, the dashed curve, and the dot curve denote the MASE of $\widehat{m}_{Mcc}(x)$, $\widehat{m}_M^*(x)$, and $\widehat{m}_{Mcp}^*(x)$, respectively.
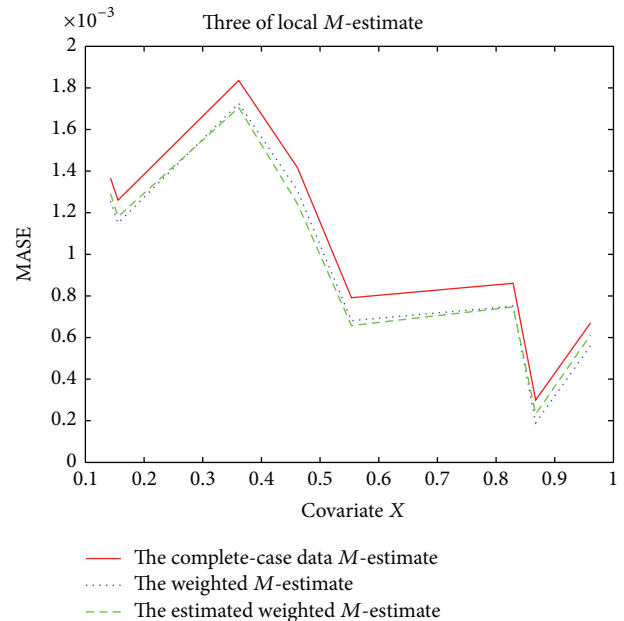
Proof. Since the second equality can be obtained from the first one, we only prove the first equality. It is obvious that

$$\frac{1}{n}\sum_{j=1}^n \psi'\left(\varepsilon_j + \eta_j\right)\delta_j K_j\left(X_j - x\right)^l$$

$$= \frac{1}{n}\sum_{j=1}^n \psi'\left(\varepsilon_j\right)\delta_j K_j\left(X_j - x\right)^l$$

$$+ \frac{1}{n} \sum_{j=1}^{n} \left[\psi' \left(\varepsilon_j + \eta_j \right) - \psi' \left(\varepsilon_j \right) \right] \delta_j K_j \left(X_j - x \right)^l$$

$$= T_{n,1} + T_{n,2},$$

(19)

where

$$T_{n,1} = \frac{1}{n} \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \delta_j K_j \left(X_j - x \right)^l,$$

$$T_{n,2} = \frac{1}{n} \sum_{j=1}^{n} \left[\psi' \left(\varepsilon_j + \eta_j \right) - \psi' \left(\varepsilon_j \right) \right] \delta_j K_j \left(X_j - x \right)^l.$$

(20)

Similar to the proof of Lemma 4 in [13], we have

$$ET_{n,1} = E \left[\frac{1}{n} \sum_{j=1}^{n} E \left(\psi' \left(\varepsilon_j \right) \mid X_j \right) \delta_j K_j \left(X_j - x \right)^l \right]$$

$$= \psi_\varepsilon (x) h_n^{l+1} f(x) p(x) s_l \left(1 + o_p (1) \right),$$

(21)

$$T_{n,1} = \psi_\varepsilon (x) h_n^{l+1} f(x) p(x) s_l \left(1 + o_p (1) \right).$$

The following only proves $T_{n,2} = o_p(h_n^{l+1})$. Let $\Delta_n = (\xi_1, \ldots, \xi_n)^T$, $D_\eta = \{ \Delta_n : |\xi_j| \le \eta, \forall j \le n \}$ for any given $\eta > 0$ and

$$V \left(\Delta_n \right) = \frac{1}{h_n^{l+1}} \sum_{j=1}^{n} \left[\psi' \left(\varepsilon_j + \xi_j \right) - \psi' \left(\varepsilon_j \right) \right] \delta_j K_j \left(X_j - x \right)^l.$$

(22)

Then we have

$$\sup_{D_n} \left| V \left(\Delta_n \right) \right|$$

$$\le \frac{1}{h_n^{l+1}} \sum_{j=1}^{n} \sup_{D_n} \left| \psi' \left(\varepsilon_j + \xi_j \right) - \psi' \left(\varepsilon_j \right) \right| \delta_j K_j \left(X_j - x \right)^l.$$

(23)

By condition (7) and noticing that $|X_j - x| \le h_n$ in the above expressions, it is not difficult to see

$$E \sup_{D_n} \left| V \left(\Delta_n \right) \right| \le a_\eta \frac{1}{h_n^{l+1}} E \sum_{j=1}^{n} \delta_j K_j \left(X_j - x \right)^l \le b_\eta,$$

(24)

where a_η and b_η are two sequences of positive numbers, tending to zero as $\eta \to 0$. Since $\sup_{1 \le j \le n} |\eta_j| = o_p(1)$, it follows that $V(\widehat{\Delta}_n) = o_p(1)$ with $\widehat{\Delta}_n = (\eta_1, \ldots, \eta_n)^T$. The conclusion is obtained coming from the fact that $T_{n,2} = h_n^{l+1} V(\widehat{\Delta}_n) = o_p(h_n^{l+1})$.   □

Lemma 4. *Under conditions (1)–(7) and for any random sequence $\{\eta_j\}_{j=1}^n$, if $\max_{1 \le j \le n} |\eta_j| = o_p(1)$ and $K_j = K((X_j - x)/h_n)$, then*

$$\frac{1}{n} \sum_{j=1}^{n} \psi' \left(\varepsilon_j + \eta_j \right) \frac{\delta_j}{p \left(X_j \right)} K_j \left(X_j - x \right)^l$$

$$= \psi_\varepsilon (x) h_n^{l+1} f(x) s_l \left(1 + o_p (1) \right),$$

$$\frac{1}{n} \sum_{j=1}^{n} \psi' \left(\varepsilon_j + \eta_j \right) R \left(X_j \right) \frac{\delta_j}{p \left(X_j \right)} K_j \left(X_j - x \right)^l$$

$$= \frac{1}{2} \psi_\varepsilon (x) h_n^{l+3} m'' (x) f(x) s_{l+2} \left(1 + o_p (1) \right),$$

(25)

where $R(X_j) = m(X_j) - m(x) - m'(x)(X_j - x)$.

Proof. The proof is similar to the proof of Lemma 3.   □

Lemma 5. *Under conditions (1)–(7), $J_n / \sqrt{nh_n}$ is asymptotically normal and*

$$\frac{J_n}{\sqrt{nh_n}} \xrightarrow{\mathscr{L}} N \left(0, D(x) \right),$$

(26)

where $J_n = \sum_{j=1}^{n} \psi(\varepsilon_j) \delta_j K((X_j - x)/h_n)$ and $D(x) = \psi_\varepsilon^2(x) f(x) p(x) \int_{-\infty}^{+\infty} K^2(u) du$.

Proof. Let $J_n = \sum_{j=1}^{n} \psi(\varepsilon_j) K((X_j - x)/h_n) \delta_j \triangleq \sum_{j=1}^{n} \xi_j$. Then J_n is a sum of i.i.d. random variables with mean zero and variance B_n^2 with

$$B_n^2 = nE \left\{ \psi^2 \left(\varepsilon_1 \right) K^2 \left(\frac{X_1 - x}{h_n} \right) \delta_1^2 \right\}.$$

(27)

Similar to the proof of Lemma 4 in [13], we can easily obtain asymptotic expression of B_n^2, namely,

$$B_n^2 = nh_n \psi_\varepsilon^2 (x) f(x) p(x) \int_{-\infty}^{+\infty} K^2 (u) du,$$

(28)

and easily verify Lyapunov's condition $(1/B_n^{2+r}) \sum_{j=1}^{n} E|\xi_j|^{2+r} \to 0$ via using condition (6). That is, J_n is asymptotically normal. With (28) we have

$$\frac{J_n}{\sqrt{nh_n}} \xrightarrow{\mathscr{L}} N \left(0, D(x) \right).$$

(29)

This completes the proof of this lemma.   □

Lemma 6. *Under conditions (1)–(7), $\sqrt{nh_n} \widehat{J}_n$ was asymptotically normal and*

$$\sqrt{nh_n} \widehat{J}_n \xrightarrow{\mathscr{L}} N \left(0, D^* (x) \right),$$

(30)

where $\widehat{J}_n = (1/nh_n) \sum_{j=1}^{n} \psi(\varepsilon_j) (\delta_j / p(X_j)) K((X_j - x)/h_n)$ and $D^*(x) = \psi_\varepsilon^2(x) (f(x)/p(x)) \int_{-\infty}^{+\infty} K^2(u) du$.

Proof. The proof is similar to the proof of Lemma 5. □

Lemma 7. *Under conditions (1)–(7), $f(\cdot)$ and $p''(\cdot)$ are bounded; furthermore, $f(\cdot)$ also satisfies Lipchitz's condition; that is, $|f(x) - f(y)| \leq c|x - y|$ for $c > 0$. If $h_n = cn^{-\beta}, 0 < \beta < 1$, then the following equalities hold:*

① $1/\widehat{p}(x) = (1/p(x))(1 + O_r(h_n + 1/\sqrt{nh_n}))$;

② $B_n = (1/nh_n) \sum_{j=1}^{n} K((X_j - x)/h_n)\varepsilon_j(\delta_j/p^2(X_j)) \times (\widehat{p}(X_j) - p(X_j)) = o_p(1)$;

③ $G_n = (1/nh_n) \sum_{j=1}^{n} \psi(\varepsilon_j)K((X_j - x)/h_n)(\delta_j/p^2(X_j)) \times (\widehat{p}(X_j) - p(X_j)) = o_p(1)$.

Proof. We prove ① firstly.

① Since

$$\frac{1}{\widehat{p}(x)} = \frac{\sum_{i=1}^{n} v_i}{\sum_{j=1}^{n} v_j \delta_j}, \tag{31}$$

equality (31) is changed into

$$\frac{1}{\widehat{p}(x)} = \frac{\sum_{i=1}^{n} v_i}{\left(\sum_{j=1}^{n} v_j \delta_j + n^{-2}\right)} + \frac{n^{-2} \sum_{i=1}^{n} v_i}{\left\{\sum_{j=1}^{n} v_j \delta_j \left(\sum_{j=1}^{n} v_j \delta_j + n^{-2}\right)\right\}}. \tag{32}$$

Now, we denote $Z_n = O_r(a_n)$ if $E|Z_n|^r = O(a_n^r)$. It is easy to see that

(i) $O_r(a_n) O_r(b_n) = O_{r/2}(a_n b_n)$,

(ii) $Z_n = E(Z_n) + O_r\left(\left(E|Z_n - EZ_n|^r\right)^{1/r}\right)$. $\tag{33}$

Then we have with the method of the kernel density estimate that

$$Em_{n,l} = nh_n^{l+1} f(x) s_l (1 + O(h_n)),$$
$$Es_{n,l} = nh_n^{l+1} f(x) p(x) s_l (1 + O(h_n)), \tag{34}$$

where

$$s_{n,l} = \sum_{j=1}^{n} \delta_j K\left(\frac{X_j - x}{h_n}\right)(X_j - x)^l, \quad l = 0, 1, 2. \tag{35}$$

By the operations property of (2) for an integer $r > 1$,

$$\frac{1}{nh_n^{l+1}} m_{n,l} = \frac{1}{nh_n^{l+1}} Em_{n,l} + O_r\left(\frac{1}{\sqrt{nh_n}}\right)$$
$$= f(x) s_l + O_r\left(h_n + \frac{1}{\sqrt{nh_n}}\right), \tag{36}$$

$$\frac{1}{nh_n^{l+1}} s_{n,l} = \frac{1}{nh_n^{l+1}} Es_{n,l} + O_r\left(\frac{1}{\sqrt{nh_n}}\right)$$
$$= f(x) p(x) s_l + O_r\left(h_n + \frac{1}{\sqrt{nh_n}}\right). \tag{37}$$

Since $s_0 = 1$ and $s_1 = 0$,

$$\sum_{i=1}^{n} v_i = m_{n,0} m_{n,2} - m_{n,1}^2$$
$$= n^2 h_n^4 f^2(x) s_2 \left(1 + O_r\left(h_n + \frac{1}{\sqrt{nh_n}}\right)\right), \tag{38}$$

$$\sum_{j=1}^{n} v_j \delta_j = s_{n,0} m_{n,2} - s_{n,1} m_{n,1}$$
$$= n^2 h_n^4 f^2(x) p(x) s_2 \left(1 + O_r\left(h_n + \frac{1}{\sqrt{nh_n}}\right)\right). \tag{39}$$

Let $W_n = (\sum_{j=1}^{n} v_j \delta_j + n^{-2})/n^2 h_n^4$, $W = p(x) f^2(x) s_2$. The following will prove

$$\frac{1}{W_n} = \frac{1}{W} + o_4(1). \tag{40}$$

In fact, (40) holds only if $E((W/W_n) - 1)^4 = o(1)$ holds. Then,

$$E\left(\frac{W}{W_n} - 1\right)^4 = E\frac{(W_n - W)^4}{W_n^4} I\left(|W_n - W| \leq \frac{W}{2}\right)$$
$$+ E\frac{(W_n - W)^4}{W_n^4} I\left(|W_n - W| > \frac{W}{2}\right)$$
$$\leq \left(\frac{W}{2}\right)^{-4} E(W_n - W)^4$$
$$+ n^{16} E(W_n - W)^4 I\left(|W_n - W| > \frac{W}{2}\right)$$
$$\triangleq A_n + B_n, \tag{41}$$

where $A_n = (W/2)^{-4} E(W_n - W)^4$ and $B_n = n^{16} E(W_n - W)^4 I(|W_n - W| > W/2)$. It follows from (39) and the definition of W_n and W that

$$W_n - W = \frac{\sum_{j=1}^{n} v_j \delta_j + n^{-2}}{n^2 h_n^4} - p(x) f^2(x) s_2$$
$$= p(x) f^2(x) s_2 O_r\left(h_n + \frac{1}{\sqrt{nh_n}}\right). \tag{42}$$

Further, (41) and (42) yield $A_n = o(1)$. Again, since $W_n \geq n^{-4}$,

$$B_n \leq n^{16}\left(\frac{W}{2}\right)^{-r} E(W_n - W)^{r+4}$$
$$= O\left(n^{16}\left(h_n + \frac{1}{\sqrt{nh_n}}\right)^{r+4}\right), \tag{43}$$

which indicates $B_n = o(1)$ if r is sufficiently large. As a result, equality (40) holds. We conclude from (38) and (39) that this completes the proof of ①.

② Similar to the proof of Theorem 1 in [4], one can easily obtain

$$E\left\{(\widehat{p}(X_i) - p(X_i)) \mid X_i\right\}$$

$$= \frac{1}{2}h_n^2 p''(X_i) \int_{-\infty}^{+\infty} u^2 K(u)\, du \left\{1 + o_p(1)\right\}$$

$$= \frac{1}{2}h_n^2 C_1(X_i)\left\{1 + o_p(1)\right\},$$

$$\mathrm{Var}\left\{(\widehat{p}(X_i) - p(X_i)) \mid X_i\right\}$$

$$= \frac{1}{nh_n} \frac{p(X_i)(1 - p(X_i))}{f(X_i)} \int_{-\infty}^{+\infty} K^2(u)\, du \left\{1 + o_p(1)\right\}$$

$$= \frac{1}{nh_n} C_2(X_i)\left\{1 + o_p(1)\right\}.$$

$$(44)$$

As a consequence, we have with (44) that

$$\mathrm{Var}\, B_n$$

$$= \frac{1}{n^2 h_n^2}\,\mathrm{Var}$$

$$\times \left\{\sum_{j=1}^n K\left(\frac{X_j - x}{h_n}\right)\varepsilon_j \frac{\delta_j}{p^2(X_j)}\left(\widehat{p}(X_j) - p(X_j)\right)\right\}$$

$$= \frac{1}{n^2 h_n^2} E$$

$$\times \left\{\sum_{j=1}^n K\left(\frac{X_j - x}{h_n}\right)\varepsilon_j \frac{\delta_j}{p^2(X_j)}\left(\widehat{p}(X_j) - p(X_j)\right)\right\}^2$$

$$+ o_p(1)$$

$$\leq \frac{1}{n^2 h_n^2} E\left\{\sum_{j=1}^n K^2\left(\frac{X_j - x}{h_n}\right)\varepsilon_j^2 \frac{1}{p^4(X_j)}\right.$$

$$\left. \times E\left[\left(\widehat{p}(X_j) - p(X_j)\right)^2 \mid X_j\right]\right\} + o_p(1)$$

$$= \frac{1}{n^2 h_n^2} E\left\{\sum_{j=1}^n K^2\left(\frac{X_j - x}{h_n}\right)\varepsilon_j^2 \frac{1}{p^4(X_j)}\right.$$

$$\left. \times \mathrm{Var}\left(\widehat{p}(X_1) - p(X_1) \mid X_j\right)\right\}$$

$$+ \frac{1}{n^2 h_n^2} E\left\{\sum_{j=1}^n K^2\left(\frac{X_j - x}{h_n}\right)\varepsilon_j^2 \frac{1}{p^4(X_j)}\right.$$

$$\left. \times E^2\left((\widehat{p}(X_j) - p(X_j)) \mid X_j\right)\right\} + o_p(1)$$

$$= \frac{1}{n^3 h_n^3} E\left\{\sum_{j=1}^n K^2\left(\frac{X_j - x}{h_n}\right)\varepsilon_j^2 \frac{1}{p^4(X_j)}C_2(X_j)\right\}$$

$$\times \left(1 + o_p(1)\right)$$

$$+ \frac{h_n^2}{4n^2} E\left\{\sum_{j=1}^n K^2\left(\frac{X_j - x}{h_n}\right)\varepsilon_j^2 \frac{1}{p^4(X_j)}C_1(X_j)\right\}$$

$$\times \left(1 + o_p(1)\right)$$

$$= \frac{1}{n^2 h_n^3} E\left\{K^2\left(\frac{X_1 - x}{h_n}\right)\varepsilon_1^2 \frac{1}{p^4(X_1)}C_2(X_1)\right\}\left(1 + o_p(1)\right)$$

$$+ \frac{h_n^2}{4n} E\left\{K^2\left(\frac{X_1 - x}{h_n}\right)\varepsilon_1^2 \frac{1}{p^4(X_1)}C_1(X_1)\right\}\left(1 + o_p(1)\right)$$

$$= o_p(1).$$

$$(45)$$

This completes the proof of ②.

③ The proof is similar to the proof of ②.

□

Lemma 8. *Under conditions (1)–(7), if $nh_n^4 \to 0$ and $nh_n^2/\log(h_n^{-1}) \to \infty$ when $n \to \infty$, then*

$$\sup_{x \in D}\left|\widehat{p}(x) - p(x) - \frac{1}{nh_n f(x)}\sum_{j=1}^n \varepsilon_j K\left(\frac{X_j - x}{h_n}\right)\right|$$

$$= O_p\left\{c_n h_n^2 + c_n^2 \log^{1/2}\left(\frac{1}{h_n}\right)\right\},$$

$$(46)$$

where $D \subseteq R$ is a compact set, $c_n = (nh_n)^{-1/2}$, and $\widehat{p}(x)$ is the estimate value of $p(x)$.

Proof. The same arguments as those of Theorem 2 in [14] can yield the proof of this lemma. □

Lemma 9. *Under conditions (1)–(7), $f(\cdot)$ and $p''(\cdot)$ are bounded; furthermore, $f(\cdot)$ also satisfies Lipchitz's condition; that is, $|f(x) - f(y)| \leq c|x - y|$ for $c > 0$. If $h_n = cn^{-\beta}, 0 < \beta < 1$, then $\sqrt{nh_n}\widetilde{J}_n$ was asymptotically normal and $\sqrt{nh_n}\widetilde{J}_n \xrightarrow{\mathscr{L}} N(0, D^*(x))$, where $\widetilde{J}_n = (1/nh_n)\sum_{j=1}^n \psi(\varepsilon_j)K((X_j - x)/h_n)(\delta_j/\widehat{p}(X_j))$ and $D^*(x) = \psi_\varepsilon^2(x)(f(x)/p(x))\int_{-\infty}^{+\infty} K^2(u)du.$*

Proof. By using the first equality of Lemmas 7 and 8, we get

$$\widetilde{J}_n - \widehat{J}_n$$

$$= \frac{1}{nh_n}\sum_{j=1}^n \left[\frac{\delta_j}{\widehat{p}(X_j)}\psi(\varepsilon_j)K\left(\frac{X_j - x}{h_n}\right)\right.$$

$$\left. - \frac{\delta_j}{p(X_j)}\psi(\varepsilon_j)K\left(\frac{X_j - x}{h_n}\right)\right]$$

$$= \frac{1}{nh_n} \sum_{j=1}^{n} \left[\psi \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right) \frac{p \left(X_j \right) - \widehat{p} \left(X_j \right)}{p \left(X_j \right) \widehat{p} \left(X_j \right)} \right]$$

$$= \frac{1}{nh_n} \sum_{j=1}^{n} \left[\psi \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right) \frac{p \left(X_j \right) - \widehat{p} \left(X_j \right)}{p^2 \left(X_j \right)} \right]$$

$$\times \left\{ 1 + o_p \left(1 \right) \right\}$$

$$= -G_n \left\{ 1 + o_p \left(1 \right) \right\}. \tag{47}$$

It then follows from the third equality of Lemmas 7 and 6 that $\widetilde{J}_n = \widehat{J}_n + o_p(1)$ and consequently $\sqrt{nh_n}\widetilde{J}_n \xrightarrow{\mathscr{L}} N(0, D^*(x))$. This completes the proof of this lemma. □

In what follows we present the proofs of Theorems 1 and 2, respectively.

Proof of Theorem 1. The proof can be proved by the following two cases.

(i) If either $\widehat{m}_M(x) = \widehat{m}_{Mcc}(x)$ or $\widehat{m}_M(x) = \widehat{m}_{Mcp}(x)$, the proof of this theorem is similar to the proof of Theorem 1 in [15].

(ii) If $\widehat{m}_M(x) = \widehat{m}_M^*(x)$, similar to the proof of Theorem 1 in [15], the proof of this theorem is obtained immediately by the first equality of Lemmas 7 and 8. It is easy to see from (i) and (ii) that we complete the proof of this theorem. □

Proof of Theorem 2. We can prove the conclusion of this theorem by the following three cases.

(1) If $\widehat{m}_M(x) = \widehat{m}_{Mcc}(x)$, let

$$R \left(X_j \right) = m \left(X_j \right) - m \left(x \right) - m' \left(x \right) \left(X_j - x \right),$$

$$\widehat{\eta}_j = R \left(X_j \right) - \left[\widehat{m}_{Mcc} \left(x \right) - m \left(x \right) \right]$$
$$- \left[\widehat{m}'_{Mcc} \left(x \right) - m' \left(x \right) \right] \left(X_j - x \right) \tag{48}$$
$$= m \left(X_j \right) - \widehat{m}_{Mcc} \left(x \right) - \widehat{m}'_{Mcc} \left(x \right) \left(X_j - x \right),$$

where $\widehat{m}'_{Mcc}(\cdot)$ is the estimate of $m'(\cdot)$. Then

$$\varepsilon_j + \widehat{\eta}_j = Y_j - \widehat{m}_{Mcc} \left(x \right) - \widehat{m}'_{Mcc} \left(x \right) \left(X_j - x \right). \tag{49}$$

Using (6), we get

$$\left\{ \psi \left(\varepsilon_j \right) + \psi' \left(\varepsilon_j \right) \widehat{\eta}_j + \left[\psi \left(\varepsilon_j + \widehat{\eta}_j \right) - \psi \left(\varepsilon_j \right) - \psi' \left(\varepsilon_j \right) \widehat{\eta}_j \right] \right\}$$
$$\times \delta_j K \left(\frac{X_j - x}{h_n} \right) = 0. \tag{50}$$

Again,

$$\sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \widehat{\eta}_j \delta_j K \left(\frac{X_j - x}{h_n} \right)$$

$$= \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) R \left(X_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right)$$

$$- \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right) \left(\widehat{m}_{Mcc} \left(x \right) - m \left(x \right) \right)$$

$$- \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right) \left(\widehat{m}'_{Mcc} \left(x \right) - m' \left(x \right) \right) \left(X_j - x \right)$$

$$= I_1 - I_2 - I_3, \tag{51}$$

where

$$I_1 = \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) R \left(X_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right),$$

$$I_2 = \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right) \left(\widehat{m}_{Mcc} \left(x \right) - m \left(x \right) \right), \tag{52}$$

$$I_3 = \sum_{j=1}^{n} \psi' \left(\varepsilon_j \right) \delta_j K \left(\frac{X_j - x}{h_n} \right)$$
$$\times \left(\widehat{m}'_{Mcc} \left(x \right) - m' \left(x \right) \right) \left(X_j - x \right).$$

By Lemma 3, we have that

$$I_1 = \frac{nh_n^3}{2} p \left(x \right) f \left(x \right) \psi_\varepsilon \left(x \right) m'' \left(x \right) s_2 \left(1 + o_p \left(1 \right) \right),$$

$$I_2 = nh_n p \left(x \right) f \left(x \right) \psi_\varepsilon \left(x \right) s_0 \left(\widehat{m}_{Mcc} \left(x \right) - m \left(x \right) \right) \tag{53}$$
$$\times \left(1 + o_p \left(1 \right) \right),$$

$$I_3 = nh_n^2 p \left(x \right) f \left(x \right) \psi_\varepsilon \left(x \right) s_1 \left(\widehat{m}'_{Mcc} \left(x \right) - m' \left(x \right) \right) \tag{54}$$
$$\times \left(1 + o_p \left(1 \right) \right).$$

It follows from the consistency shown in (15) and condition (7) that

$$\sum_{j=1}^{n} \left[\psi \left(\varepsilon_j + \widehat{\eta}_j \right) - \psi \left(\varepsilon_j \right) - \psi' \left(\varepsilon_j \right) \widehat{\eta}_j \right] \delta_j K \left(\frac{X_j - x}{h_n} \right) \tag{55}$$
$$= o_p \left(I_2 \right).$$

Since $s_0 = 1$ and $s_1 = 0$, we have with (50) and (54) that

$$nh_n p \left(x \right) f \left(x \right) \psi_\varepsilon \left(x \right) \left(\widehat{m}_{Mcc} \left(x \right) - m \left(x \right) \right) \left(1 + o_p \left(1 \right) \right)$$

$$= \frac{nh_n^3}{2} p \left(x \right) f \left(x \right) \psi_\varepsilon \left(x \right) m'' \left(x \right) s_2 \left(1 + o_p \left(1 \right) \right) + J_n. \tag{56}$$

Therefore,

$$\widehat{m}_{Mcc}(x) - m(x) = \frac{h_n^2 m''(x)}{2} s_2 \left(1 + o_p(1)\right)$$

$$+ \frac{J_n}{nh_n \psi_\varepsilon(x) f(x) p(x)} \left(1 + o_p(1)\right)$$

$$= C_n + \frac{J_n}{nh_n \psi_\varepsilon(x) f(x) p(x)} \left(1 + o_p(1)\right). \tag{57}$$

By Lemma 5 and Slutsky's Theorem, we get (16); that is,

$$\sqrt{nh_n} \left(\widehat{m}_{Mcc}(x) - m(x) - C_n\right) \xrightarrow{\mathscr{L}} N(0, \pi(x)). \tag{58}$$

(2) If $\widehat{m}_M(x) = \widehat{m}_{Mcp}(x)$, similar to the proof case (1), the proof can be obtained immediately by Lemmas 4 and 6.

(3) The following will prove the case when $\widehat{m}_M(x) = \widehat{m}_M^*(x)$. Let

$$R(X_j) = m(X_j) - m(x) - m'(x)(X_j - x),$$

$$\widehat{\eta}_j^* = R(X_j) - [\widehat{m}_M^*(x) - m(x)]$$

$$- [\widehat{m}_M^{*\prime}(x) - m'(x)](X_j - x) \tag{59}$$

$$= m(X_j) - \widehat{m}_M^*(x) - \widehat{m}_M^{*\prime}(x)(X_j - x),$$

where $\widehat{m}_M^{*\prime}(\cdot)$ is the estimate of $m'(\cdot)$. Then we have

$$\varepsilon_j + \widehat{\eta}_j^* = Y_j - \widehat{m}_M^*(x) - \widehat{m}_M^{*\prime}(x)(X_j - x). \tag{60}$$

By (13), one can get

$$\left\{\psi(\varepsilon_j) + \psi'(\varepsilon_j)\widehat{\eta}_j^* + [\psi(\varepsilon_j + \widehat{\eta}_j^*) - \psi(\varepsilon_j) - \psi'(\varepsilon_j)\widehat{\eta}_j^*]\right\}$$

$$\times \frac{\delta_j}{\widehat{p}(X_j)} K\left(\frac{X_j - x}{h_n}\right) = 0. \tag{61}$$

It then follows from the first equality of Lemma 7, Lemma 8, and (61) that

$$\frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) \widehat{\eta}_j^* \frac{\delta_j}{\widehat{p}(X_j)} K\left(\frac{X_j - x}{h_n}\right)$$

$$= \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) R(X_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right)$$

$$- \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right) (\widehat{m}_M^*(x) - m(x))$$

$$- \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right)$$

$$\times (\widehat{m}_M^{*\prime}(x) - m'(x))(X_j - x) + o_p(1)$$

$$= I_1 - I_2 - I_3 + o_p(1), \tag{62}$$

where

$$I_1 = \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) R(X_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right),$$

$$I_2 = \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right)$$

$$\times (\widehat{m}_M^*(x) - m(x)), \tag{63}$$

$$I_3 = \frac{1}{nh_n} \sum_{j=1}^n \psi'(\varepsilon_j) \frac{\delta_j}{p(X_j)} K\left(\frac{X_j - x}{h_n}\right)$$

$$\times (\widehat{m}_M^{*\prime}(x) - m'(x))(X_j - x).$$

By Lemma 4, it is easy to see

$$I_1 = \frac{h_n^2}{2} f(x) \psi_\varepsilon(x) m''(x_0) s_2 \left(1 + o_p(1)\right), \tag{64}$$

$$I_2 = f(x) \psi_\varepsilon(x) s_0 (\widehat{m}_M^*(x) - m(x))\left(1 + o_p(1)\right), \tag{65}$$

$$I_3 = h_n f(x) \psi_\varepsilon(x) s_1 (\widehat{m}_M^{*\prime}(x) - m'(x))\left(1 + o_p(1)\right). \tag{66}$$

According to the first equality of Lemmas 7 and 8, the consistency shown in (15), and the condition (7), we have

$$\frac{1}{nh_n} \sum_{j=1}^n \left[\psi(\varepsilon_j + \widehat{\eta}_j^*) - \psi(\varepsilon_j) - \psi'(\varepsilon_j)\widehat{\eta}_j^*\right]$$

$$\times \frac{\delta_j}{\widehat{p}(X_j)} K\left(\frac{X_j - x}{h_n}\right) = o_p(I_2). \tag{67}$$

Since $s_0 = 1$ and $s_1 = 0$, we have with (61) and (65) that

$$f(x) \psi_\varepsilon(x) (\widehat{m}_M^*(x) - m(x))\left(1 + o_p(1)\right)$$

$$= \frac{h_n^2}{2} f(x) \psi_\varepsilon(x) m''(x) s_2 \left(1 + o_p(1)\right) + \widetilde{J}_n. \tag{68}$$

Therefore,

$$\widehat{m}_M^*(x) - m(x) = \frac{h_n^2 m''(x)}{2} s_2 \left(1 + o_p(1)\right)$$

$$+ \frac{\widetilde{J}_n}{\psi_\varepsilon(x) f(x) p(x)} \left(1 + o_p(1)\right)$$

$$= C_n + \frac{\widetilde{J}_n}{\psi_\varepsilon(x) f(x) p(x)} \left(1 + o_p(1)\right). \tag{69}$$

By Lemma 9 and Slutsky's Theorem, we get (16); that is,

$$\sqrt{nh_n} \left(\widehat{m}_M^*(x) - m(x) - C_n\right) \xrightarrow{\mathscr{L}} N(0, \pi(x)). \tag{70}$$

It is easy to see from cases (1), (2), and (3) that we complete the proof. □

Conflict of Interests

Acknowledgments

References

[1] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London, UK, 1996.

[2] P. J. Green and B. W. Silverman, *Kernel Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall, London, UK, 1994.

[3] D. M. Titterington and G. M. Mill, "Kernel-based density estimates from incomplete data," *Journal of the Royal Statistical Society B: Methodological*, vol. 45, no. 2, pp. 258–266, 1983.

[4] J. Fan, "Local linear regression smoothers and their minimax efficiencies," *The Annals of Statistics*, vol. 21, no. 1, pp. 196–216, 1993.

[5] T. Hastie and C. Loader, "Local regression: automatic kernel estimators of regression curves," *Annals of Statistics*, vol. 15, pp. 182–201, 1993.

[6] T. Orchard and M. A. Woodbury, "A missing information principle: theory and applications," in *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 3, pp. 697–715, University of California, June-July 1970.

[7] D. Ruppert and M. P. Wand, "Multivariate locally weighted least squares regression," *The Annals of Statistics*, vol. 22, no. 3, pp. 1346–1370, 1994.

[8] P. E. Cheng, "Nonparametric estimation of mean functionals with data missing at random," *Journal of the American Statistical Association*, vol. 89, no. 425, pp. 81–87, 1994.

[9] K. Hirano, G. W. Imbens, and G. Ridder, "Efficient estimation of average treatment effects using the estimated propensity score," *Econometrica*, vol. 71, no. 4, pp. 1161–1189, 2003.

[10] Q. Wang, O. Linton, and W. Härdle, "Semiparametric regression analysis with missing response at random," *Journal of the American Statistical Association*, vol. 99, no. 466, pp. 334–345, 2004.

[11] H. Liang, "Generalized partially linear models with missing covariates," *Journal of Multivariate Analysis*, vol. 99, no. 5, pp. 880–895, 2008.

[12] Q. Wang and Z. Sun, "Estimation in partially linear models with missing responses at random," *Journal of Multivariate Analysis*, vol. 98, no. 7, pp. 1470–1493, 2007.

[13] J. Fan and I. Gijbels, "Variable bandwidth and local linear regression smoothers," *The Annals of Statistics*, vol. 20, no. 4, pp. 2008–2036, 1992.

[14] R. J. Carroll, J. Fan, I. Gijbels, and M. P. Wand, "Generalized partially linear single-index models," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 477–489, 1997.

[15] J. Fan and J. Jiang, "Variable bandwidth and one-step local M-estimator," *Science in China A*, vol. 29, no. 1, pp. 688–702, 1999.