

Reweighted Nadaraya-Watson Estimator of the Regression Mean

Raid B. Salha¹ & Hazem I. El Shekh Ahmed²

¹Department of Mathematics, the Islamic University of Gaza, Palestine

²Department of Mathematics, Al Quds Open University, Palestine

Correspondence: Hazem I. El Shekh Ahmed, Department of Mathematics, Al Quds Open University, Palestine.

E-mail: hshaikhahmad@qou.edu

Received: August 12, 2014 Accepted: December 12, 2014 Online Published: January 26, 2015

doi:10.5539/ijsp.v4n1p138

URL: <http://dx.doi.org/10.5539/ijsp.v4n1p138>

Abstract

In this paper, the estimation of the regression mean using the Reweighted Nadaraya-Watson (RNW) estimator has been considered. The RNW is a modification of the Nadaraya-Watson (NW) estimator in order to obtain some more refinement estimator. We have considered some conditions under which the asymptotic normality of the proposed estimator has been derived. Then we generalized this result to the multivariate case by considering the estimation of the regression mean at distinct points.

Keywords: Nadaraya-Watson estimator, Reweighted Nadaraya-Watson estimator, regression mean, weighted coefficient, asymptotic normality.

1. Introduction

Theory and methodology for nonparametric regression is now well developed for the case of the estimation of the regression mean, to motivate the problem, consider a sequence of independent and identically distributed real random variables $\{(X_i, Y_i)\}_{i=1}^n$ with a joint pdf $f(x, y)$ as a bivariate random variable (X, Y) . The simple nonparametric regression function is written as

$$Y_i = m(X_i) + \varepsilon_i \quad (1.1)$$

where $X_i: i = 1, 2, \dots, n$ are called the predictors, $Y_i: i = 1, 2, \dots, n$ are the corresponding responses, $m(X_i)$ are the unknown regression mean functions to be estimated nonparametrically, and $\varepsilon_i: i = 1, 2, \dots, n$ denote the measurement errors, where $\varepsilon \sim N(0, \sigma^2)$.

The regression mean function $m(x)$ is the conditional mean, which is given by

$$m(x) = E(Y|x) = \int \frac{y f(x, y)}{g(x)} dy, \quad (1.2)$$

where $g(x)$ is the marginal density function of X . Using the kernel estimation, the regression mean function $m(x)$ is estimated by $\hat{m}(x)$, where

$$\hat{m}(x) = \int \frac{y \hat{f}(x, y)}{\hat{g}(x)} dy. \quad (1.3)$$

Conditional density estimation was introduced by Rosenblatt (1969). A bias correction was proposed by Hyndman et al. (1996). Fan et al. (1996) proposed a direct estimator based on a local polynomial estimation. One of the most widely used and studied estimators in the literature is the one proposed independently by Nadaraya (1964) and Watson (1964). Nadaraya-Watson kernel estimation is denoted by $\hat{f}_{NW}(Y|x)$ and defined as,

$$\hat{f}_{NW}(y|x) = \sum_{i=1}^n K_h(y - Y_i) w_i^{NW}(x) \quad (1.4)$$

where,

$$w_i^{NW}(x) = \frac{K_h(x - X_i)}{\sum_{i=1}^n K_h(x - X_i)}. \quad (1.5)$$

This gives the NW estimator $\hat{m}(X_i)$ of $m(X_i)$, where

$$\hat{m}(x) = \hat{E}(Y|X = x) = \frac{\sum_{i=1}^n Y_i K_h(x - X_i)}{\sum_{i=1}^n K_h(x - X_i)} \quad (1.6)$$

2. Reweighted Nadaraya-Watson Estimator

The large bias and boundary effects are considered to be the most important defects in the case of the Nadaraya-Watson estimator. The Nadaraya-Watson estimator was treated and modified in order to obtain some more refinement estimator, which is called the Reweighted Nadaraya-Watson (RNW) estimator, see Cai (2001), De Gooiger and Zerom (2003).

The (RNW) estimator is derived by a slight modification of the well-known Nadaraya-Watson estimator, to solve such disadvantages. The (RNW) conditional density estimator is defined by,

$$\hat{f}_{RNW}(y|x) = \sum_{i=1}^n K_h(y - Y_i) w_i^{RNW}(x), \quad (2.1)$$

where,

$$w_i^{RNW}(x) = \frac{\tau_i(x) K_h(x - X_i)}{\sum_{i=1}^n \tau_i(x) K_h(x - X_i)}, \quad (2.2)$$

where the probability weights $\tau_i(x)$ are considered with the following constraints,

$$\tau_i(x) \geq 0, \quad \sum_{i=1}^n \tau_i(x) = 1 \quad \text{and} \quad \sum_{i=1}^n \tau_i(x) (X_i - x) K_h(x - X_i) = 0. \quad (2.3)$$

These probability weights $\tau_i(x)$ will introduce a force to the Nadaraya-Watson weights $w_i^{NW}(x)$, and the next step is how to evaluate these weights $\tau_i(x)$ using the conditions (2.3).

The way to indicate the unique solution of $\tau_i(x)$ by maximizing the function $\sum_{i=1}^n \log(\tau_i(x))$ subjected to the previous constraints via Lagrange multipliers.

$$\text{Let } G = \sum_{i=1}^n \log(\tau_i(x)) + \kappa \left(1 - \sum_{i=1}^n \tau_i(x) \right) - n\lambda \sum_{i=1}^n \tau_i(x) (x_i - X) K_h(x_i - X).$$

Now differentiating the function G with respect to $\tau_i(x)$, we obtain

$$\frac{\partial G}{\partial \tau_i(x)} = \frac{1}{\tau_i(x)} - \kappa(n) - n\lambda \sum_{i=1}^n (x_i - X) K_h(x_i - X).$$

The solution can be derived directly by solving the equation $\frac{\partial G}{\partial \tau_i(x)} = 0$,

Salha and Shekh Ahmed (2009) have showed that

$$\tau_i(x) = \frac{1}{n(\lambda(x_i - X) K_h(x_i - X) + 1)}, \quad (2.4)$$

where λ can be derived as a unique minimizer of $L(\lambda)$, where $L(\lambda)$ is defined by

$$L(\lambda) = -\sum_{i=1}^n \log(1 + \lambda(X_i - x)K_h(x - X_i)) \quad (2.5)$$

The weighted coefficient $\tau_i(x)$ depends on the computed λ value and that makes (RNW) estimator is more available in the practice.

3. Main Results

In this section, we state some conditions under which we derive the main two results in this paper. The first result is stated in Theorem 1, where the asymptotic normality of the (RNW) estimator is shown. The second result is presented in Theorem 2, where the result of Theorem 1 has been generalized to the case of multivariate.

First, we consider the following notations.

$$\begin{aligned} \mu_j &= \int u^j K(u) du \\ \nu_j &= \int u^j K^2(u) du \\ \sigma^2(x) &= \text{Var}(Y_i | X_i = x) \end{aligned}$$

4. Conditions

Consider the following conditions:

C1. The kernel $K(\cdot)$ is a symmetric and bounded density with a bounded support $[-1, 1]$, and satisfies

$$\int u K(u) du = 0, \quad \int u^2 K(u) du < \infty.$$

C2. For fixed x , $g(x) > 0$, $g(\cdot)$ and $\sigma^2(\cdot)$ are continuous at x and $m(\cdot)$ has continuous second order derivatives in neighborhood of x .

C3. The conditional density function of Y given $X = x$ is bounded.

C4. $E(|Y_i|^\delta | X_i = U) \leq M < \infty$ for some $\delta > 2$, in a neighborhood of x .

C5. $\lim_{h \rightarrow 0} h = 0$, and $\lim_{n \rightarrow \infty} nh^3 = 0$.

Now, we state and prove the first main result in this paper.

Theorem 1. Under the conditions C1-C5, the following holds:

$$\sqrt{nh} \left[\hat{m}(x) - m(x) - \frac{1}{2} h^2 \mu_2 m''(x) + O_p(h^2) \right] \xrightarrow{D} N\left(0, g(x) \int K^2(u) du\right)$$

Proof.

$$\begin{aligned} [\hat{m}(x) - m(x)] &= \frac{\sum_{i=1}^n [Y_i - m(x)] w_i(x) K_h(x - X_i)}{\sum_{i=1}^n w_i(x) K_h(x - X_i)} \\ &= \frac{\sum_{i=1}^n [Y_i - m(X_i) + m(X_i) - m(x)] w_i(x) K_h(x - X_i)}{\sum_{i=1}^n w_i(x) K_h(x - X_i)} \\ &= \left\{ (nh)^{-\frac{1}{2}} J_1 + J_2 \right\} J_3^{-1} \{1 + o_p(1)\}, \end{aligned} \quad (3.1)$$

where

$$J_1 = \sqrt{\frac{h}{n}} \sum_{i=1}^n [Y_i - m(X_i)] w_i(x) K_h(x - X_i),$$

$$J_2 = \frac{1}{n} \sum_{i=1}^n [m(X_i) - m(x)] w_i(x) K_h(x - X_i),$$

and

$$J_3 = \frac{1}{n} \sum_{i=1}^n w_i(x) K_h(x - X_i).$$

Lemma 1. Under the conditions C1, C2 and C5, the following holds,

$$w_i(x) = n^{-1} b_i(x) (1 + o_p(1)),$$

where

$$b_i(x) = \left[1 + \frac{h \mu_2 g^{(1)}(x)}{g(x)} (X_i - x) K_h(x - X_i) K_h(x - X_i) \right]^{-1},$$

Proof. See De Gooiger and Zerom (2003).

Let $\varepsilon_i = Y_i - m(X_i)$

$$J_1 = \sqrt{\frac{h}{n}} \sum_{i=1}^n b_i(x) \varepsilon_i(x) K_h(x - X_i).$$

$$J_2 = \frac{1}{n} \sum_{i=1}^n [m(X_i) - m(x)] b_i(x) K_h(x - X_i). \quad (3.2)$$

$$J_3 = \frac{1}{n} \sum_{i=1}^n b_i(x) K_h(x - X_i).$$

Using Taylor expansion, we have

$$m(X_i) = m(x) + (X_i - x) m'(x) + \frac{(X_i - x)^2}{2!} m''(x) + o_p(h^2). \quad (3.3)$$

Lemma 2. Under the conditions C1-C5, the following hold

1. $J_2 = \frac{h^2}{2} \mu_2 m''(x) g(x) + O_p(h^2)$
2. $J_3 = g(x) + o_p(1).$

Proof.

i) By Equations (3.2) and (3.3), we have

$$J_2 = \frac{1}{2n} \sum_{i=1}^n m''(x) (X_i - x)^2 b_i(x) K_h(x - X_i) + O_p(h^2)$$

$$\begin{aligned}
&= \frac{1}{2} m''(x) \int (u-x)^2 K_h(x-u) g(u) du \\
&= \frac{h^2}{2} \mu_2 m''(x) g(x) + O_p(h).
\end{aligned}$$

ii) Similarly

$$\begin{aligned}
J_3 &= \frac{1}{n} \sum_{i=1}^n b_i(x) K_h(x-u) g(u) du = \int K_h(x-u) g(u) du + O_p(1) \\
&= g(x) \int K_h(u) du + O_p(1) = g(x) + O_p(1).
\end{aligned}$$

Now, $J_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Delta_i$, $\Delta_i = \sqrt{h} b_i(x) \varepsilon_i K_h(x - X_i)$.

Since, $E(\varepsilon_i | X_i) = 0$, $E(\Delta_i) = 0$, this implies that $E(J_1) = 0$.

Lemma 3. Under the conditions C1- C5, the following hold

$$J_1 \xrightarrow{D} N\left(0, \sigma^2(x) g(x) \int K^2(u) du\right).$$

Proof.

$$\begin{aligned}
Var(J_1) &= Var\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \Delta_i\right) = \frac{1}{n} \sum_{i=1}^n Var(\Delta_i) = Var(\Delta_i) = E(\Delta_i)^2 \\
&= h \iint b_i^2(x) (Y_i - m(X_i))^2 K_h^2(x - X_i) f(X_i, Y_i) dX_i dY_i \\
&= h \int Y_i^2 f(Y_i | X_i = x) dY_i \int K_h^2(x - X_i) g(X_i) dX_i + o_p(1) \\
&= E(Y_i^2 | X_i = x) \int K_h^2(u) g(x - hu) du + o_p(1) \\
&= Var(Y_i | X_i = x) g(x) \int K_h^2(u) du + o_p(1).
\end{aligned}$$

To show the asymptotic normality distribution of J_1 , we use Liapounov's theorem. It is sufficient to show that,

for $\delta > 0$,

$$\begin{aligned}
\rho_n &= \frac{\sum_{i=1}^n E|\Delta_i - E(\Delta_i)|^{2+\delta}}{\left[\sum_{i=1}^n Var(\Delta_i)\right]^{1+\frac{\delta}{2}}} \rightarrow 0, \text{ as } n \rightarrow \infty. \\
\rho_n &= \frac{\sum_{i=1}^n E|\Delta_i|^{2+\delta}}{\left[n \sum_{i=1}^n Var(\Delta_i)\right]^{1+\frac{\delta}{2}}} = \frac{(n^{-1})^{1+\frac{\delta}{2}} \sum_{i=1}^n E|\Delta_i|^{2+\delta}}{\left[\sum_{i=1}^n Var(\Delta_i)\right]^{1+\frac{\delta}{2}}} \rightarrow 0
\end{aligned}$$

Now,

$$\left(n^{-1}\right)^{1+\frac{\delta}{2}} \sum_{i=1}^n E\left|\Delta_i\right|^{2+\delta}=\left(n^{-1}\right)^{\frac{\delta}{2}} E\left|\Delta_i\right|^{2+\delta}=n^{-\frac{\delta}{2}} E\left|\Delta_i\right|^{2+\delta} \rightarrow 0, \text { by C4.}$$

Therefore, we have $\rho_n \rightarrow 0$, this completes the proof of the lemma. A combination of lemmas (1)-(3) and Equation (3.1) completes the proof of Theorem 1.

In the next Theorem, we will generalize Theorem 1 to the case of multivariate. Schuster (1972) has generalized the asymptotic normality of the Nadaraya-Watson estimator that was shown in Nadaraya (1965) by evaluating it at distinct points.

Theorem 2. Suppose x_1, x_2, \dots, x_k are distinct points and $g(x_i) > 0$, for $i = 1, 2, \dots, k$. Then under the conditions C1 to C5, we have

$$\sqrt{nh}(\hat{m}(x_1, x_2, \dots, x_k) - m(x_1, x_2, \dots, x_k))^T \rightarrow Z^*,$$

where Z^* has a multivariate normal distribution with mean vector $\mathbf{0}$ and a diagonal covariance matrix

$$C = [c_{ij}], \text{ where } c_{ii} = \frac{\sigma^2(x_i)}{g(x_i)} \int K^2(u) du, i = 1, 2, \dots, k.$$

Proof. For simplicity, we will prove the theorem in the case $k = 2$. Also, we will use the techniques of Schuster (1972).

For $i = 1, 2, \dots, n$ and $s = 1, 2$, define the followings:

$$U_{ni}^*(x_s) = \frac{1}{h} K\left(\frac{x_s - X_i}{h}\right) w_i(x_s), \quad U_{ni}(x_s) = \sqrt{h}(U_{ni}^*(x_s) - EU_{ni}^*(x_s)),$$

$$V_{ni}^*(x_s) = Y_i U_{ni}^*(x_s), \quad V_{ni}(x_s) = \sqrt{h}(V_{ni}^*(x_s) - EV_{ni}^*(x_s)),$$

$$U_n(x_s) = \sum_{i=1}^n U_{ni}(x_s), \quad V_n(x_s) = \sum_{i=1}^n V_{ni}(x_s)$$

$$W_{ni}(x_s) = (U_{ni}(x_1), V_{ni}(x_1), U_{ni}(x_2), V_{ni}(x_2)),$$

$$\sqrt{n}Z_n = (U_n(x_1), V_n(x_1), U_n(x_2), V_n(x_2))^T$$

$$A = \int K^2(u) du \begin{bmatrix} g(x_1) & W(x_1) & 0 & 0 \\ W(x_1) & V(x_1) & 0 & 0 \\ 0 & 0 & g(x_2) & W(x_2) \\ 0 & 0 & W(x_2) & V(x_2) \end{bmatrix},$$

where, $g(x) = \int f(x, y) dy$, $W(x) = \int y f(x, y) dy$ and $V(x) = \int y^2 f(x, y) dy$.

Let Z be a four-variate normal random variable with zero mean vector and covariance matrix A .

Lemma 4. Under the conditions C1 to C5, we have $Z_n \xrightarrow{D} Z$.

Proof.

Using Cràmer-Wold theorem, let $C = (c_1, d_1, c_2, d_2)$ in \mathbf{R}^4 . We want to show that $CZ_n^T \xrightarrow{D} CZ^T$.

The following hold for $s = 1, 2$ and $r = 1, 2, s \neq r$

- 1- $EU_{n_i}^2(x_s) = g(x_s) \int K^2(u) du + O(h).$
- 2- $EV_{n_i}^2(x_s) = V(x_s) \int K^2(u) du + O(h).$
- 3- $EU_{n_i}(x_s)V_{n_i}(x_s) = w(x_s) \int K^2(u) du + O(h).$
- 4- $EU_{n_i}(x_s)U_{n_i}(x_r) = O(h).$
- 5- $EV_{n_i}(x_s)V_{n_i}(x_r) = O(h).$
- 6- $EU_{n_i}(x_s)V_{n_i}(x_r) = O(h).$

We will prove (2) to illustrate the method (the proof of the remaining completes using the same techniques.

$$\begin{aligned}
 (2) \quad EV_{n_i}^2(x_s) &= h \left(EV_{n_i}^{*2}(x_s) - (EV_{n_i}^*(x_s))^2 \right) \\
 &= h \left[\iint \frac{Y_i^2}{h^2} K^2 \left(\frac{x_s - X_i}{h} \right) w_i(x_s) f(X_i, Y_i) dX_i dY_i \right. \\
 &\quad \left. - \left(\iint \frac{Y_i^2}{h} K \left(\frac{x_s - X_i}{h} \right) w_i(x_s) f(X_i, Y_i) dX_i dY_i \right)^2 \right] \\
 &= h \left[\iint \frac{Y_i^2}{h} K^2(u) w_i(x_s) f(x_s - hu, Y_i) du dY_i \right. \\
 &\quad \left. - \left(\iint Y_i K(u) w_i(x_s) f(x_s - hu, Y_i) du dY_i \right)^2 \right] \\
 &= \iint Y_i^2 K^2(u) w_i(x_s) [f(x_s, Y_i) - h u f'(x_s, Y_i) du dY_i] + O(h) \\
 &= \int Y_i^2 f(x_s, Y_i) dY_i \int K^2(u) du + O(h) \\
 &= V(x_s) \int K^2(u) du + O(h).
 \end{aligned}$$

The proof of Lemma 4 can be obtained by using the same techniques of the proof of Lemma 1 in Schuster (1972).

Now, let

$$Z_n^* = \sqrt{\frac{h}{n}} \left(\sum_{i=1}^n [U_{n_i}^*(x_i) - g(x_i)], \sum_{i=1}^n [V_{n_i}^*(x_1) - w(x_1)], \sum_{i=1}^n [U_{n_i}^*(x_2) - g(x_2)], \sum_{i=1}^n [V_{n_i}^*(x_2) - w(x_2)] \right)^T$$

Lemma 5. Under the condition of Lemma 4, the following holds $Z_n^* \xrightarrow{D} Z$.

Proof. Let $B_n = (g(x_1) - EU_{n_1}^*(x_1), w(x_1) - EV_{n_1}^*(x_1), g(x_2) - EU_{n_1}^*(x_2), w(x_2) - EV_{n_1}^*(x_2))^T$

For $s = 1, 2$

$$\begin{aligned} |EU_{n_1}^*(x_s) - g(x_s)| &= \left| \int \frac{1}{h} K\left(\frac{x_s - u}{h}\right) w_1(x_s) g(u) du - g(x_s) \right| \\ &\approx \left| \int K(u) g(x_s - hu) du - g(x_s) \right| \\ &= \left| \int K(u) \left[g(x_s) - h u g'(x_s) + \frac{h^2 u^2 g''(x_s)}{2} \right] du + O(h^2) - g(x_s) \right| \\ &= \left| g(x_s) + \frac{h^2 g''(x_s)}{2} \int u^2 K(u) du + O(h^2) - g(x_s) + O(h^2) \right| \\ &\leq \sup_x |g''(x)|^2 \frac{h^2}{2} \int u^2 K(u) du + O(h^2) - g(x_s) + O(h^2) = O(h^2) \\ |EU_{n_1}(x_s) - g(x_s)| &= O(h^2). \end{aligned}$$

Similarly,

$$\begin{aligned} |EV_{n_1}^*(x_s) - w(x_s)| &= \left| \int \int Y_i \frac{1}{h} K\left(\frac{x_s - u_i}{h}\right) w_i(x_s) f(X_i, Y_i) dX_i dY_i - w(x_s) \right| \\ &= \left| \int Y_i \int K(u) g(x_s - hu) f(x_s - hu, Y_i) du dY_i - w(x_s) \right| \\ &= \left| \int Y_i f(Y_i | x_s - hw) \int K(u) g(x_s - hu) du \right| = O(h^2). \end{aligned}$$

Therefore, $B_n = O(h^2)$. This implies that, $Z_n - Z_n^* = (nh_n^5)^{\frac{1}{2}} = o(1)$, using C5.

To complete the proof, define the function H from \mathbf{R}^4 and \mathbf{R}^2 by

$$H(y_1, y_2, y_3, y_4) = \left(\frac{y_2}{y_1}, \frac{y_4}{y_3} \right)^T.$$

Let $\theta = (g(x_1), w(x_1), g(x_2), w(x_2))$, and write $Z_n^* = (nh)^{\frac{1}{2}}(T_n - \theta)^T$, where

$$T_n = \frac{1}{n} \left(\sum_{i=1}^n U_{n_i}^*(x_1), \sum_{i=1}^n V_{n_i}^*(x_1), \sum_{i=1}^n U_{n_i}^*(x_2), \sum_{i=1}^n V_{n_i}^*(x_2) \right).$$

Now, the proof of the theorem completes using the Mann-Wald Theorem.

5. Application (Simulation Study)

In this section, the performance of the RNW kernel estimator in estimating the regression mean function is tested using two simulated data. The performance of the estimator has been tested using the mean squared error (MSE), which is defined by

$$MSE = \frac{SSE}{n},$$

and the correlation coefficients between the predicted values \hat{y} and the actual values y , $R_{y,\hat{y}}^2$, which is defined by

$$R_{y,\hat{y}}^2 = 1 - \frac{SSE}{SSTO},$$

where, $SSE = \sum (y_i - \hat{y})^2$ denotes the total sum of errors, $SSTO = \sum (y_i - \bar{y})^2$ denotes the total sum of squares and \bar{y} denotes the mean of actual values.

Also, a comparison between the RNW kernel estimator and the NW kernel estimator has been given. Two samples of size 400 are simulated from the following two models.

The first model is

$$y = \sin 2\pi(1-x^2) + xe,$$

where $e \sim N(0, 1)$ and $x \sim \text{uniform}[0, 1]$, and the second is

$$y = x \sin 2\pi x + e,$$

where $e \sim N(0, 0.1)$ and $x \sim \text{uniform}[0, 1]$.

Figure 1 and Figure 2 present two scatter plots of the simulated data, the perfect curve, the RNW estimator and the NW estimator for the first and second models, respectively. The results of the simulation studies are collected in Table 1.

Table 1. Results of the simulation studies.

Model	Estimator	$R_{y,\hat{y}}^2$	MSE
First	RNW	0.9676385	0.0122138
	NW	0.9573843	0.0160839
Second	RNW	0.9485958	0.0068545
	NW	0.9411828	0.0078424

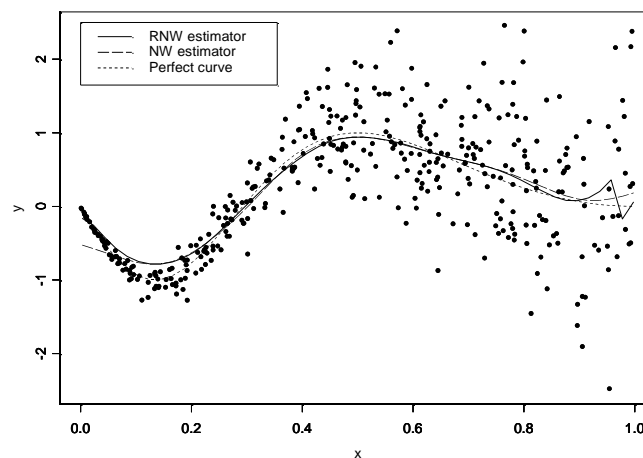


Figure 1. RNW and NW estimator for the first model

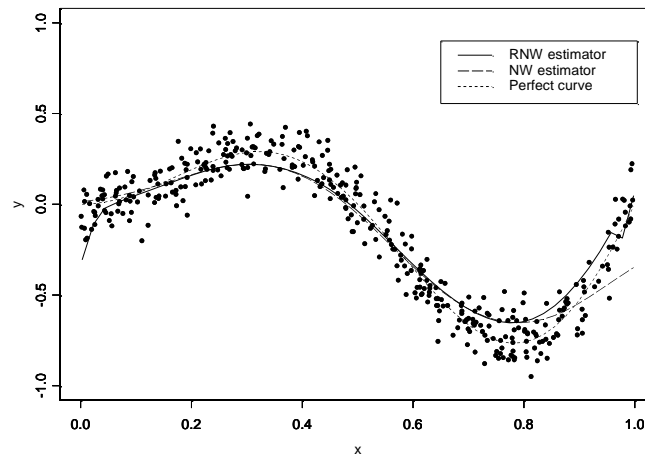


Figure 2. RNW and NW estimator for the second model

6. Conclusion

In this paper, we considered the RNW kernel estimator of the regression mean function. We derive the asymptotic normality of the regression mean function at different conditional points. Two applications using simulated data indicate that the performance of the RNW kernel estimator is reasonably good and it is better than the NW kernel estimator.

References

- Cai, Z. (2001). Weighted Nadaraya-Watson regression estimation. *Statistics and Probability Letters*, 51, 307-318.
- De Gooiger, J., & Zerom, D. (2003). On conditional Density Estimation. *Statistical Neerlandica*, 57, 159-176.
- Fan, J., Yao, Q., & Tong, H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika*, 83, 189-206.
- Hyndman, R. J., Bashtannyk, D. M., & Grunwald, G. K. (1996). Estimating and visualizing conditional densities. *J. Comput. Graph. Statist.*, 5, 315-336.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9, 141-142.
- Nadaraya, E. A. (1965). On nonparametric estimation of density functions and regression curves. *Theory of Probability and its Applications*, 10, 186-190.
- Rosenblatt, M. (1969). Conditional probability density and regression estimates in Multivariate Analysis II, Ed. P.R. Krishnaiah, pp. 25-31. New York: Academic Press.
- Salha, R., & Shekh Ahmed, H. (2009). On the Kernel Estimation of the Conditional Mode. *Asian Journal for Mathematics and Statistics*, 2(1), 1-8.
- Schuster, E. (1972). Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *The Annals of Mathematical Statistics*, 43(1), 84-88.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhya Ser. A.*, 26, 359-372.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).