Communications in Statistics - Simulation and Computation

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lssp20

Adaptive Weighted Nadaraya-Watson Estimation of the Conditional Quantiles by Varying Bandwidth

Hazem I. El Shekh Ahmed\textsuperscript{a}, Raid B. Salha\textsuperscript{b} & Hossam O. EL-Sayed\textsuperscript{c}
\textsuperscript{a} Department of Mathematics Al Quds Open University Palestine
\textsuperscript{b} Department of Mathematics The Islamic University of Gaza Palestine
\textsuperscript{c} Department of Education UNRWA Palestine
Accepted author version posted online: 24 Jun 2015.


To link to this article: http://dx.doi.org/10.1080/03610918.2015.1048880

Disclaimer: This is a version of an unedited manuscript that has been accepted for publication. As a service to authors and researchers we are providing this version of the accepted manuscript (AM). Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). During production and pre-press, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal relate to this version also.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
Adaptive Weighted Nadaraya-Watson Estimation of the Conditional Quantiles by Varying Bandwidth

Hazem I. El Shekh Ahmed  
Department of Mathematics  
Al Quds Open University  
Palestine  
hshaikhahmad@qou.edu

Raid B. Salha  
Department of Mathematics  
The Islamic University of Gaza  
Palestine  
rbsalha@mail.iugaza.edu

Hossam O. EL-Sayed  
Department of Education  
UNRWA  
Palestine

Abstract

In this paper, we define the adaptive Weighted Nadaraya-Watson estimation (AWNW) of the conditional distribution function (cdf) for independent and identically distributed (iid) data using varying bandwidth. The asymptotic normality of the proposed estimator is investigated. The results of the simulation studies show that the proposed estimation have better performance than the Weighted Nadaraya-Watson estimation with fixed bandwidth.

Keywords: Quantile regression, Weighted Nadaraya-Watson estimate, conditional distribution, kernel estimation, asymptotic normality.
1. Introduction

Suppose \(\{(X_i, Y_i)\}_{i=1}^{n}\) are \(\mathbb{R} \times \mathbb{R}\) random variables such that \(Y_i = m(X_i) + \epsilon_i, \ i = 1, 2, \ldots n,\) where \(m(X_i)\) is an unknown regression function. We assume that the response variable \(Y\) depends on an independent random variable \(X\) with common probability density function \(f\) and \(m(x)\) is the conditional mean curve, where

\[
m(x) = E(Y|X = x) = \int \frac{yf(x,y)}{f(x)} dy,
\]

and a nonparametric kernel estimation of the regression function can be obtained as

\[
\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right).
\]

Silverman [7] gave an algorithm with three steps for adaptive kernel estimation of density function. At the first, a prior kernel estimator with a fixed bandwidth is obtained. The second step, the local bandwidth factor \(\lambda_i\) is defined by,

\[
\lambda_i = \left\{ \frac{\hat{f}(X_i)}{g} \right\}^{-\alpha}, \quad 0 \leq \alpha \leq 1
\]

where \(g\) (assuming \(g \neq 0\)) is the geometric mean of \(\hat{f}(X_i)\). Finally, for one variable the conditional kernel estimation is given by

\[
\hat{f}_U(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h \lambda_i} K\left(\frac{x - X_i}{h \lambda_i}\right).
\]

The adaptive kernel estimation is equivalent to the kernel estimation with fixed bandwidth when \(\alpha\) is equal to 0. When \(\alpha = 1\), then the adaptive kernel estimation is equivalent to the nearest neighbor estimation.

by setting \( g = \sum_{i=1}^{n} \hat{f}(X_i)/n \).

The conditional distribution function has an important role for quantile regression. See Yu and Jones [9] and Hall et al [4].


Abberger [1] discussed the bias, variance and the mean square error of this estimator, and proved that the asymptotic normality of the proposed estimator.

Cai [2] defined the weight functions \( p_i(x) \) is of the data such that for \( 1 \leq i \leq n \),

\[
p_i(x) \geq 0, \quad \sum_{i=1}^{n} p_i(x) = 1, \quad \text{and} \quad \sum_{i=1}^{n} (X_i - x)p_i(x)K\left(\frac{x - X_i}{h}\right) = 0.
\]

He considered a Weighted Nadaraya-Watson estimation of the conditional distribution function as

\[
\hat{F}(y|x) = \frac{\sum_{i=1}^{n} p_i(x)K\left(\frac{x - X_i}{h}\right)I\{Y_i \leq y\}}{\sum_{i=1}^{n} p_i(x)K\left(\frac{x - X_i}{h}\right)}.
\]

The conditional \( \alpha \)- quantile is given by

\[
\hat{q}_\alpha(x) = \inf\{y \in R \mid \hat{F}(y|x) \geq \alpha\}, \quad 0 < \alpha < 1.
\]

From the previous, we have noticed a lot of statisticians use the kernel estimation of the density and conditional distribution function, but a question is raised: Is there any adjustments or improvements to this estimator who gives better results? In this paper, we will use the adaptive Weighted Nadaraya-Watson estimator of the conditional distribution function to
estimate the conditional quantiles.

We proposed the following estimator of the conditional distribution function as

\[
F_n(y|x) = \frac{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x) K(\frac{x - X_i}{h\lambda_i}) I_{(Y_i \leq y)}}{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x) K(\frac{x - X_i}{h\lambda_i})},
\]

where, we define the local bandwidth factor \( \lambda_i \) by

\[
\lambda_i = \left[ \frac{f_n(X_i)}{g} \right]^{-\alpha}, \quad 0 \leq \alpha \leq 1
\]

where \( g \) is the arithmetic mean of \( f_n(X_i) \).

Now, we define the adaptive of the conditional \( \alpha \)- quantile by

\[
q_{n,\alpha}(x) = \inf \{ y \in \mathbb{R} | F_n(y|x) \geq \alpha \}, \quad 0 < \alpha < 1.
\]

2. Preliminaries

In this section, we will consider assumptions on the kernel function, the bandwidth and the conditional distribution function. These assumptions are:

**(A1)** \( h = h_n \) is sequence of positive number satisfies the following:

(i) \( h \to 0, \quad \text{for} \quad n \to \infty; \)

(ii) \( nh \to \infty, \quad \text{for} \quad n \to \infty; \)

**(A2)** The kernel \( K \) is a Borel function and satisfies the following:

(i) \( K \) has compact support;

(ii) \( K \) is symmetric;
(iii) \( K \) is Lipschitz-continuous;

(iv) \( \int K(u)du = 1; \)

(v) \( K \) is bounded;

(A3) For fixed \( y \in \mathbb{R} \) there exists \( F''(y|x) = \frac{\partial^2 F(y|x)}{\partial x^2} \) in a neighborhood of \( x \).

(A4) For fixed \( y \in \mathbb{R} \), the marginal density \( g(\cdot) > 0 \), continuous at \( x \) and bounded

(A5) The probability weights \( p_i(x) \) is considered with the following properties:

(i) \( p_i(x) \geq 0 \),

(ii) \( \sum_{i=1}^{n} p_i(x) = 1 \),

(iii) \( \sum_{i=1}^{n} \frac{1}{h \lambda_i} (X_i - x)p_i(x)K(\frac{x - X_i}{h}) = 0. \)

Now, we state and prove some lemmas that will help us to achieve the main result in this paper.

**Lemma 1.** For any random variable \( X \) with density \( f \), we have

\[
E[f_n(x)] - f(x) = \frac{1}{2} h^2 \mu_2(x) f''(x) + o(h^2)
\]

**Proof:**

Assume that \( \int_{-\infty}^{\infty} K(u)du = 1 \), then we have \( \int_{-\infty}^{\infty} uK(u)du = 0. \)
And we assume that \( \mu_2(x) = \int_{-\infty}^{\infty} u^2 K(u) du < \infty. \)

\[
E[f_n(x)] = E\left[ \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{x-X_i}{h} \right) \right]
= \frac{1}{nh} \int_{-\infty}^{\infty} K\left( \frac{x-X_i}{h} \right) f(y) dy
\]

Let \( \frac{x-y}{h} = z, \) then \( y = x - hz, \) and so

\[
E[f_n(x)] = \frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{\infty} K(z) f(x - zh) dz
= \int_{-\infty}^{\infty} K(z) f(x - zh) dz
\]

Now, using Taylor’s expansion for \( f(x - zh) \) to get

\[
f(x - zh) = f(x) - zh f'(x) + \frac{(zh)^2}{2} f''(x) + o(h^2)
\]

Now,

\[
E[f_n(x)] = \int_{-\infty}^{\infty} K(z) \left[ f(x) - zh f'(x) + \frac{(zh)^2}{2} f''(x) + o(h^2) \right] dz
= \int_{-\infty}^{\infty} K(z) f(x) dz - h f'(x) \int_{-\infty}^{\infty} z K(z) dz + \frac{h^2}{2} f''(x) \int_{-\infty}^{\infty} z^2 K(z) dz
+ \int_{-\infty}^{\infty} o(h^2) K(z) dz
= f(x) + \frac{1}{2} h^2 \mu_2(x) f''(x) + o(h^2).
\]

This complete the proof of the lemma.

**Lemma 2.** (Integral approximation of the sum over the kernel function)

Using (A2)(i), Lipschitz-continuity, and the mean value theorem of integration, it follows:
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \frac{1}{nh_n} K(U_i) = \int_{-\infty}^{\infty} K(u)du
\]
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \frac{1}{nh_n} K^2(U_i) = \int_{-\infty}^{\infty} K^2(u)du
\]
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \frac{1}{nh_n} U_iK(U_i) = \int_{-\infty}^{\infty} uK(u)du
\]

**Proof**: See Abberger [1]

Now, the next lemma shows how the weight function \( p_i(x) \) can be calculated.

**Lemma 3.** Under the assumption \( A(1) \) and \( A(5) \), we have

\[
p_i(x) = n^{-1}\left\{ 1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left( \frac{X_i - x}{h\lambda_i} \right) \right\}^{-1}
\]

**Proof**:  

First, introduce the empirical log-likelihood function as

\[
L = \sum_{i=1}^{n} \log(p_i(x)).
\]

By maximizing \( L \) and using Lagrange multipliers, we have

\[
F(\tau_1, \tau_2, p_1(x), \ldots, p_n(x)) = \sum_{i=1}^{n} \log(p_i(x)) + \tau_1 (1 - \sum_{i=1}^{n} p_i(x)) + \tau_2 \sum_{i=1}^{n} \frac{1}{h\lambda_i} (X_i - x)p_i(x)K\left( \frac{X_i - x}{h\lambda_i} \right),
\]

where \( \tau_1, \tau_2 \) are the Lagrange multipliers.
Now differentiate with respect to \( p_i(x) \), to get

\[
\frac{\partial F}{\partial p_i(x)} = np_i^{-1}(x) - n\tau_1 + n \frac{\tau_2}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i}).
\]

Setting \( \frac{\partial F}{\partial p_i(x)} = 0 \), we have that

\[
\frac{\partial F}{\partial p_i(x)} = p_i^{-1}(x) - \tau_1 + \frac{\tau_2}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i}) = 0.
\]

This implies that,

\[
p_i(x) = \{\tau_1 - \frac{\tau_2}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}
\]

\[
= \tau_1^{-1}(1 - \frac{\tau_2}{\tau_1h\lambda_i} (X_i - xK(\frac{x - X_i}{h\lambda_i}))^{-1}
\]

choose \( \tau = -\frac{\tau_2}{\tau_1} \) to get,

\[
p_i(x) = \tau_1^{-1}\{1 + \frac{\tau}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}
\]

(1)

So,

\[
\tau_1 p_i(x) = \{1 + \frac{\tau}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}
\]

Take the sum over two sides, to get

\[
\tau_1 = \sum_{i=1}^{n} \{1 + \frac{\tau}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}
\]

(2)

This implies that

\[
p_i(x) = \frac{\{1 + \frac{\tau}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}}{\sum_{i=1}^{n} \{1 + \frac{\tau}{h\lambda_i} (X_i - x)K(\frac{x - X_i}{h\lambda_i})\}^{-1}}
\]
Thus
\[ p_i(x)\left(1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right) = \frac{1}{\sum_{i=1}^{n} \left[1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]^{-1}}. \]

Summing over \( i = 1, \ldots, n \) to get,
\[ \sum_{i=1}^{n} p_i(x) + \sum_{i=1}^{n} \frac{\tau}{h\lambda_i} (X_i - x)p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right) = \frac{n}{\sum_{i=1}^{n} \left[1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]^{-1}}. \]

That is,
\[ \sum_{i=1}^{n} \left[1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]^{-1} = n. \]

Thus,
\[ p_i(x) = n^{-1}\left[1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]^{-1}. \]

Cai [2] considered that, \( \tau \) is chosen to maximize
\[ L_n(\tau) = \frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{h\lambda_i} \left[1 + \frac{\tau}{h\lambda_i} (X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]. \]

**Lemma 4.** Under the assumption \( A(3) \) and \( A(4) \), we have

1. \( \tau = -\frac{h\lambda_i\mu_2g'(x)}{2\nu_2g(x)}[1 + o_p(1)], \)
2. \( p_i(x) = b_i(x)[1 + o_p(1)], \)

where
\[ b_i(x) = \left[1 - \frac{h\lambda_i\mu_2g'(x)}{2\nu_2g(x)}(X_i - x)K\left(\frac{x - X_i}{h\lambda_i}\right)\right]^{-1}, \]
and
\[ \mu_2 = \int_{-\infty}^{\infty} u^2 K(u) du \quad \text{and} \quad \nu_2 = \int_{-\infty}^{\infty} u^2 K^2(u) du. \]

**Proof:** See Cai [2] and Salha [6].

3. **Asymptotic normality**

In this section, we show the asymptotic normality of \((F_n(y|x) - F(y|x))\), and we consider the mean squared error of \(F_n(y|x)\)

**Theorem 1.** Suppose that \(A(3)\) and \(A(4)\) hold. Then as \(n \to \infty\), we have

\[
F_n(y|x) - F(y|x) = \frac{1}{2} (h\lambda)^2 \mu_2 F''(y|x) + o(h\lambda)^2 + O((nh\lambda)^{-\frac{1}{2}}) \tag{3}
\]

and

\[
\sqrt{nh\lambda}[F_n(y|x) - F(y|x) - B(y|x) + o(h\lambda)^2] \to N(0, \sigma^2(y|x)) \tag{4}
\]

where \(B(y|x) = \frac{1}{2} (h\lambda)^2 \mu_2 F''(y|x), \sigma^2(y|x) = \nu_0 F(y|x)[1 - F(y|x)] / g(x)\)

and \(\nu_0 = \int_{-\infty}^{\infty} K^2(u) du. \)

**Proof:**
Since

\[
F_n(y|x) - F(y|x) = \sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right) I_{[Y_i \leq y]} - F(y|x)
\]

\[
= \frac{1}{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right)} \sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right) I_{[Y_i \leq y]} - F(y|x)
\]

\[
= \frac{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right) \left( I_{[Y_i \leq y]} - F(y|X_i) \right) + (F(y|X_i) - F(y|x))}{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right)}
\]

\[
= J_3^{-1} \left[ (nh\lambda_i)^{-\frac{1}{2}} J_1 + J_2 \right] \left[ 1 + o(1) \right], \quad (5)
\]

where

\[
J_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sqrt{h\lambda_i}} b_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right) I_{[Y_i \leq y]} - F(y|X_i)
\]

\[
J_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h\lambda_i} \left[ F(y|X_i) - F(y|x) \right] b_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right)
\]

\[
J_3 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h\lambda_i} b_i(x)K\left(\frac{x - X_i}{h\lambda_i}\right).
\]

Now the proof of the theorem will be given via the following lemmas.

**Lemma 5.**

\[
Var(J_1) = v_0 F(y|x) [1 - F(y|x)] g(x)
\]

**Proof:**
We note that
\[ \text{Var}(J_1) = E(J_1^2) - [E(J_1)]^2. \]

Since
\[
E[ I_{\{Y_i \leq y\}} - F(y|X_i) ] = \int_{-\infty}^{\infty} [ I_{\{w \leq w\}} - F(y|X_i) ] f(w|X_i)dw
\]
\[ = \int_{-\infty}^{\infty} I_{\{w \leq w\}}dw - \int_{-\infty}^{\infty} F(y|X_i)f(w|X_i)dw = F(y|X_i) - F(y|X_i) = 0. \]

That is \( E(J_1) = 0. \)

Now,
\[
J_1^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h\lambda_i} b_i^2(x) K^2 \left( \frac{x - X_i}{h\lambda_i} \right) [ I_{\{Y_i \leq y\}} - F(y|X_i) ]^2
\]
\[ + \sum_{i<j} \frac{1}{h^2\lambda_i\lambda_j} b_i(x)b_j(x)K \left( \frac{x - X_i}{h\lambda_i} \right) K \left( \frac{x - X_j}{h\lambda_j} \right) [ I_{\{Y_i \leq y\}} - F(y|X_i) ] \]

Then,
\[
J_1^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h\lambda_i} b_i^2(x) K^2 \left( \frac{x - X_i}{h\lambda_i} \right) [ I_{\{Y_i \leq y\}} - F(y|X_i) ]^2
\]

Since \( \lim_{n \to \infty} b_i^2(x) = 1 \), then \( E(b_i^2(x)) = 1 + o(1) \).

Now,
\[
E[ I_{\{Y_i \leq y\}} - F(y|X_i) ]^2 = E[ \frac{1}{n} \sum_{i=1}^{n} b_i^2(x) K^2 \left( \frac{x - X_i}{h\lambda_i} \right) [ I_{\{Y_i \leq y\}} - F(y|X_i) ]^2
\]
\[ = \int_{-\infty}^{\infty} I_{\{w \leq w\}}dw - 2F(y|X_i) \int_{-\infty}^{\infty} I_{\{w \leq w\}}dw + F^2(y|X_i) \]
\[ = F(y|X_i) - 2F^2(y|X_i) + F^2(y|X_i) \]
This implies that
\[ E \left[ I_{\{Y_i \leq y\}} - F(y|X_i) \right]^2 = F(y|X_i)[1 - F(y|X_i)] \]

Now,
\[ E \left[ \frac{1}{h\lambda_i} K^2 \left( \frac{x - X_i}{h\lambda_i} \right) \right] = \frac{1}{h\lambda_i} \int_{-\infty}^{\infty} K^2 \left( \frac{x - y}{h\lambda_i} \right) g(y) dy \]

set \( u = \frac{x - y}{h\lambda_i} \), to get

\[ E \left[ \frac{1}{h\lambda_i} K^2 \left( \frac{x - X_i}{h\lambda_i} \right) \right] = \int_{-\infty}^{\infty} K^2(u)g(x - h\lambda_i)du \]

Using Taylor expansion to get

\[ E \left[ \frac{1}{h\lambda_i} K^2 \left( \frac{x - X_i}{h\lambda_i} \right) \right] = g(x) \int_{-\infty}^{\infty} K^2(u)du + o(1) = v_0 g(x) + o(1) \]

Thus, we have the result.

**Lemma 6.**

(1) \[ J_2 = g(x)B(y|x) + o(h\lambda_i)^2 \]

(2) \[ J_3 = g(x) + o(1) \]

**Proof:**

To prove (1), we use Taylor expansion for \( F(y|X_i) \) to get

\[ F(y|X_i) = F(y|x) + (X_i - x) \frac{\partial F(y|x)}{\partial x} + \frac{1}{2} (X_i - x)^2 \frac{\partial^2 F(y|x)}{\partial x^2} + o(h\lambda_i)^2 \]

That is,

\[ F(y|X_i) - F(y|x) = (X_i - x) \frac{\partial F(y|x)}{\partial x} + \frac{1}{2} (X_i - x)^2 \frac{\partial^2 F(y|x)}{\partial x^2} + o(h\lambda_i)^2 \]
So,

\[ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h \lambda_i} [F(y|X_i) - F(y|x)] b_i(x) K(\frac{x - X_i}{h \lambda_i}) \]

\[= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h \lambda_i} b_i(x) K(\frac{x - X_i}{h \lambda_i} | (X_i - x) \partial F(y|x) \partial x + \frac{1}{2} (X_i - x)^2 \frac{\partial^2 F(y|x)}{\partial x^2} + o(h \lambda_i)^2) \]

\[= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h \lambda_i} b_i(x) K(\frac{x - X_i}{h \lambda_i} | \frac{1}{2} (X_i - x)^2 \frac{\partial^2 F(y|x)}{\partial x^2} + o(h \lambda_i)^2) \]

Then,

\[ E(J_2) = \int_{-\infty}^{\infty} \frac{1}{2h \lambda_i} (y - x)^2 \frac{\partial^2 F(y|x)}{\partial x^2} K(\frac{x - y}{h \lambda_i}) g(y) dy + o(h \lambda_i)^2 \]

Let \( u = \frac{x - y}{h \lambda_i} \), to get

\[ E(J_2) = \int_{-\infty}^{\infty} \frac{1}{2} u^2 (h \lambda_i)^2 K(u) \frac{\partial^2 F(y|x)}{\partial x^2} g(x - uh \lambda_i) du + o(h \lambda_i)^2 \]

\[= g(x) \left[ \frac{1}{2} (h \lambda_i)^2 \frac{\partial^2 F(y|x)}{\partial x^2} \right] \int_{-\infty}^{\infty} u^2 K(u) du + o(h \lambda_i)^2 \]

Thus we have the result.

To prove (2),
\[ E(J_3) = E \left[ \sum_{i=1}^{n} \frac{1}{nh\lambda_i} b_i(x) K\left(\frac{x - X_i}{h\lambda_i}\right) \right] \]

\[ = \int_{-\infty}^{\infty} \frac{1}{h\lambda_i} b_i(x) K\left(\frac{x - y}{h\lambda_i}\right) g(y) dy + o(1) \]

\[ = \int_{-\infty}^{\infty} K(z) g(x - h\lambda_i) dz + o(1) \]

\[ = \int_{-\infty}^{\infty} K(z) g(x) dz + o(1) \]

\[ = g(x) + o(1) \]

From equation (5), we have

\[ [F_n(y|x) - F(y|x)] \approx (nh\lambda)^{-\frac{1}{2}} J_1 + g(x) B(y|x) \]

Therefore,

\[ \sqrt{nh\lambda}[F_n(y|x) - F(y|x) - B(y|x) + o(h\lambda)^2] = g^{-1}(x) J_1 + o(1) \] (6)

Since the mean square error (MSE) is used to measure the error when estimating the density function at a single point. It is defined by

\[ MS E[f_n(x)] = E[f_n(x) - f(x)]^2. \] (7)

That is, the last equation can be written as

\[ MS E[f_n(x)] = (Ef_n(x) - f(x))^2 + Var f_n(x). \]
So, from Theorem (1), the mean square error can be calculated as the next Corollary.

**Corollary 1.** Under the assumptions of Theorem 1, we have

\[
\text{MSE}[F_n(y|x)] = \frac{1}{2} (h\lambda_i)^2 F''(y|x) \int_{-\infty}^{\infty} u^2 K(u) du^2 + \int_{-\infty}^{\infty} K^2(u) du F(y|x)[1 - F(y|x)] / g(x)
\]

(8)

Now, the next Corollary shows the asymptotic normality \((nh\lambda_i)^{1/2} (F_n(y|x) - [F(y|x)])\).

**Corollary 2.** Under the assumptions of Theorem 1 and \(h \to 0\), we have

\[
\sqrt{nh\lambda_i} [F_n(y|x) - F(y|x)] \to N(0, \sigma^2(y|x))
\]

(9)

Now, the next theorem shows that the convergence of the conditional quantiles.

**Theorem 2.** Suppose that A(3) and A(4) hold. Then as \(n \to \infty\), we have

\[
q_{n,\alpha} \to q_{\alpha} \quad \text{in probability}
\]

**Proof:**

From equation (3), we have, for all \(x\) and \(y\)

\[
F_n(y|x) \to F(y|x) \quad \text{in probability.}
\]
Now, by Cai [2] Theorem 3, we have

$$\sup_{y \in \mathbb{R}} |F_n(y|x) - F(y|x)| \longrightarrow 0 \quad \text{in probability.}$$

Since \(q_{\alpha}(x)\) is unique, that is for any fixed \(x\), there is an \(\epsilon = \epsilon(x) > 0\) such that

$$\delta = \delta(\epsilon) = \min\{\alpha - F(q_{\alpha}(x) - \epsilon \mid x), F(q_{\alpha}(x) + \epsilon \mid x) - \alpha\} > 0.$$ 

Now,

$$P[|q_{n,\alpha}(x) - q_{\alpha}(x)| > \epsilon] \leq P[|F(q_{n,\alpha}(x)|x) - \alpha| > \delta]$$

$$\leq P\{\sup_{y \in \mathbb{R}} |F_n(y|x) - F(y|x)| > \delta\}$$

which tends to zero. Thus we have the result.

4. Simulation Studies

In this section, two simulation studies were conducted to compare the performance of the WNW estimation of the conditional distribution function and the AWNW estimation of the conditional distribution function.

In this studies, 400 data points are simulated from R program.

Case Study 1

We generated samples using R program of size 25, 50, 75, 100, 250, 400, 700 from the regression function \(y = \sin 2\pi(1 - x)^2 + ex\), where the \(x_i\) were drawn from a uniform distribution based on the interval \([0, 1]\), and the \(e\) have a normal distribution \(N(0, 1)\).

The fixed bandwidth \(h\) was computed by Silverman [6] and Wand, Jones [7] as

$$h_{opt} = 1.06 \text{ sd(t)} n^{-\frac{1}{5}}.$$
Now, we define a pilot estimate $f_n(X_i)$ that satisfies $f_n(X_i) \geq 0$, where $\{X_i\}$ is a family of iid.

In this paper we select the Epanechnikov kernel

$$K(t) = \frac{3}{4}(1 - \frac{1}{5}t^2)/\sqrt{5} \quad \text{for } |t| < \sqrt{5}.$$ 

We define the local bandwidth factor $\lambda_i$ by

$$\lambda_i = \left[ \frac{f_n(X_i)}{g} \right]^{-\alpha}, \quad 0 \leq \alpha \leq 1$$

where $g$ is the arithmetic mean of $f_n(X_i)$. We have considered that the equation

$$F_n(y|x) = \frac{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K(\frac{x - X_i}{h\lambda_i})I_{(Y_i \leq y)}}{\sum_{i=1}^{n} \frac{1}{h\lambda_i} p_i(x)K(\frac{x - X_i}{h\lambda_i})},$$

where $p_i(x)$ was computed as in Lemma 4.

The graph of the real regression function of the estimation of the regression function and the adaptive regression function.

The mean square error was computed as

From Table 1, for all sample sizes, we see that the AWNW estimators using varying bandwidths have smaller MSE values than the WNW estimator with fixed bandwidth. In each case, it is seen that MSE has the best performance.

**Case Study 2**

Another study was conducted to compare the performance of the weighted Nadaraya-Watson estimation of the conditional distribution function and the adaptive weighted Nadaraya-
Watson estimation of the conditional distribution function.

We generated samples using R program of size 25, 50, 75, 150, 250, 400, 700 from the regression function \( y = -0.03x^3 + 0.001x^2 + 1.25x + e \), where the \( x_i \) were drawn from a uniform distribution based on the interval \([0, 1]\), and the \( e \) have a normal distribution \( N(0, 1) \) and \( h, \lambda \) computed as the above simulation.

The mean square error was computed as

5. **Conclusion**

In this paper, we have studied the adaptive Weighted Nadaraya-Watson estimation of the conditional quantiles and we concluded the mean square error and asymptotically distribution of the proposed estimator.

The results of the simulation studies showed the adaptive weighted Nadaraya-Watson estimation of the conditional quantile with varying bandwidths provide better estimates than the conditional quantile estimate with fixed bandwidth which indicate that the AWNW is better than the WNW estimators.
References


Table 1: MSE and MSE Adaptive for samples

<table>
<thead>
<tr>
<th>n</th>
<th>MSE</th>
<th>MSE Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.06584803</td>
<td>0.05274891*</td>
</tr>
<tr>
<td>50</td>
<td>0.04262526</td>
<td>0.03490612*</td>
</tr>
<tr>
<td>75</td>
<td>0.04045895</td>
<td>0.02991416*</td>
</tr>
<tr>
<td>150</td>
<td>0.03757347</td>
<td>0.02849025*</td>
</tr>
<tr>
<td>250</td>
<td>0.0374387</td>
<td>0.02662148*</td>
</tr>
<tr>
<td>400</td>
<td>0.03647519</td>
<td>0.02602952*</td>
</tr>
<tr>
<td>700</td>
<td>0.03602435</td>
<td>0.0256541*</td>
</tr>
</tbody>
</table>

* Minimum MSE in each row.

Table 2: MSE and MSE Adaptive for samples

<table>
<thead>
<tr>
<th>n</th>
<th>MSE</th>
<th>MSE Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.04864595</td>
<td>0.04423345*</td>
</tr>
<tr>
<td>50</td>
<td>0.03177171</td>
<td>0.02536480*</td>
</tr>
<tr>
<td>75</td>
<td>0.03025748</td>
<td>0.02302679*</td>
</tr>
<tr>
<td>150</td>
<td>0.02876352</td>
<td>0.01976609*</td>
</tr>
<tr>
<td>250</td>
<td>0.02800730</td>
<td>0.0191666*</td>
</tr>
<tr>
<td>400</td>
<td>0.02797401</td>
<td>0.01846351*</td>
</tr>
<tr>
<td>700</td>
<td>0.0279603</td>
<td>0.01833646*</td>
</tr>
</tbody>
</table>

* Minimum MSE in each row.
Figure 1: shows the plot of the data together which the true regression curve, the Weighted Nadaraya-Watson estimator and the adaptive Weighted Nadaraya-Watson estimator (This graph at \(n = 75\)). The continuous line indicates that the true regression, the dotted line indicates that WNW estimator and the intermittent line indicates that the AWNW estimator.
Figure 2: shows the plot of the data together which the true regression curve, the WNW estimator and the AWNW estimator (This graph at n = 75). The continuous line indicates that the true regression, the dotted line indicates that WNW estimator and the intermittent line indicates that the AWNW estimator.