

دراسة مقارنة بين مقدر ندارايا- واتسن ومقدر ندارايا- واتسن المرجح لدالة متوسط الانحدار

A comparison Study Between the Nadaraya-Watson and the Reweighted Nadaraya-Watson Estimators of the Regression Mean Function

Raid B. Salha

Associate Prof. of Statistics, Department of Mathematics,
The Islamic University, Gaza, Palestine
rbsalha@iugaza.edu.ps

Received

7/6/2015

Accepted

25/8/2015

الملخص:

تستخدم مقدرات النواة اللامعلمية بكثرة في تقدير الانحدار. من أهم مقدرات النواة لدالة متوسط الانحدار هو مقدر ندارايا - واتسن. مقدر آخر يسمى مقدر ندارايا- واتسن المرجح قد تم اقتراحه لتحسين أداء مقدر ندارايا - واتسن. في هذا البحث نقارن بين المقدرين نظريا بمقارنة تقارب كل من تحيزهم، تباينهم و متوسط مربعات الخطأ. نتائج هذه المقارنة أشارت إلى أن تحيز مقدر ندارايا - واتسن المرجح كان الأفضل. بالإضافة لذلك فإنه تم مقارنة الأداء العملي للمقدين باستخدام بيانات محاكاة وأخرى حقيقية و النتائج كانت متسقة مع نتائج المقارنة النظرية وأشارت إلى أن مقدر ندارايا - واتسن المرجح قد عدل جوانب الضعف في مقدر ندارايا - واتسن.

Abstract:

Nonparametric kernel estimators are widely used in regression estimation. One of the most important kernel estimators of the regression mean function is the Nadaraya-Watson estimator. Another estimator, which is called the Reweighted Nadaraya-Watson estimator has been proposed to improve the performance of the Nadaraya-Watson estimator. In this paper, we have compared theoretically between the two estimators by looking at their asymptotic bias, variance and the mean squared error. The results of this comparison

indicated that the bias of the Reweighted Nadaraya-Watson estimator is better than that of the Nadaraya-Watson estimator.

Also, a comparison of the practical performance of the two estimators based on simulated and real data has been given. The results of this comparison was consistent with the results of the theoretical comparison and indicated that, the Reweighted Nadaraya-Watson estimator has modified the weakness of the Nadaraya-Watson estimator.

Keywords: Kernel estimation, regression mean function, asymptotic bias, asymptotic variance, mean squared error.

1. Introduction

Let $\{(X_i, Y_i)\}_{i=1}^n$ be a random sample which is distributed as a random variable (X, Y) with joint probability density function (pdf) $f(x, y)$ and X has a marginal pdf $g(x)$. Suppose the relationship between the dependent variable Y and the independent variable X can be written as,

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, \dots, n,$$

where $m(x)$ is the regression mean function, $m(x) = E(Y|X = x) = \int_{-\infty}^{\infty} y f(y|x) dy$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are independent random variables with $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2 < \infty$. If $m(x)$ is unknown, then its estimation can be done by estimating the conditional density function $f(y|x)$.

Several nonparametric methods have been proposed for estimating $f(y|x)$ based on a random sample $\{(X_i, Y_i)\}_{i=1}^n$. A class of kernel-type estimators is called the Nadaraya-Watson (NW) estimator which is one of the most widely known and used for estimating $f(y|x)$. The NW estimator has been created independently by Watson (1964) and Nadaraya (1964). The NW estimator of the conditional pdf, $f(y|x)$ is defined as follows.

Definition 1. The Nadaraya-Watson estimator, $\hat{f}_{NW}(y|x)$ of $f(y|x)$ is given by

$$\hat{f}_{NW}(y|x) = \frac{\sum_{i=1}^n K_h(x - X_i) K_h(y - Y_i)}{\sum_{i=1}^n K_h(x - X_i)},$$

On the Nadaraya-Watson Estimators of the Regression Mean Function

where K is a kernel function, $h = h_n$ is a sequence of positive numbers converges to zero and it is called the bandwidth and $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$.

The NW estimator of the regression mean function, $m(x)$, can be defined as

Definition 2. The Nadaraya-Watson estimator, $\hat{m}_{NW}(x)$, of $m(x)$ is given by

$$\hat{m}_{NW}(x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)}.$$

But the NW estimator suffers from some weakness when the estimation is done near the boundary points or the design density is highly clustered. For more details, see Hall et al. (1999), Cai (2001, 2002) and Yu and Stander (2003).

To overcome these weakness of the NW estimator a new class of the kernel estimation of the conditional density function and the regression mean function has been proposed by Hall et al. (1999). The new estimator is called the Rewieghted Nadaraya-Watson (RNW) estimator. It is defined as follows.

Definition 3. The Rewieghted Nadaraya-Watson estimator, $\hat{f}_{RNW}(y|x)$ of $f(y|x)$ is given by

$$\hat{f}_{RNW}(y|x) = \frac{\sum_{i=1}^n \tau_i(x) K_h(x - X_i) K_h(y - Y_i)}{\sum_{i=1}^n \tau_i(x) K_h(x - X_i)},$$

where $\tau_i(x)$ are probability weights functions satisfying the following properties

1. $\tau_i(x) \geq 0$,
2. $\sum_{i=1}^n \tau_i(x) = 1$,
3. $\sum_{i=1}^n \tau_i(x) (X_i - x) K_h(x - X_i) = 1$.

$$\tau_i(x) = \frac{1}{n(\lambda(X_i - x) K_h(x - X_i) + 1)}$$

where λ is the unique minimizer with respect to λ of $L(\lambda)$, where

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

Definition 4. The Reweighted Nadaraya-Watson estimator, $\hat{m}_{RNW}(x)$ of $m(x)$ is given by

$$\hat{m}_{RNW}(x) = \frac{\sum_{i=1}^n \tau_i(x) K_h(x - X_i) Y_i}{\sum_{i=1}^n \tau_i(x) K_h(x - X_i)}.$$

Recently, Salha and EL Shekh Ahmed (2015) studied the RNW and proved the joint asymptotic normality of it estimated at a finite number of conditional points.

Although there are many papers studied the two estimators, until now there is no comparison study between the two estimators, which covers the theoretical and applicable performance of the two estimators. Therefore, in this paper, we compare between the two estimators theoretically by considering the bias, variance and the mean squared error (MSE), then we compare between them by applications using simulated and real data. From these comparisons, we study where and at what cases the performance of the two estimators is different and if the results from the applications are consistent with the theoretical results.

The remaining of this paper consists of three sections. Section 2, contains a brief theoretical comparison between the estimators $\hat{m}_{NW}(x)$ and $\hat{m}_{RNW}(x)$. A practical comparison between the two estimators is given in Section 3. In Section 4, we discuss the results of the comparisons and the simulation studies and we give some conclusions.

2. A theoretical comparison

In this section, we make a brief theoretical comparison of the NW estimators, $\hat{f}_{NW}(y|x)$, $\hat{m}_{NW}(x)$ and the RNW estimator, $\hat{f}_{RNW}(y|x)$, $\hat{m}_{RNW}(x)$ by looking at their asymptotic bias, variance and the MSEs as the sample size $n \rightarrow \infty$, $h = h_n \rightarrow 0$, and $nh \rightarrow \infty$.

The asymptotic biases, variances and MSE of $\hat{f}_{NW}(y|x)$ and $\hat{f}_{RNW}(y|x)$ are given in Table 1. The results are taken from De Gooijer and Zerom (2003).

Table 1. A theoretical comparison between $\hat{f}_{NW}(y|x)$ and $\hat{f}_{RNW}(y|x)$.

	$\hat{f}_{NW}(y x)$	$\hat{f}_{RNW}(y x)$
Bias	$\frac{h^2}{2} Q(x) \mu_2(K)$	$\frac{h^2}{2} \mu_2(K) [f^{(2,0)}(y x) + f^{(0,2)}(y x)]$
Variance	$\frac{R(K) f(y x)}{n h^2 g(x)}$	$\frac{R(K) f(y x)}{n h^2 g(x)}$
MSE	$\frac{R(K) f(y x)}{n h^2 g(x)} + \frac{h^4}{4} Q^2(x) \mu_2^2(K)$	$\frac{R(K) f(y x)}{n h^2 g(x)} + \frac{h^4}{4} \mu_2^2(K) [f^{(2,0)}(y x) + f^{(0,2)}(y x)]^2$

where,

$$Q(x) = \left[f^{(2,0)}(y|x) + f^{(0,2)}(y|x) + 2 \frac{g'(x)}{g(x)} f^{(1,0)}(y|x) \right],$$

$$\mu_2(K) = \int u^2 K(u) du \text{ and } R(K) = \int K^2(u) du.$$

From Table 1, we have the following remarks:

1. The two variances are the same. Therefore, the difference in the MSEs between the two estimators depends only on their respective biases
2. The bias of $\hat{f}_{NW}(y|x)$ is large if either $\left| \frac{g'(x)}{g(x)} \right|$ or $|f^{(1,0)}(y|x)|$ are large.
3. The term $\left| \frac{g'(x)}{g(x)} \right|$ becomes large when the design density $g(x)$ is highly clustered.
4. When the design density is uniform then the term $\left| \frac{g'(x)}{g(x)} \right|$ vanishes which implies that the two biases are the same.
5. Since $\hat{f}_{RNW}(y|x)$ does not depend on the design density $g(x)$, it is called a design adaptive.
6. If $f(y|x)$ depends on x only through a location parameter, for example the regression mean $m(x)$, therefore $f(y|x) = f(y - m(x)|x)$ which implies that

$$f^{(1,0)}(y|x) = m'(x)f^{(1,0)}(y - m(x)|x).$$

This means $|f^{(1,0)}(y|x)|$ is large whenever $|m'(x)|$ is large and vanishes when $m'(x) = 0$ and this happens if $m(x)$ is flat or has maximum, minimum or inflection point at x .

This implies that the bias of $\hat{f}_{NW}(y|x)$ is larger than that of $\hat{f}_{RNW}(y|x)$ at the points x where $|m'(x)|$ is large and the two biases are the same at x , where $m(x)$ is flat or has maximum, minimum or inflection points. From the theoretical comparison, we note that

1. $\hat{f}_{RNW}(y|x)$ is better than $\hat{f}_{NW}(y|x)$ because of its bias and design adaptation.
2. $\hat{f}_{RNW}(y|x)$ does not suffer from boundary effect and therefore does not require modifications at the boundary.

The asymptotic biases, variances and MSE of $\hat{m}_{NW}(x)$ and $\hat{m}_{RNW}(x)$ are given in Table 2. Results for $\hat{m}_{NW}(x)$ are taken from Fan (1992) and for $\hat{m}_{RNW}(x)$ from Cai (2001). From Table 2, we have the following remarks:

1. The two estimators have the same variance and different biases.
2. The bias of $\hat{m}_{NW}(x)$ is large if either $|m''(x)|$ or $|\frac{g'(x)}{g(x)}|$ are large and it vanishes when $m''(x)$ and $g'(x)$ or $m'(x) = 0$.
3. $\hat{m}_{NW}(x)$ is not suitable to be used when the design density $g(x)$ is highly clustered where the term $|\frac{g'(x)}{g(x)}|$ becomes large.
4. $\hat{m}_{RNW}(x)$ is called a design adaptive since it does not depend on the density $g(x)$ of X .
5. The bias of $\hat{m}_{RNW}(x)$ is large if $|m''(x)|$ is large and vanishes when $m''(x) = 0$.
6. The variances of the two estimators get small as the sample size increases.

One special case where the biases are the same is at the points where $m'(x) = 0$ and this happens on flat portions, minima, maxima, points of inflection. Another case, if $g'(x) = 0$.

In addition if $m''(x) = 0$, or $|g'(x)m'(x)| > 0$ then $MSE(\hat{m}_{RNW}(x)) \leq MSE(\hat{m}_{NW}(x))$.

Table 2. A theoretical comparison between $\hat{m}_{NW}(x)$ and $\hat{m}_{RNW}(x)$.

	$\hat{m}_{NW}(x)$	$\hat{m}_{RNW}(x)$
Bias	$\frac{h^2}{2} A(x) \mu_2(K)$	$\frac{h^2}{2} \mu_2(K) m''(x)$
Variance	$\frac{R(K) \sigma^2(x)}{n h g(x)}$	$\frac{R(K) \sigma^2(x)}{n h g(x)}$
MSE	$\frac{R(K) \sigma^2(x)}{n h g(x)} + \frac{h^4}{4} A^2(x) \mu_2^2(K)$	$\frac{R(K) \sigma^2(x)}{n h g(x)} + \frac{h^4}{4} \mu_2^2(K) (m''(x))^2$

where, $A(x) = \left[m''(x) + \frac{2 g'(x) m'(x)}{g(x)} \right]$, $\sigma^2(x) = Var(y|x)$.

3. Practical comparison

In this section, we compared between $\hat{m}_{RNW}(x)$ and $\hat{m}_{NW}(x)$ through applications using simulated and real data. For each estimator, we computed the MSE and the correlation coefficients, $R_{y,\hat{y}}^2$, where

$$\begin{aligned} \text{MSE} &= \frac{SSE}{n} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \quad R_{y,\hat{y}}^2 = 1 - \frac{SSE}{SSTO} \\ &= 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \end{aligned}$$

where y_i denotes the true value, \hat{y}_i denotes its predicted value and \bar{y} is the mean of the data.

The bandwidth is computed using the following equation from Silverman (1986)

$$h_n = 1.06 s n^{-0.2},$$

where n is the sample size and s is the standard deviation of the sample.

This section consists of two subsections. In the first, we used simulated data while in the second subsection, we used two well-known international real data.

3.1 Simulation Studies

In this subsection, we used three simulated data from three different models to compare between $\hat{m}_{NW}(x)$ and $\hat{m}_{RNW}(x)$ and to study

where and at what cases the performance of the two estimators is different.

1. Simulation study 1

Two samples of sizes 200 are simulated from the model

$$y = 1.5 x^2 + e, \quad e \sim N(0, 0.25).$$

We used two different design densities, $g(x)$. In the first sample, we used $g(x) = \text{Uniform}[0,1]$, while in the second one we used $g(x) = N(0,1)$.

Table 3 contains the results of the comparison. Figure 1 and Figure 2 show the scatter plot of the simulated data together with the perfect curve and its estimation using the RNW and the NW estimators for the first and second simulated data.

From Table 3, we see that the MSE of $\hat{m}_{RNW}(x)$ is less than that of $\hat{m}_{NW}(x)$ and the correlation coefficient of $\hat{m}_{RNW}(x)$ is greater than that of $\hat{m}_{NW}(x)$. Also, from Figure 1 and Figure 2, we see that the two estimators has almost the same performance in the interior region around $x = 0$, where the value of $m'(x)$ is closed to zero. At the boundaries, we see the RNW is very closed to the perfect curve and its performance is much better than that of the NW. This results from the simulated data are consistent with the theoretical results from the second section.

Table 3. MSE and $R^2_{y,\hat{y}}$ for $\hat{m}_{RNW}(x)$ and $\hat{m}_{NW}(x)$ (Simulation study 1).

	Design pdf, $g(x)$		$\hat{m}_{RNW}(x)$	$\hat{m}_{NW}(x)$
First simulated data	Uniform [-1, 1]	MSE	0.008175	0.025389
		$R^2_{y,\hat{y}}$	0.964012	0.888236
First simulated data	N (0, 1)	MSE	0.032320	0.666388
		$R^2_{y,\hat{y}}$	0.997233	0.942960

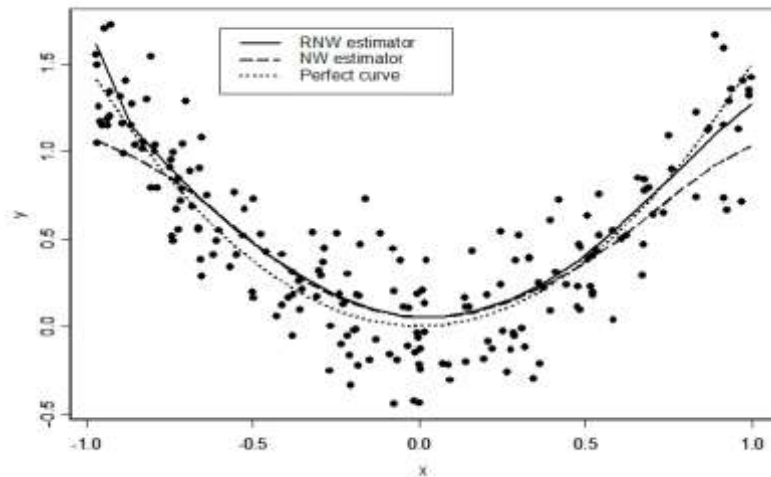


Figure 1. A scatter plot of the first simulated data (Simulation study 1) together with the perfect curve, the RNW and the NW estimator.

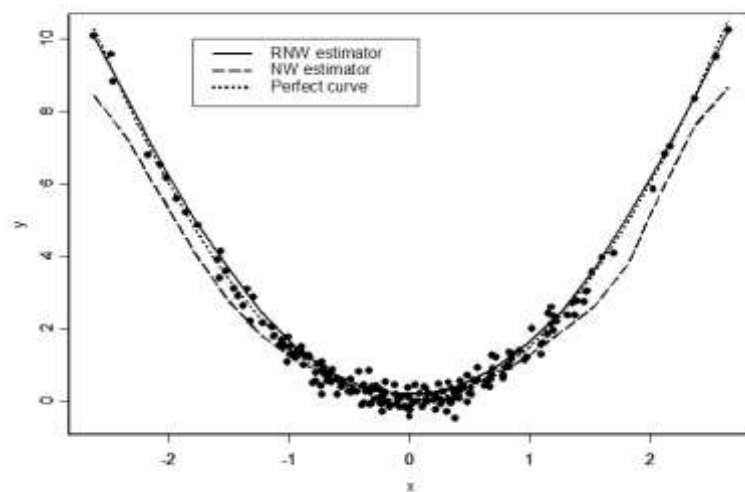


Figure 2. A scatter plot of the second simulated data (Simulation study 1) together with the perfect curve, the RNW and the NW estimator.

2. Simulation study 2

Two samples of sizes 200 are simulated from the model

$$y = 1 + 1.5x + e, \quad e \sim N(0, 0.25).$$

In the first sample we used $x \sim \text{Uniform}[-1, 1]$, while in the second one we used $x \sim N(0, 1)$.

Table 4 contains the results of the comparison. Figure 3 and Figure 4 show the scatter plot together with the perfect curve and its estimation using the RNW and the NW estimators for the first and second simulated data.

From Table 4, we see that the MSE of $\hat{m}_{RNW}(x)$ is less than that of $\hat{m}_{NW}(x)$ and on the other hand the correlation coefficient of $\hat{m}_{RNW}(x)$ is greater than that of $\hat{m}_{NW}(x)$.

Also, from Figure 3 and Figure 4, we see that the two estimators has almost the same performance in the interior region. At the boundaries, we see the RNW is very closed to the perfect curve and its performance is much better than that of the NW.

Table 4. MSE and $R^2_{y,\hat{y}}$ for $\hat{m}_{RNW}(x)$ and $\hat{m}_{NW}(x)$ (Simulation study 2).

	Design pdf $g(x)$		$\hat{m}_{RNW}(x)$	$\hat{m}_{NW}(x)$
First simulated data	Uniform [-1, 1]	MSE	0.004722	0.013994
		$R^2_{y,\hat{y}}$	0.994107	0.982536
First simulated data	N (0, 1)	MSE	0.010857	0.081240
		$R^2_{y,\hat{y}}$	0.998113	0.985878

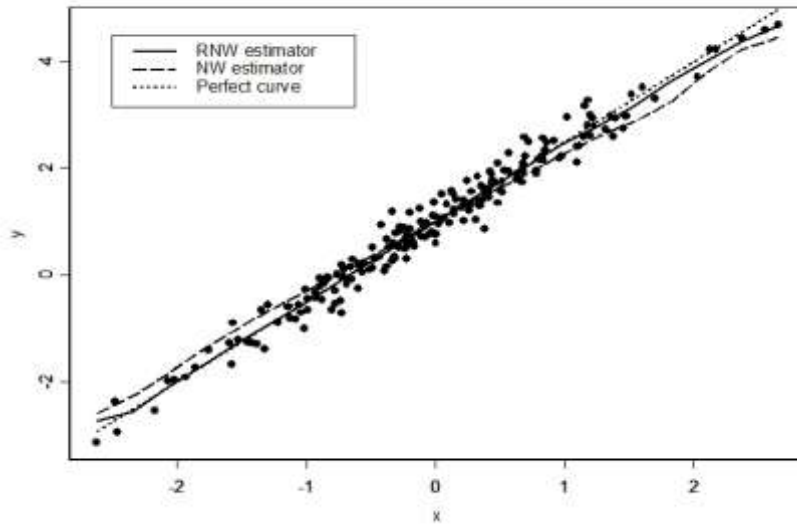


Figure 3. A scatter plot of the first simulated data (Simulation study 2) together with the perfect curve, the RNW and the NW estimators.

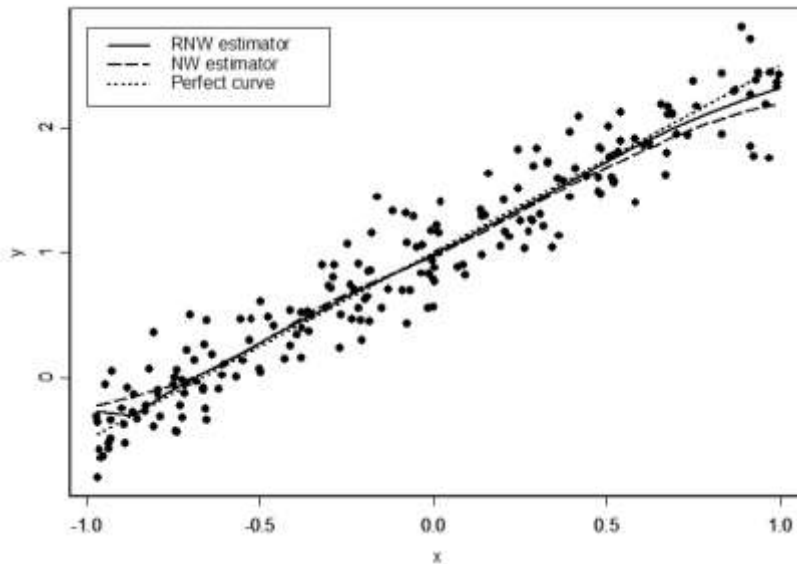


Figure 4. A scatter plot of the second simulated data (Simulation study 2) together with the perfect curve, the RNW and the NW estimators.

3. Simulation study 3

A samples of size 200 is simulated from the model

$$y = 2x^{\frac{1}{2}} - \sin(2\pi(1-x)^2) + e, \\ x \sim \text{Uniform}[0,1], \quad e \sim N(0, 0.25).$$

The results are listed in Table 5. Figure 5 shows the simulated data together with the perfect curve and its estimation using the RNW and the NW estimators. From Table 5, we see the performance of the two estimators is almost the same. The RNW is slightly better than the NW. Also from Figure 5, we see that the two estimators have the same performance at the points where $m(x)$ has minima, maxima, or inflection points.

Table 5. MSE and $R^2_{y,\hat{y}}$ for $\hat{m}_{RNW}(x)$ and $\hat{m}_{NW}(x)$ (Simulation study 3).

	$\hat{m}_{RNW}(x)$	$\hat{m}_{NW}(x)$
MSE	0.031169	0.032204
$R^2_{y,\hat{y}}$	0.915955	0.913162

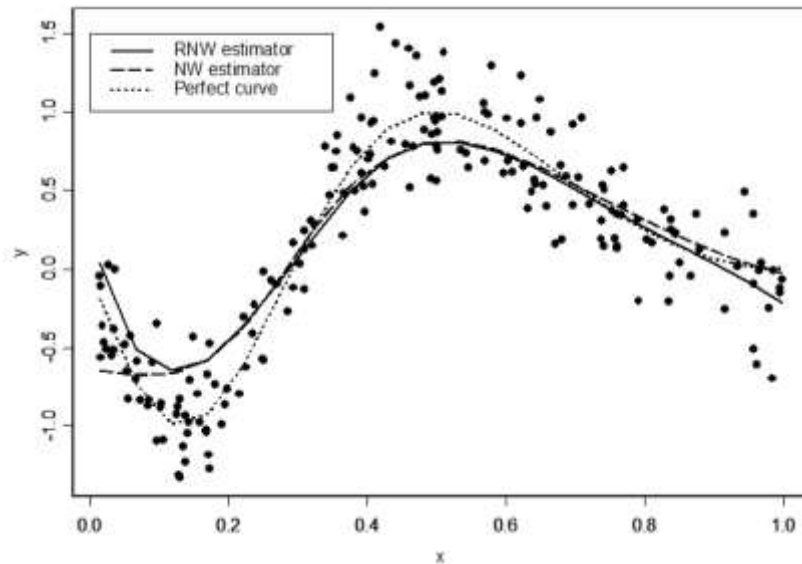


Figure 5. A scatter plot of the simulated data (Simulation study 3) together with the perfect curve, the RNW and the NW estimator.

3.2 Real data

In this section, we compared between the RNW and NW estimators of the regression mean function through applications using two international real data.

1. Airline data

In the first application using real data, we consider the international airline passenger data (Monthly totals in thousands from January 1949 to December 1960) of Box et al. (1994). Figure 6 shows the time plot of the data. It is clear from the graph that the variability increases with the level of the observations. We have transformed the data by taking the logarithm, then we used the first 137 observations to predict the last 6 observations using the RNW, $\hat{m}_{RNW}(x)$ and the NW, $\hat{m}_{NW}(x)$ estimators. We have used a bandwidth, $h = 0.2525712..$ The MSE of the predicted values using the two estimators are as follows:

$$\text{MSE}(\hat{m}_{RNW}(x)) = 0.01268932, \text{MSE}(\hat{m}_{NW}(x)) = 0.05887032.$$

Figure 7 shows the tail of the transformed data with the predicted values. Also, the results has been listed in Table 6. The results indicate that the RNW is better than the NW.

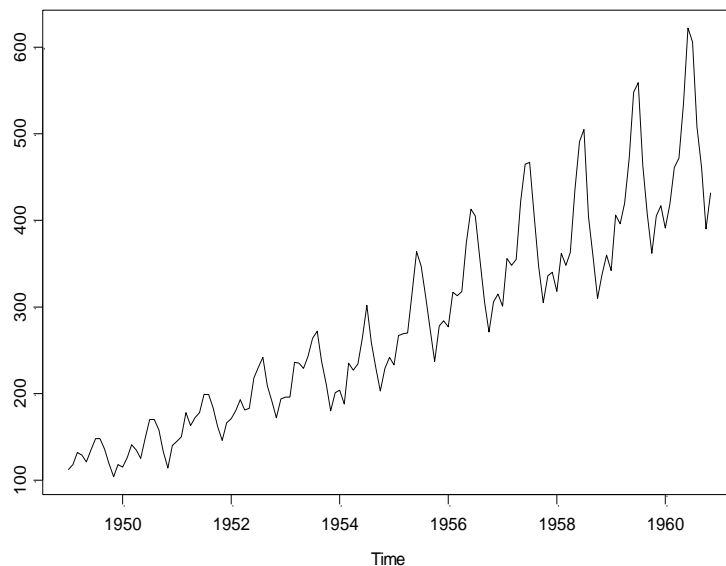


Figure 6. Time series plot of the airline data.

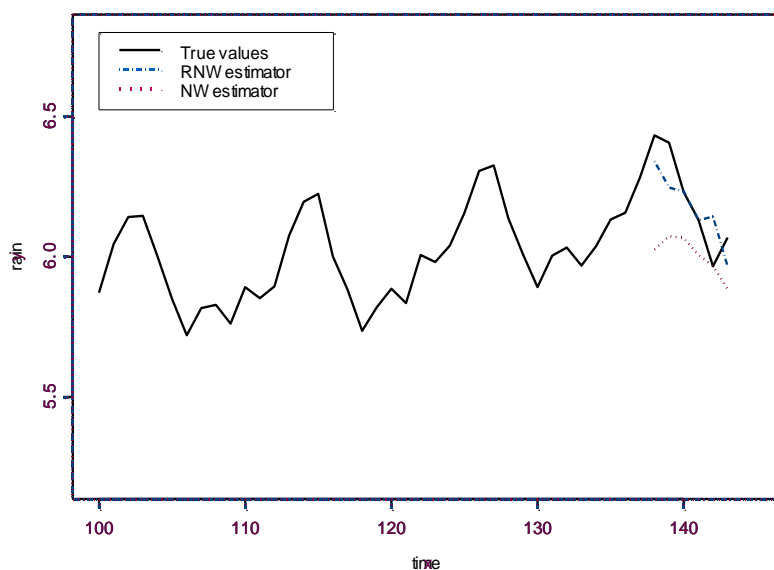


Figure 7. A plot of the transformed airline data with the predicted values

Table 6. The results for the prediction of the transformed airline data.

i	y_i	$\hat{m}_{RNW}(x)$	$\hat{m}_{NW}(x)$
138	6.43294	6.340948	5.965150
139	6.40688	6.246350	6.073604
140	6.230481	6.233550	6.066088
141	6.230481	6.131481	6.006746
142	5.966147	6.142851	5.967386
143	6.068426	5.965150	5.886037

2. Canadian Lynx Trappings

In the second application using real data, we consider the lynx data which is built in S-Plus program (The data give annual number of lynx trappings in the Mackenzie River District of North-West Canada for the period 1821 to 1934). Figure 8 shows the time plot of the data. We have transformed the data by taking the logarithm, then we used the first 107 observations to predict the last 7 observations using the RNW, $\hat{m}_{RNW}(x)$ and the NW, $\hat{m}_{NW}(x)$ estimators. We have used a bandwidth, $h = 0.8232566$. The MSE of the predicted values using the two estimators are as follows:

$$\text{MSE}(\hat{m}_{RNW}(x)) = 0.03402359, \quad \text{MSE}(\hat{m}_{NW}(x)) = 0.07619853.$$

Figure 9 shows the tail of the transformed data with the predicted values. Also, the results has been listed in Table 7. The results indicate that the RNW is better than the NW.

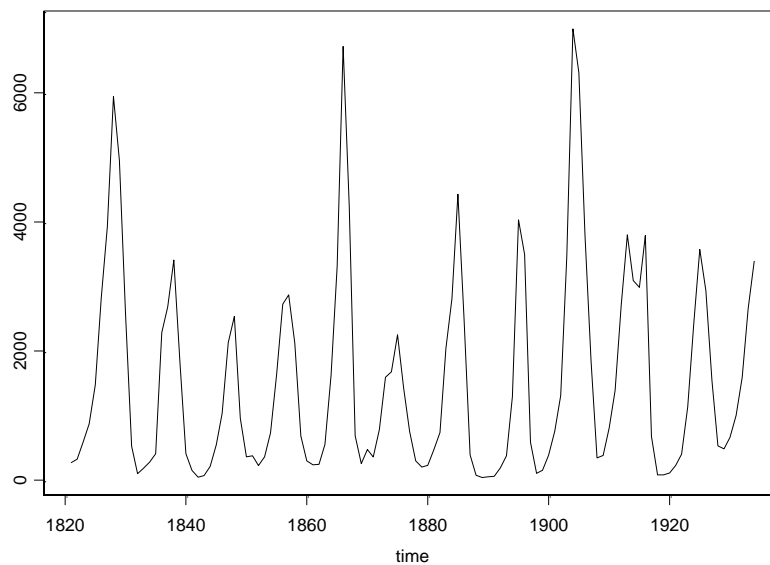


Figure 8. Time series plot of the lynx data.

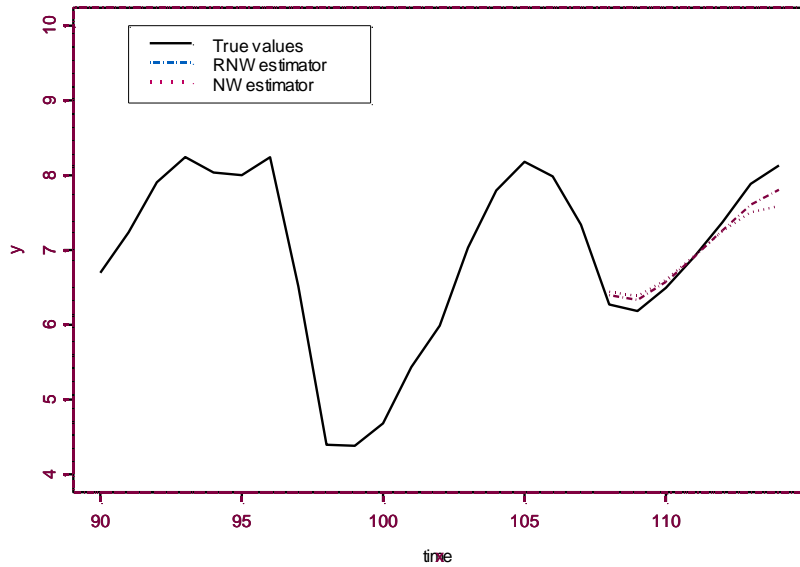


Figure 9. A plot of the transformed lynx data with the predicted values

Table 7. The results for the prediction of the transformed lynx data.

i	y_i	$\hat{m}_{RNW}(x)$	$\hat{m}_{NW}(x)$
108	6.270988	6.397733	6.441620
109	6.184149	6.331402	6.384472
110	6.495266	6.574513	6.600657
111	6.907755	6.905330	6.920557
112	7.371489	7.262187	7.251139
113	7.884953	7.606901	7.504852
114	8.130354	7.806515	7.588216

4. Discussion and Conclusions

In this paper, we compared theoretically between the RNW and the NW estimators of the conditional density and the regression mean function by looking at their asymptotic bias, variance and the mean

On the Nadaraya-Watson Estimators of the Regression Mean Function

squared error. This comparison has showed that the RNW has less bias and it is a design adaptive since it does not depend on the design density $g(x)$.

Also, a comparison of practical performance of the two estimators based on simulated and real data has been given. The results of the comparison indicated that the bias and the performance of the RNW estimator are better than that of the NW estimator. Three simulation studies with different regression mean function, $m(x)$, and design density function, $g(x)$, has been given. The results of these studies were consistent with the results of the theoretical comparison.

Firstly, we studied the influence of the mean regression function $m(x)$, on the performance of the two estimators. For that, we have used $m_1(x) = 1.5x^2$ in the first simulation study and $m_2(x) = 1 + 1.5x$ in the second one. Since $m'_1(x) = 0$ when $x = 0$, we have seen in Figure 1 and Figure 2, that the performance of the two estimators is very close in the interior around $x = 0$ where $m_1(x)$ has a minimum point. As we go away from $x = 0$ to the boundaries the performance of the RNW becomes better than that of the NW. We noted that from Table 3 and Table 4, for the same design density function $g(x)$ the MSE of the two estimators in the case of $m_1(x)$ is larger than MSE in the case of $m_2(x)$. This happens since $m''_1(x) \neq 0$ and $m''_2(x) = 0$, which is consistent with the theoretical results in Section 2.

Secondly, we studied the influence of the design density function $g(x)$, on the performance of the two estimators. For that, we have used $g_1(x) = \text{Uiniform}[-1, 1]$ and $g_2(x) = N(0, 1)$. From Table 3 and Table 4, we noted that the value of $\frac{\text{MSE}(\hat{m}_{RNW}(x))}{\text{MSE}(\hat{m}_{NW}(x))}$ in the case of $g_2(x)$ is much bigger than its value in the case of $g_1(x)$. The reason of this is that $g'_2(x) \neq 0$, while $g'_1(x) = 0$, which is consistent with the theoretical results in Section 2.

Thirdly, we have used $m_3(x) = 2x^{\frac{1}{2}} - \sin(2\pi(1-x)^2)$ which has minima, maxima, and inflection points. Table 5 and Figure 5 indicated that the performance of the two estimators is very close.

Finally, two applications using international real data have showed that the performance of the RNW is better than that of NW.

We conclude from this study that using the RNW estimator gives estimations better than that of the NW estimators especially at the

boundaries and when the design density $g(x)$ is highly clustered. The performance of the two estimators is approximately the same in the interior where the regression mean function $m(x)$ has maxima, minima or inflection points.

References:

- Box, G. E., Jenkins, G. M. and Reinsel, G. C. (1994). Time series analysis, third edition. Prentice-Hall.
- Cai, Z., (2002). Regression quantiles for time series data. *Econometric Theory*, 18, 169-192.
- Cai Z. (2001). Weighted Nadaraya-Watson regression estimation. *Statistics and probability letters*, 51, 307-318.
- De Gooijer J.G., Zerom, D. (2003). On conditional density estimation, *Statistica Neerlandica*, 57, 159-176.
- Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, 87, 998-1004.
- Hall, P., Wolff, R. and Yao, Q. (1999). Methods for estimating a conditional distribution. *Journal of the American Statistical Association*, 94, 154-163.
- Nadaraya, E. A. (1964). One estimating regression. *Theory of the Probability of its Applications*, 9, 141-142.
- Salha, R. and EL Shekh Ahmed, H. (2015). Reweighted Nadaraya-Watson estimator of regression mean. *International Journal of Statistics and Probability*, 4(1), 138-147.
- Silverman, B. W. (1986). Density estimation for statistics and data analysis. Chapman and Hall.
- Watson, G.S. (1964). Smooth regression analysis. *Sankhya Series A*, 26, 359-372.
- Yu, K and Stander, J. (2003). Quantile regression: applications and current research areas. *The Statistician*, 52 (3), 331-350.