Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 10(1), 2015, pages 739–750.
DOI: http://dx.doi.org/1016929.as.2015.739.66

General method of boundary correction in kernel regression estimation

Souraya Kheireddine, Abdallah Sayah and Djabrane Yahia[1]

Laboratory of Applied Mathematics, Mohamed Khider University of Biskra P.B. 145, 07000, Algeria

Received January 31, 2015; Accepted October 26, 2015

Abstract. Kernel estimators of both density and regression functions are not consistent near the finite end points of their supports. In other words, boundary effects seriously affect the performance of these estimators. In this paper, we combine the transformation and the reflection methods in order to introduce a new general method of boundary correction when estimating the mean function. The asymptotic mean squared error of the proposed estimator is obtained. Simulations show that our method performes quite well with respect to some other existing methods.

Résumé. Les estimateurs à noyau des fonctions de densité et de régression présentent des problèmes de convergence aux bords de leurs supports. En d'autre termes, les effets de bord affectent sérieusement les performances de leurs estimateurs. Dans cet article, nous combinons les méthodes de transformation et de réflexion, pour introduire une nouvelle méthode générale de correction de l'effet de bord lors de l'estimation de la moyenne. L'erreur quadratique moyenne asymptotique de l'estimateur proposé est obtenue. Les simulations montrent que notre méthode se comporte assez bien par rapport à d'autres méthodes existantes.

Key words: Boundary; Kernel estimation; Mean squared error; Regression; Estimation.
AMS 2010 Mathematics Subject Classification : 62G07; 62G20.

1. Introduction

Let Y be a real random variable (rv), and let X be a continuous covariable with probability density function f which is supported within $[0, \infty)$. The real rv's Y and X are repectively called variable of interest and predictor. Our goal is to estimate the regression function, which is the conditional expectation $m(x) := E(Y|X = x)$ (assuming $f(x) \neq 0$). Then the model can be written as

$$Y = m(X) + \epsilon \tag{1}$$

[1]Corresponding author Djabrane Yahia: yahia_dj@yahoo.fr

where ϵ is a rv such that $E(\epsilon|X) = 0$ and $Var(\epsilon|X) = \sigma^2 < \infty$.

There exist many interesting nonparametric estimators for the unknown regression function m. Examples of these last can be found in, for instance, Gasser and Müller (1979), Eubank (1988) and Fan and Gijbels (1996). Given a sample of independent replicates of (X, Y), the popular Nadaraya-Watson estimator Nadaraya (1964) and Watson (1964) of m is given by

$$m_n(x) = \frac{\sum_{i=1}^{n} Y_i K_h(x - X_i)}{\sum_{i=1}^{n} K_h(x - X_i)} \tag{2}$$

where $h := h_n$ ($h \to 0$ and $nh \to \infty$) is the bandwidth and $K_h(.) := K(./h)$, where K is an integrable smoothing kernel which usually is nonnegative, i.e., a symmetric probability density function with compact support. There have been numerous activities to study $m_n(x)$, see Härdle (1990) and Wand and Jones (1995) for a review.

Boundary effects are a well known problem in the nonparametric curve estimation setup, no matter if we think density estimation or regression. Moreover, both density and regression estimator usually show a sharp which increase in bias and variance when estimating them at points near the boundary region, i.e., for $x \in [0, h)$, this phenomenon is referred as "boundary effects". In the context of the regression function estimation, Gasser and Müller (1979) identified the unsatisfactory behavior of (2) for points in the boundary region. They proposed optimal boundary kernels but did not give any formulas. However, Gasser and Müller (1979) and Müller (2012) suggested multiplying the truncated kernel at the boundary zone or region by a linear function. Rice (1984) proposed another approach using a generalized jackknife. Schuster (1985) introduced a reflection technique for density estimation. Eubank and Speackman (1991) presented a method for removing boundary effects using a bias reduction theorem. Müller (1991) proposed an explicit construction of a boundary kernel which is the solution of a variational problem under asymmetric support. Moreover, Müller and Wang (1994) gave explicit formulas for a new class of polynomial boundary kernels and showed that these new kernels have some advantages over the smooth optimum boundary kernels in Müller (1991), i.e., these new kernels have higher mean squared error (MSE) efficiency. The local linear methods developed recently have become increasingly popular in this context (cf. Fan and Gijbels, 1996). More recently, in Dai and Sperlich (2010) a simple and effective boundary correction for kernel densitiy and regression estimator is proposed, by applying local bandwidth variation at the boundaries.

To remove the boundary effects a variety of methods have been developed in the literature, the most widely used is the reflection method, the boundary kernel method, the transformation method, the pseudo-data method and the local linear method. They all have their advantages and disadvantages. One of the drawbacks is that some of them (especially boundary kernels), can produce negative estimators. The recent work of Karunamuni and Alberts (2005) provides excellent selective review article on boundary kernel methods and their statistical properties in nonparametric density estimation. In the latter reference, a new boundary correction methodology in density estimation is proposed and studied. It is the purpose of this paper to extend this approach to the regression case.

The rest of the paper is organized as follows. Section 2 introduces our new nonparametric regression estimator and presents some asymptotic results. In Section 3, extensive simulations are carried out to compare the proposed estimator with outher ons. Proofs are relegated to Section 4.

2. Main results

In this paper, we combine the transformation and reflection boundary correction methods to estimate the mean function $m_n(x)$. At each point in the boundary region (i.e., for $x = ch$, $0 \leq c \leq 1$), we propose to investigate a class of estimators of the form

$$\widetilde{m}_n(x) = \frac{\sum_{i=1}^{n} Y_i \left\{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)\right\}}{\sum_{i=1}^{n} \left\{K_h\left(x + g_2\left(X_i\right)\right) + K_h\left(x - g_2\left(X_i\right)\right)\right\}}$$
$$:= \frac{\widetilde{\varphi}_n(x)}{\widetilde{f}_n(x)} \tag{3}$$

where h is the bandwidth, $K_h\left(.\right) := K\left(./h\right)$ and K is a kernel function and g_1, g_2 are two transformations that need to be determined. Also, let the kernel function K in (3) be a non-negative, symmetric function with support $[-1, 1]$, and satisfying

$$\int K\left(t\right) dt = 1, \int tK\left(t\right) dt = 0, \text{ and } 0 < \int t^2 K\left(t\right) dt < \infty,$$

that is, K is a kernel of order 2.

For $x \geq h$, $\widetilde{m}_n(x)$ reduces to the traditional kernel estimator $m_n(x)$ given in (2). Thus $\widetilde{m}_n(x)$ is a natural boundary continuation of the usual kernel estimator (2). Moreover, estimator (3) is non-negative as long as the kernel K is non-negative. Most importantly, the proposed estimator improves the bias while the variance remains almost unchanged.

We assume that the transformations g_1, g_2 in (3) are non-negative, continuous and monotonically increasing functions defined on $[0, \infty)$. Further assume that g_k^{-1} exists, $g_k(0) = 0$, $g_k' = 1$, and that g_k'' and g_k''' exist and are continuous on $[0, \infty)$, where g_k^{-1} denoting the inverse function of g_k (for $k = 1, 2$). Particularly, suppose that

$$g_1''\left(0\right) = \frac{\varphi'\left(0\right)}{\varphi\left(0\right)} C_{K,c} \quad \text{and} \quad g_2''\left(0\right) = \frac{f'\left(0\right)}{f\left(0\right)} C_{K,c} \tag{4}$$

where

$$C_{K,c} := 2\int_c^1 \left(t - c\right) K\left(t\right) dt \left(2\int_c^1 \left(t - c\right) K\left(t\right) dt + c\right)^{-1}.$$

Suppose further that, $f^{(j)}$, $\varphi^{(j)}$ and $m^{(j)}$ the $j^{th}-$derivatives of f, φ and m exist and are continued on $[0, \infty)$, $j = 0, 1, 2$, with $f^{(0)} = f$, $\varphi^{(0)} = \varphi$ and $m^{(0)} = m$.
The bias and variance of our estimator are given in the following theorem, which is the main result of this paper.

Theorem 1. *Under the above conditions on f, φ, m, g_1, g_2, and K. For the estimate $\widetilde{m}_n(x)$ defined in (3), we have for $x = ch$, $0 \leq c \leq 1$:*

$$Bias(\widetilde{m}_n(x)) = \frac{h^2\left(A_1 - m\left(x\right)A_2\right)}{f\left(x\right)} + o(h^2), \tag{5}$$

and

$$Var\left(\widetilde{m}_n(x)\right) = \frac{f(0)\sigma^2(0)}{nhf^2(x)}\left(\int_{-1}^{1} K^2\left(t\right)dt + 2\int_{c}^{1} K\left(t\right)K\left(2c-t\right)dt\right) + o\left(\frac{1}{nh}\right). \tag{6}$$

where

$$A_1 := \varphi''\left(0\right)\int_{-1}^{1} t^2K\left(t\right)dt - \left[g_1'''\left(0\right)\varphi\left(0\right) + 3g_1''\left(0\right)\left(\varphi^{(1)}\left(0\right) - g_1''\left(0\right)\varphi\left(0\right)\right)\right]\left(\int_{-1}^{1} t^2K\left(t\right)dt + c^2\right), \tag{7}$$

$$A_2 := f''\left(0\right)\int_{-1}^{1} t^2K\left(t\right)dt - \left[g_2'''\left(0\right)f\left(0\right) + 3g_2''\left(0\right)\left(f'\left(0\right) - g_2''\left(0\right)f\left(0\right)\right)\right]\left(\int_{-1}^{1} t^2K\left(t\right)dt + c^2\right) \tag{8}$$

and $\sigma^2(x) = Var\left(Y/X = x\right)$.

Hence, the MSE of $\widetilde{m}_n(x)$ is

$$MSE\left(\widetilde{m}_n(x)\right) = Bias^2(\widetilde{m}_n(x)) + Var\left(\widetilde{m}_n(x)\right)$$

The asymptotic MSE of $\widetilde{m}_n(x)$ is

$$AMSE\left(\widetilde{m}_n(x)\right) = \frac{h^4\left(A_1 - m\left(x\right)A_2\right)^2}{f^2\left(x\right)} + \frac{f(0)\sigma^2(0)}{nhf^2(x)}\left(\int_{-1}^{1} K^2\left(t\right)dt + 2\int_{c}^{1} K\left(t\right)K\left(2c-t\right)dt\right).$$

On the basis of Theorem 1, the asymptotic optimal bandwidth that minimizes the $AMSE$ is

$$h_{opt} = Cn^{-1/5} \quad \text{with} \quad C = \left(\frac{\sigma^2\left(0\right)f(0)\left(2\int_c^1 K\left(t\right)K\left(2c-t\right)dt + \int K^2\left(t\right)dt\right)}{4\left(A_1 - m\left(x\right)A_2\right)^2}\right)^{1/5}. \tag{9}$$

Remark 1. Functions satisfying the conditions (4) can be easily constructed. We employ the following transformation in our investigation. For $0 \le c \le 1$, define

$$g_k(y) = y + \frac{1}{2}d_ky^2 + \lambda_0 d_k^2 y^3, \qquad k = 1, 2 \tag{10}$$

where $d_1 = g_1''\left(0\right)$ (resp. $d_2 = g_2''\left(0\right)$) and λ_0 is a positive constant such that $12\lambda_0 > 1$. This condition on λ_0 is necessary for $g_k(y)$ of (10) to be an increasing function in y.

Remark 2. The choice h_{opt} of h is only possible in a simulation study, when all required quantities are known, but not in a real data situation. To select the bandwidth for the new method in practice, we can replace the unknown quantities in (9) by their estimates. Another method is to use leave-one-out cross-validation (cf. Härdle and Vieu, 1992) to select the bandwidth h, i.e., we find h by minimizing

$$CV\left(h\right) = \sum\nolimits_{i=1}^{n}\left(y_j - \widetilde{m}_{i,h}(x_i)\right)^2,$$

here $\widetilde{m}_{i,h}(.)$ is the proposed regression estimate by leaving the ith observation out.

3. Simulation results

In this section, we present some simulation results which are designed to illustrate the performance of our estimator (3) for small sample and large sizes. For comparison purposes, the local linear and the classical Nadaraya–Watson estimators (2) were also considered. Recently, local polynomial fitting, and particularly its special case - local linear fitting - have become increasingly popular in light of recent works by Cleveland and Loader (1996) Fan (1992) and several others. It has the advantages of achieving full asymptotic minimax efeciency and automatically correcting for boundary bias. A review of local polynomial smoothing is given in Fan and Gijbels (1996). The local linear regression estimator is given by

$$\widehat{m}_n(x) = \sum_{j=1}^{n} w_j Y_j / w_j, \quad w_j := K_h \left(X_j - x\right)\left(S_{n,2} - S_{n,1}\left(X_j - x\right)\right),$$

where $S_{n,k} := \sum_{j=1}^{n} K_h \left(X_j - x\right)\left(X_j - x\right)^k$, for $k = 1, 2$.

To assess the effect of the correction methods near the boundaries, the following models are investigated:

$$Model\ 1 : m_1\left(x\right) = 2x + 1 \quad \text{and} \quad Model\ 2 : m_2\left(x\right) = 2x^2 + 3x + 1$$

and errors ε_j, assumed to be standard normally distributed independent rv's. Likewise, consider two cases of density f with support $[0, \infty)$ of the continuous covariable X :

$$density\ 1 : f_1\left(x\right) = \exp(-x) \quad \text{and} \quad density\ 2 : f_2\left(x\right) = \frac{2}{\pi\left(1 + x^2\right)0}, \quad x \geq 0.$$

For each density f_1, f_2 and models m_1, m_2 we calculate the absolute biases and $MSE's$ of the proposed general transformation and reflection (GTR), the local linear (LL) and Nadaraya-Watson (NW) estimators, in left boundary region (i.e., $x = ch$; for $c =$ 0.1, 0.2, 0.3, 0.4 and 0.5). The bandwidth selection is based on cross-validation procedure. The main reason for this choice is that it provides a fair basis for comparison among the different estimators regardless of bandwidth effects.

Throughout our simulations, we use the Epanechnikov kernel (cf. Epanechnikov, 1969)

$$K\left(t\right) = (3/4)\left(1 - t^2\right) 1_{[-1,1]}\left(t\right),$$

where $1_A\left(.\right)$ denotes the indicator function of a set A.

The simulated sample sizes are $n = 50$ (small) and $n = 500$ (large). All results are calculated by averaging over 1000 simulation runs. For each model and each density, we calculate the absolute bias and the MSE of the estimators at the points in the mentioned boundary region. The results are shown in Tables **1** and **2**. We see that in all cases the standard Nadaraya-Watson estimator $m_n(x)$ is the worst one. This is clearly due to the boundary effect. Furthermore, when looking at the $MSE's$, our new method outperforms the others. The bias is about the same for our method and the local linear one.

n		$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE
		$c = .1$		$c = .2$		$c = .3$		$c = .4$		$c = .5$	
	GTR	.0141	.0613	.0381	.0631	.0461	.0584	.0512	.0571	.0499	.0580
50	NW	.2678	.1362	.2192	.1120	.1775	.0894	.1365	.0700	.0977	.0617
	LL	.0375	1.237	.0064	.3649	.0064	.1468	.0101	.1167	.0011	.1265
					Model 1						
	GTR	.0109	.0083	.0126	.0090	.0162	.0091	.0199	.0084	.0229	.0080
500	NW	.1747	.0393	.1503	.0313	.1217	.0233	.0954	.0163	.0690	.0115
	LL	.0127	.0386	.0025	.0240	.0006	.0169	.0049	.0123	.0013	.0114
	GTR	.1361	.0595	.1877	.0786	.1841	.0785	.1304	.0656	.0873	.0715
50	NW	.5705	.3934	.4356	.2540	.3413	.1782	.2923	.1453	.2550	.1295
	LL	.0694	.1252	.1334	.0903	.1991	.1002	.2158	.1054	.2674	.1279
					Model 2						
	GTR	.0940	.0141	.0948	.0146	.0778	.0131	.0582	.0107	.0337	.0087
500	NW	.3520	.1320	.2854	.0887	.2343	.0627	.1951	.0454	.1809	.0397
	LL	.0458	.0300	.0955	.0263	.1327	.0305	.1517	.0358	.1780	.0428

Table 1. Bias and MSE of the indicated regression estimators at boundary, case of density1

n		$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE
		$c = .1$		$c = .2$		$c = .3$		$c = .4$		$c = .5$	
	GTR	.1131	.1189	.1064	.1084	.0876	.0813	.1216	.0769	.1166	.0728
50	NW	.8697	.8014	.6174	.4329	.4455	.2529	.2630	.1279	.1205	.0734
	LL	.2662	.7402	.0818	.2944	.0245	.2785	.0044	.1955	.0060	.1803
					Model 1						
	GRT	.0496	.0196	.0520	.0167	.0455	.0131	.0391	.0107	.0481	.0098
500	NW	.7162	.5180	.5016	.2577	.3621	.1374	.2511	.0690	.1352	.0251
	LL	.0063	.0601	.0054	.0373	.0040	.0238	.0007	.0159	.0009	.0152
	GRT	.1257	.1758	.1205	.1411	.1339	.1119	.1479	.0946	.2261	.1172
50	NW	.6272	.4902	.6220	.4794	.6345	.4856	.7076	.5848	.7271	.6305
	LL	.1849	1.840	.0657	.2573	.1288	.1719	.2294	.1398	.4228	.2557
					Model 2						
	GRT	.0505	.0260	.0562	.0201	.0663	.0179	.0575	.0135	.0801	.0137
500	NW	.3744	.1526	.3511	.1344	.3218	.1136	.3077	.1037	.3646	.1408
	LL	.0004	.0752	.0274	.0489	.0806	.0359	.1517	.0418	.2958	.1023

Table 2. Bias and MSE of the indicated regression estimators at boundary, case of density2

4. Proofs

Proof of (5). For $x = ch$, $0 \leq c \leq 1$, we have

$$\begin{aligned}\widetilde{m}_n(x) &= \frac{\sum_{i=1}^{n} Y_i \left\{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)\right\}}{\sum_{i=1}^{n} \left\{K_h\left(x + g_2\left(X_i\right)\right) + K_h\left(x - g_2\left(X_i\right)\right)\right\}} \\ &:= \frac{\widetilde{\varphi}_n(x)}{\widetilde{f}_n(x)},\end{aligned}$$

where g_1 and g_2 are given in (4). For the numerator $\widetilde{\varphi}_n(x)$, we have

$$\begin{aligned}
E\left[\widetilde{\varphi}_n(x)\right] &= \frac{1}{h}\int\int \left\{K_h\left(x+g_1(u)\right)+K_h\left(x-g_1(u)\right)\right\} yf\left(u,y\right)dydu \\
&= \frac{1}{h}\int \left\{K_h\left(x+g_1(u)\right)+K_h\left(x-g_1(u)\right)\right\}\varphi\left(u\right)du \\
&= \frac{1}{h}\int K_h\left(x+g_1(u)\right)\varphi\left(u\right)du + \frac{1}{h}\int K_h\left(x-g_1(u)\right)\varphi\left(u\right)du \\
&=: I_1 + I_2,
\end{aligned}$$

where $\varphi\left(u\right) = \int yf\left(u,y\right)dy$.

Let $t = \left(x+g_1(u)\right)/h$, then

$$I_1 = \int_c^1 K\left(t\right)\frac{\varphi\left(g_1^{-1}\left(h\left(t-c\right)\right)\right)}{g_1^{(1)}\left(g_1^{-1}\left(h\left(t-c\right)\right)\right)}dt.$$

A Taylor expansion of order 2 of the function $\varphi\left(g_1^{-1}\left(.\right)\right)/g_1^{(1)}\left(g_1^{-1}\left(.\right)\right)$ at $t=c$ gives

$$\begin{aligned}
I_1 &= \int_c^1 K\left(t\right)\left[\varphi\left(0\right)+h\left(t-c\right)\left(\varphi'\left(0\right)-g_1''\left(0\right)\varphi\left(0\right)\right)\right. \\
&\quad \left. +\frac{h^2\left(t-c\right)^2}{2}\left\{\varphi''\left(0\right)-g_1'''\left(0\right)\varphi\left(0\right)-3g_1''\left(0\right)\left(\varphi'\left(0\right)-g''\left(0\right)\varphi\left(0\right)\right)\right\}\right]dt + o\left(h^2\right) \\
&= \varphi\left(0\right)\int_c^1 K\left(t\right)dt + h\left\{\varphi'\left(0\right)-g_1''\left(0\right)\varphi\left(0\right)\right\}\int_c^1\left(t-c\right)K\left(t\right)dt \\
&\quad + \frac{h^2}{2}\left\{\varphi''\left(0\right)-g_1'''\left(0\right)\varphi\left(0\right)-3g_1'''\left(0\right)\left(\varphi'\left(0\right)-g_1''\left(0\right)\varphi\left(0\right)\right)\right\}\int_c^1\left(t-c\right)^2K\left(t\right) \\
&\quad + o\left(h^2\right).
\end{aligned} \tag{11}$$

Similarly,

$$\begin{aligned}
I_2 &= \varphi\left(0\right)\int_{-1}^c K\left(t\right)dt - h\left\{\varphi'\left(0\right)-g_1''\left(0\right)\varphi\left(0\right)\right\}\int_{-1}^c\left(t-c\right)K\left(t\right)dt \\
&\quad + \frac{h^2}{2}\left\{\varphi''\left(0\right)-g_1'''\left(0\right)\varphi\left(0\right)-3g_1''\left(0\right)\left(\varphi'\left(0\right)-g_1''\left(0\right)\varphi\left(0\right)\right)\right\}\int_{-1}^c\left(t-c\right)^2K\left(t\right)dt \\
&\quad + o\left(h^2\right).
\end{aligned} \tag{12}$$

Using the properties of K, we have

$$\int_{-1}^c tK\left(t\right)dt = -\int_c^1 K\left(t\right)dt \quad \text{and} \quad \int_{-1}^c K\left(t\right)dt = 1-\int_c^1 K\left(t\right)dt.$$

Also, by the existence and the continuity of $\varphi''(.)$ near 0, we have for $x = ch$,

$$\begin{array}{l}\varphi(0) = \varphi(x) - ch\varphi'(x) + \frac{(ch)^2}{2}\varphi''(x) + o\left(h^2\right), \\ \varphi'(x) = \varphi'(0) + ch\varphi''(0) + o(h), \\ \varphi''(x) = \varphi''(0) + o(1).\end{array} \tag{13}$$

Now combining (11) and (12) and using the properties of K along with (13), we have for $x = ch$, $0 \le c \le 1$

$$\begin{aligned}E\left[\widetilde{\varphi}_n(x)\right] &= \frac{1}{h}E\left[K_h\left(x+g_1\left(X_i\right)\right)Y_i\right] + \frac{1}{h}E\left[K_h\left(x-g_1\left(X_i\right)\right)Y_i\right] \\ &= \varphi(0)\int_c^1 K(t)\,dt + \varphi(0)\int_{-1}^c K(t)\,dt + h\left\{\varphi'(0) - g_1''(0)\varphi(0)\right\}\int_c^1 (t-c)K(t)\,dt \\ &- h\left\{\varphi'(0) - g_1''(0)\varphi(0)\right\}\int_{-1}^c (t-c)K(t)\,dt \\ &+ \frac{h^2}{2}\left\{\varphi'(0) - g_1'''(0)\varphi(0) - 3g_1''(0)\left(\varphi'(0) - g_1''(0)\varphi(0)\right)\right\}\int_c^1 (t-c)^2 K(t) \\ &+ \frac{h^2}{2}\left\{\varphi''(0) - g_1'''(0)\varphi(0) - 3g_1''(0)\left(\varphi'(0) - g_1''(0)\varphi(0)\right)\right\}\int_c^1 (t-c)^2 K(t)\,dt \\ &+ o\left(h^2\right).\end{aligned} \tag{14}$$

Furthermore, the kernel K provides

$$\int_{-1}^1 (t-c)^2 K(t)\,dt = \int_{-1}^1 t^2 K(t)\,dt + c^2,$$

and

$$\int_c^1 (t-c)K(t)\,dt - \int_{-1}^c (t-c)K(t)\,dt = 2\int_c^1 (t-c)K(t)\,dt + c.$$

From (14) we have

$$
\begin{aligned}
E\left[\widetilde{\varphi}_n(x)\right] &= \varphi(x) + h\left\{\varphi'(0) - g_1'(0)\varphi(0)\right\}\{2\int_c^1 (t-c)K(t)\,dt + c\} \\
&+ \frac{h^2}{2}\left\{\varphi''(0) - g_1'''(0)\varphi(0) - 3g_1''(0)\left(\varphi'(0) - g_1''(0)\varphi(0)\right)\right\} \\
&\times \{\int_{-1}^1 t^2 K(t)\,dt + c^2\} + o\left(h^2\right) \\
&= \varphi(x) + h\left\{2\varphi'(0)\int_c^1 (t-c)K(t)\,dt - g_1''(0)\varphi(0)\{2\int_c^1 (t-c)K(t)\,dt + c\}\right\} \\
&+ \frac{h^2}{2}\left\{\varphi''(0)\int_{-1}^1 t^2K(t)\,dt\right. \\
&\left. - \left[g_1'''(0)\varphi(0) + 3g_1''(0)\left(\varphi'(0) - g_1''(0)\varphi(0)\right)\right](\int_{-1}^1 t^2K(t)\,dt + c^2)\right\} \\
&+ o\left(h^2\right).
\end{aligned} \tag{15}
$$

Under the condition (4) on the transformation g_1, the second order term of the right-hand side of (15) is zero. It can be shown that

$$E\left[\widetilde{\varphi}_n(x)\right] - \varphi(x) =: h^2 A_1 + o\left(h^2\right),$$

where A_1 is given in (7).

Similarly, we can get

$$
\begin{aligned}
E\left[\widetilde{f}_n(x)\right] &= f(x) + h\left\{2f'(0)\int_c^1 (t-c)K(t)\,dt - g_2''(0)f(0)\{2\int_c^1 (t-c)K(t)\,dt + c\}\right\} \\
&+ \frac{h^2}{2}\left\{f''(0)\int_{-1}^1 t^2K(t)\,dt\right. \\
&\left. - \left[g_2''(0)f(0) + 3g_2''(0)\left(f'(0) - g_2''(0)f(0)\right)\right]\{\int_{-1}^1 t^2K(t)\,dt + c^2\}\right\} \\
&+ o\left(h^2\right)
\end{aligned} \tag{16}
$$

Substitute $g_2''(0)$, the second term of the right-hand side of (16) is zero. Then

$$E\left[\widetilde{f}_n(x)\right] - f(x) =: h^2 A_2 + o\left(h^2\right)$$

where A_2 is given in (8). Hence

$$\widetilde{m}_n(x) = \frac{h^2 A_1 + o\left(h^2\right)}{h^2 A_2 + o\left(h^2\right)} = m(x) + \frac{h^2\left(A_1 - m(x) A_2\right)}{f(x)} + o\left(h^2\right).$$

The asymptotic bias result (5) follows directly.

Proof of (6). In order to find the asymptotic variance of the proposed estimator (3), we may write

$$\widetilde{m}_n(x) = \sum_{i=1}^{n} W_{ni}(x) Y_i,$$

with

$$W_{ni}(x) = \frac{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)}{\sum_{i=1}^{n}\left\{K_h\left(x + g_2\left(X_i\right)\right) + K_h\left(x - g_2\left(X_i\right)\right)\right\}}.$$

The weights $W_{ni}(x)$ are nonnegative and satisfy $\sum_{i=1}^{n} W_{ni}(x) = 1$, for all $x \in \mathbb{R}$. Moreover, we have

$$\begin{aligned}\widetilde{m}_n(x) - m(x) &= \sum_{i=1}^{n} W_{ni}(x)\left\{Y_i - m\left(X_i\right)\right\} + \sum_{i=1}^{n} W_{ni}(x)\left\{m\left(X_i\right) - m(x)\right\} \\ &=: \mathcal{J}_1 + \mathcal{J}_2.\end{aligned}$$

Here $\mathcal{J}_1$ is the variance which is study here. Recall that the predictable quadratic variation of $\mathcal{J}_1$ equals

$$\widetilde{f}_n^2(x) \sum_{i=1}^{n} W_{ni}^2(x) \sigma^2(X_i) = (nh)^{-2} \sum_{i=1}^{n} \sigma^2(X_i)\left\{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)\right\}^2,$$

where $\sigma^2(.)$ is the conditional variance i.e., $\sigma^2(\cdot) = Var\left(Y|X = \cdot\right)$.

For $x = ch$, $0 \leq c \leq 1$, we have, using a Taylor expansion of order 2

$$\begin{aligned}&E\left[(nh)^{-2} \sum_{i=1}^{n} \sigma^2(X_i)\left\{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)\right\}^2\right] \\ &= \frac{1}{nh^2} E\left[\sigma^2(X_i)\left\{K_h\left(x + g_1\left(X_i\right)\right) + K_h\left(x - g_1\left(X_i\right)\right)\right\}^2\right] \\ &= \frac{1}{nh^2} \int \sigma^2(u)\left\{K_h\left(x + g_1(u)\right) + K_h\left(x - g_1(u)\right)\right\}^2 f(u) du \\ &= \frac{1}{nh^2}\left[\int \sigma^2(u) K_h^2\left(x + g_1(u)\right) f(u) du + \int \sigma^2(u) K_h^2\left(x - g_1(u)\right) f(u) du\right] \\ &+ \frac{2}{nh^2} \int \sigma^2(u) K_h\left(x + g_1(u)\right) K_h\left(x - g_1(u)\right) f(u) du \\ &=: \mathcal{J}_{11} + \mathcal{J}_{12}.\end{aligned}$$

Firstly,

$$\begin{aligned}
\mathcal{J}_{11} &= \frac{1}{nh^2}\left[h\int_c^1 \sigma^2(g_1^{-1}((t-c)h))K^2(t)\frac{f\left(g_1^{-1}((t-c)h)\right)}{g_1'\left(g_1^{-1}((t-c)h)\right)}dt\right. \\
&\left.+h\int_{-1}^{c} \sigma^2(g_1^{-1}((c-t)h))K^2(t)\frac{f\left(g_1^{-1}((c-t)h)\right)}{g_1'\left(g_1^{-1}((c-t)h)\right)}dt\right] \\
&= \frac{f(0)\sigma^2(0)}{nh}\int_{-1}^{1} K^2(t)\,dt + o(\frac{1}{nh}). \qquad (17)
\end{aligned}$$

Next we consider $\mathcal{J}_{12}$. By the continuity property of g_1'' and by a Taylor expansion of order 2 of g_1, we have

$$\begin{aligned}
g_1((c-t)h) &= g_1(0) + (t-c)(-h)g_1'(0) + O\left(h^2\right) \\
&= (c-t)h + O\left(h^2\right),
\end{aligned}$$

since $g_1(0) = 0$ and $g_1'(0) = 1$. Using (17) and by the change of variables, $x + g_1(y) = ht$, we obtain

$$\begin{aligned}
\mathcal{J}_{12} &= \frac{2}{nh^2}\int_0^\infty \sigma^2(u)K_h(x+g_1(X_i))K_h(x-g_1(X_i))f(u)du \\
&= \frac{2}{nh}\int_c^1 \sigma^2\left(g_1^{-1}(th-x)\right)K(t)K_h\left(x-g_1(g_1^{-1}(th-x))\right)f(g_1^{-1}(th-x))dt \\
&= \frac{2}{nh}\int_c^1 \sigma^2\left(g_1^{-1}(th-x)\right)K(t)K_h\left(x-(t-c)h+O\left(h^2\right)\right)f(g_1^{-1}(th-x))dt \\
&= \frac{2}{nh}\int_c^1 \sigma^2(0)K(t)K(2c-t+O(h))(f(0)+O(h))\,dt \\
&= \frac{2\sigma^2(0)f(0)}{nh}\int_c^1 K(t)K(2c-t)\,dt + o\left(\frac{1}{nh}\right). \qquad (18)
\end{aligned}$$

The proof of (6) now follows from (17) and (18), which achieves the proof of Theorem 1.

Acknowledgement. *We thank the editor and the referee for their constructive and useful comments that led to a much improved paper.*

References

Cleveland, W. S. and Loader, C. (1996). Smoothing by local regression: Principles and methods. *In Statistical Theory and Computational Aspects of Smoothing* (W. Härdle and M. G. Schimek, eds.) 10–49. Physica, Heidelberg.

Dai, J., Sperlich, S. (2010). Simple and effective boundary correction for kernel densities and regression with an application to the world income and Engel curve estimation. *Computational Statistics & Data Analysis*, 54(11), 2487-2497.

Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, 14(1), 153-158.

Eubank, R. L. (1988). *Spline smoothing and nonparametric regression.* Marcel Dekker, New York.

Eubank, R. L., Speckman, P. L. (1991). A bias reduction theorem with applications in nonparametric regression. *Scandinavian Journal of Statistics*, 211-222.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American statistical Association*, 87(420), 998-1004.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications.* Chapman and Hall, London.

Gasser, T., Müller, H. G. (1979). *Kernel estimation of regression functions* (pp. 23-68). Springer Berlin Heidelberg.

Härdle, W. (1990). *Applied nonparametric regression* (Vol. 27). Cambridge: Cambridge university press.

Härdle, W. and Vieu, P. (1992). Kernel regression smoothing of time series. *Journal of Time Series Analysis*, 13(3), 209-232.

Karunamuni, R. J., Alberts, T. (2005). On boundary correction in kernel density estimation. *Statistical Methodology,* 2(3), 191-212.

Müller, H. G. (2012). *Nonparametric regression analysis of longitudinal data* (Vol. 46). Springer Science & Business Media. New York.

Müller, H. G. (1991). Smooth optimum kernel estimators near endpoints. *Biometrika*, 78(3), 521-530.

Müller, H. G., Wang, J. L. (1994). Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics*, 61-76.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1), 141-142.

Rice, J. (1984). Boundary modification for kernel regression. *Communications in Statistics-Theory and Methods*, 13(7), 893-900.

Schuster, E. F. (1985). Incorporating support constraints into nonparametric estimators of densities. *Communications in Statistics-Theory and Methods*, 14(5), 1123-1136.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*, London: Chapman and Hall.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics*, Series A, 359-372.