Generalized Kernel Regression Estimator for Dependent Size-Biased Data

Yogendra P. Chaubey
Department of Mathematics and Statistics
Concordia University, Montréal, (QC) Canada H3G 1M8
E-mail:chaubey@alcor.concordia.ca

Naâmane Laïb*
L.S.T.A., Université Paris 6, France
E-mail: naamane.laib@upmc.fr

Jun Li
Department of Mathematics and Statistics
Concordia University, Montréal, (QC) Canada H3G 1M8
E-mail:lijun@mathstat.concordia.ca

January 31, 2011

Abstract

This paper considers nonparametric regression estimation in the context of dependent biased non-negative data using a generalized asymmetric kernel. It may be applied to a wider variety of practical situations, such as the length and size biased data. We derive theoretical results using a deep asymptotic analysis of the behavior of the estimator that provides consistency and asymptotic normality in addition to the evaluation of the asymptotic bias term. The asymptotic mean squared error is also derived in order to obtain the optimal value of smoothing parameters required in the proposed estimator. The results are stated under a stationary ergodic assumption, without assuming any traditional mixing conditions. A simulation study is carried out to compare the proposed estimator with the local linear regression estimate.

---

*Corresponding author

1   Introduction

Ordinary kernel regression does not provide admissible values of the regression, or its functionals at the boundaries for restricted support regressions (see, e.g., Chaubey *et al.* (2010) for further discussion). In this paper consider the setting of biased data that typically features non-negative observations. We extend the methodology given in Chaubey *et al.* (2010) to this setup, for constructing a nonparametric regression estimator that allows to deal with the boundary bias problem. In this setup the density of biased data is given by a target density weighted by some function of the observations. For example, consider the bivariate case where (U, V) is a 2-dimensional random vector with probability density function (pdf) $f(u, v)$. Suppose that the data are collected from another random vector (X, Y) with pdf $f_w(x, y)$, which is related to f as follows:

$$f_w(x, y) = \frac{\kappa(x) w(x, y) f(x, y)}{\mu}, \tag{1.1}$$

where $\mu = \mathbb{E}(\kappa(U) w(U, V))$ (which is assumed to be finite), $w(\cdot, \cdot)$ and $\kappa(\cdot)$ are nonnegative functions but $\kappa(\cdot)$ not necessarily needs to be known. Then f_w is known as a weighted density and the resulting data is known as *weighted* or *biased* data.

This covers a wide variety of practical cases, since the weighted data appear in a variety of situations such as that of missing data, damaged observations, sociological studies, econometrics, survival analysis, biomedicine and physics, among others. As a particular case, it can be applied to length biased data sampling, which is the most frequently analyzed in the literature on biased data. In such a case the probability of observing an individual at a given site is proportional to the individual's length of stay at that site.

Notice that the length biased data appear particularly in situations related to the renewal process. Consider, for instance in an univariate context, a natural process to generate a random variable X with density $f(x)$. To select a random sample of observations from X, we use a selection procedure which gives the same chance of each observation to be selected by the original mechanism. In practice it may happen that drawing a direct sample from X is impossible. In fact, an observation $X = x$, may be included with relative chance proportional to its length x. Therefore, the common probability density g of recorded biased observation, say $U_1, \ldots, U_n$, is given by $g(x) = xf(x)/\mu$ where $x > 0$ and μ is the expected mean corresponding to the density f, which is assumed to be finite. It should be noted that size biasing makes sense only for positive data (see Patil and Rao (1978)), that is our motivation to consider regression for the weighted data for non-negative random variables.

There are many other practical situations that lead to biased data sets. In an industrial setting, Cox (1969) studied the problem of sampling fibres and the estimation of fibre length

distribution. In the area of forestry, the size measure is usually proportional to either length or area (see Muttlak and McDonald (1990)). In economic context, Olave *et al.* (1998) studied the relationship between the time of unemployment of the members of a population and some covariates, such as age, starting from a sample of unemployed people at a specific moment in time. For other interesting examples of weighted distributions in practice, one may refer to Patil and Rao (1978) and Patil *et al.* (1988).

An example to show the usefulness of the introduction of the function κ in the general weighting scheme in Eq. (1.1) is given in Sköld (1999). For an overview of the present state of the art and more practical examples, one may refer to the article by Cristóbal and Alcalà (2001). Model (1.1) due to its general form may be applied to a wider variety of practical situations. It is studied by Sköld (1999) who gives the mean square convergence for both classical kernel estimator and the local linear estimator. Firstly, Ahmad (1995) considered nonparametric regression for the biased data, however in a special case when $\kappa(u) = 1$. The same problem has also been studied by Cristóbal and Alcalá (2000) when $\kappa(x) = 1$ and $w(x, y) = y$, where the authors proposed several estimators for the regression function and studied their asymptotic optimal bandwidth and asymptotic mean squared error. Some other papers, such as Jones (1991), Ojeda *al.* (2004), Vardi (1982) and Wu and Mao (1996), may also be of interest to the reader in this connection, though mostly in the independent identically distributed (*iid*) setup. However, a great deal of data in econometrics, engineering and natural sciences, among other areas, occur in the form of time series in which observations are dependent. Our goal, therefore, is to consider the case of biased data sampled from a stationary ergodic process to allow generality in the dependence structure. We avoid the widely used strong mixing condition and its variants as a dependence measure.

Following the idea in Chaubey *et al.* (2010), we consider a perturbed version of the regression function estimate that allows to deal with boundary bias problem of at 0. This estimator is very simple and may be easily generalized to multivariate case. Section 2 outlines the motivation and the form of the new estimator and Section 3 presents the main results concerning the asymptotic behavior of the estimator, including consistency, asymptotic normality and evaluation of the bias term. The asymptotic mean squared error is also derived and the optimal value of smoothing parameter is discussed, at the interior points of the interval as well as on the boundaries. Our results may applied for both mixing and non mixing processes. In this context, the martingale techniques play a vital role that allow us to obtain optimal results as in the *iid* setting. Section 4 is devoted to a simulation study comparing the performance of the new estimator with that of the local linear estimator and the traditional kernel estimator in this context. The proofs of the results in Section 3 are relegated to Section 5.

2   Smooth Estimator of the Regression Function

Let $Z_i = (X_i, Y_i)_{i \in \mathbb{N}^*}$ be a $\mathbb{R}^+ \times \mathbb{R}^+$-valued strictly stationary ergodic process defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Let $\mathbb{E}_w(\cdot)$ and $\mathrm{Var}_w(\cdot)$ be the moments when these are calculated with the density $f_w(\cdot, \cdot)$ given in (1.1) or with its marginal $f_w(\cdot)$ defined below. Let $f(\cdot)$ be

the pdf of $U_1, \ldots, U_n$ and $f_w(\cdot)$ that of $X_1, \ldots, X_n$, which are assumed to be bounded and continuous on $[0, \infty)$. Let $\phi(\cdot)$ be a Boral function of $\mathbb{R}^+$ into $\mathbb{R}$ and such that $\mathbb{E}(\phi(V)) < \infty$ and let $m(u) := \mathbb{E}(\phi(V_1)|U_1 = u)$ be the conditional mean function of $\phi(V_1)$ given $U_1 = u$, which is assumed to be bounded. Denote by $f_w(\cdot|x)$ the conditional density of Y given $X = x$ and $f(\cdot|u)$ the conditional density of V given $U = u$, when theses quantities exist. In what follows we suppose $\text{supp}(f) = \text{supp}(f_w) \subset \text{supp}(w)$.

To define an estimate for $m(\cdot)$ from the observed data Z_i, $i = 1, \ldots, n$, observe first that

$$f_w(x) = \frac{\kappa(x)\gamma(x)f(x)}{\mathbb{E}(\kappa(X)w(X,Y))} \quad \text{and} \quad f_w(y|x) = \frac{w(x,y)f(y|x)}{\int w(x,y)f(y|x)dy}, \tag{2.1}$$

where $\gamma(x) := \mathbb{E}(w(X,Y)|X = x) > 0$. It immediately follows, for $j = 1, 2$, that

$$\mathbb{E}_w\left(\phi(Y)^{j-1}w(X,Y)^{-1}|X = x\right) = \frac{(m(x))^{j-1}}{\gamma(x)}, \tag{2.2}$$

whenever these quantities exist. And, therefore

$$m(x) = \frac{\mathbb{E}_w\left(\phi(Y)w(X,Y)^{-1}|X = x\right)}{\mathbb{E}_w\left(w(X,Y)^{-1}|X = x\right)} := \frac{s(x)}{t(x)}. \tag{2.3}$$

Define the following Generalized smooth estimator for $m(\cdot)$ based on the data (X_i, Y_i)

$$m_n(x) := \frac{\sum_{i=1}^n \phi(Y_i)w(X_i, Y_i)^{-1}Q_{x,v_n}(X_i)}{\sum_{i=1}^n w(X_i, Y_i)^{-1}Q_{x,v_n}(X_i)}, \tag{2.4}$$

when the denominator is not 0. Here v_n $(0 < v_n < 1)$ is the bandwidth parameter satisfying $v_n \to 0$ and $nv_n \to \infty$ as $n \to \infty$, and $Q_{x,v_n}(\cdot)$ is a density function with mean x and variance $(xv_n)^2$. The usual kernel estimator may be obtained is a special case of (2.4) by taking $Q_{v_n}(\frac{t}{x}) = \frac{1}{xv_n}K\left(\frac{t-x}{xv_n}\right)$, where $K(\cdot)$ is a density function with mean zero and variance 1.

Note however that this estimator may note be provide a consistently estimate of $m(0)$ (Chaubey $al.$, 2010). To alleviate this situation we consider the following perturbed version

$$r_n(x) := m_n(x + \epsilon) = \frac{\sum_{i=1}^n \phi(Y_i)w(X_i, Y_i)^{-1}\Delta_i(x + \epsilon)}{\sum_{i=1}^n w(X_i, Y_i)^{-1}\Delta_i(x + \epsilon)} := \frac{r_{n,2}(x)}{r_{n,1}(x)}, \quad x \geq 0, \tag{2.5}$$

where $\epsilon := \epsilon_n$ is a positive real number that goes to 0 at an appropriate rate as $n \to \infty$,

$$r_{n,j}(x) := n^{-1}\sum_{i=1}^n \phi(Y_i)^{j-1}w(X_i, Y_i)^{-1}\Delta_i(x + \epsilon), \quad \text{for} \quad j = 1, 2, \tag{2.6}$$

4

and $\Delta_i(x + \epsilon_n) = Q_{x+\epsilon_n, v_n}(X_i)$ with

$$Q_{x+\epsilon_n, v_n}(t) = \frac{1}{\beta_{x+\epsilon_n}^{\alpha_n} \, \Gamma(\alpha_n)} t^{\alpha_n - 1} \, e^{-\alpha_n t/(x+\epsilon_n)}, \text{ where } \alpha_n = 1/v_n^2, \ \beta_{x+\epsilon_n} = v_n^2(x + \epsilon_n) \quad (2.7)$$

is a gamma density with mean $x + \epsilon$ and variance $(v_n(x+\epsilon))^2$. This estimator is nonnegative and naturally asymmetric to cope with discontinuity at $t = 0$. The choose of the gamma density is motivated by the fact that the biased length data are typically nonnegative.

Notice that the choose of the function $\phi = I_{[0, \ t]}, \quad t \in \mathbb{R}^+$, where I_A stands for the indicator function of the set A, permits to construct an estimate of the conditional distribution function $\mathbb{F}(t|u)$.

2.1   Notation and assumptions

In order to state our results we introduce the following notations. Let $\mathcal{F}_i$ be the σ-field generated by $((X_1, Y_1), \ldots, (X_i, Y_i))$ and $\mathcal{G}_i$ that generated by $((X_1, Y_1), \ldots, (X_i, Y_i), X_{i+1})$. For $i \in \mathbb{N}$, let $f_{w,i}(\cdot) := f_w(\cdot|\mathcal{F}_{i-1})$ be the conditional density of X_i given $\mathcal{F}_{i-1}$. Let $\mathcal{C}_0(\mathbb{R})$ be the space of continuous functions going to zero at infinity and $\| \cdot \|$ be the sup norm. From now on, the notation $\xrightarrow{\mathcal{D}}$ stands for the convergence in distribution of random variables and $\xrightarrow{\mathbb{P}}$ the convergence in probability. Denote by $o_{\text{a.s.}}(u)$ a random function l such that $l(u)/u$ converges to zero almost surely as $u \to 0$. Similarly, define $O_{\text{a.s.}}(u)$ as a random function l such that $l(u)/u$ is almost surely bounded.

Our results are stated under the following assumptions, which are gathered here for easy reference:

(A1) $v_n \to 0$, $\epsilon \to 0$, $nv_n \to \infty$ and $nv_n\epsilon \to \infty$ as $n \to \infty$.

(A2) For all $i \in \mathbb{N}$, $f_w(\cdot) \in \mathcal{C}_0(\mathbb{R})$ and $f_{w,i}(\cdot) \in \mathcal{C}_0(\mathbb{R})$.

(A3) The sequence $\{n^{-1} \sum_{i=1}^n f_{w,i}(u)\}$ converges uniformly in u to $f_w(u)$ almost surely (a.s.) as $n \to \infty$.

(A4) For $j = 0, 1$

  i) $\mathbb{E}_w\left(|\frac{\phi(Y)^j}{w(X,Y)}|\right) < \infty$ and $\mathbb{E}_w\left(\frac{1}{w(X,Y)^2}\right) < \infty$.

  ii) The function $\tilde{s}(u) = \mathbb{E}_w\left(w(X,Y)^{-2}|X = u\right)$ is continuous bounded.

(A5) The conditional mean of the quantities $\phi(Y_i)$, $\phi^j(Y_i)w^{-1}(X_i, Y_i)$ $(j = 0, 1)$ and $\left(\frac{\phi(Y_i) - m(x)}{w(X_i, Y_i)}\right)^2$ given the sigma-field $\mathcal{G}_{i-1}$ exists and only depends on X_i, i.e., for any $i \geq 1$,

  (i) $\mathbb{E}\left(\phi(Y_i) \mid \mathcal{G}_{i-1}\right) = m(X_i)$ a.s.

5

(ii) $\mathbb{E}_w\left(w(X_i, Y_i)^{-1} \mid \mathcal{G}_{i-1}\right) = \mathbb{E}_w\left(w(X_i, Y_i)^{-1} \mid X_i\right) = t(X_i)$ a.s.

(iii) $\mathbb{E}_w\left(\phi(Y_i)w(X_i, Y_i)^{-1} \mid \mathcal{G}_{i-1}\right) = \mathbb{E}_w\left(\phi(Y_i)w(X_i, Y_i)^{-1} \mid X_i\right) = s(X_i)$ a.s.

(iv)

$$\mathbb{E}_w\left[\left(\frac{\phi(Y_i) - m(x)}{w(X_i, Y_i)}\right)^2 \mid \mathcal{G}_{i-1}\right] = \mathbb{E}_w\left[\left(\frac{\phi(Y_i) - m(x)}{w(X_i, Y_i)}\right)^2 \mid X_i\right] := g(X_i)$$ a.s.

(A6) There exists some $\delta > 2$ such that $\mathbb{E}_w\left(\left|\frac{\phi(Y)-m(x)}{w(X,Y)}\right|^{2+\delta}\right) < \infty$ and for any fixed x, the

function $g(u) := \mathbb{E}_w\left(\left|\frac{\phi(Y)-m(x)}{w(X,Y)}\right|^{2+\delta} \mid X = u\right)$ is continuous bounded .

(A7) i) The function $f_w(\cdot)$ admits a bounded derivative.

   ii) The function $m(\cdot)$ has bounded derivatives up to order two.

Condition (A1) and (A7) are very common in the framework of regression estimation, while (A1) and (A2) involve the ergodic nature of the data and are related to the application of the ergodic theorem. (A4) is a weaker condition than those proposed elsewhere in the literature. Condition (A5) is satisfied, for instance, by letting $Y_i = X_{i+1}$ with $\{X_i\}$ being a Markov process. It is also satisfied when we consider the heteroscedastic regression models

$$w(X_i, Y_i)^{-j}\phi^k(Y_i) = h_{k,j}(X_i) + \vartheta(X_i)\epsilon_i, \quad j = 1, 2, \qquad k = 0, 1, \tag{2.8}$$

where,

$$\vartheta^2(x) = \mathrm{Var}_w\left(\frac{\phi(Y)}{w(X, Y)} \mid X = x\right) \text{ a.s}, \quad h_{1,1}(x) = s(x),$$

and

$$h_{k,2}(x) = \mathbb{E}_w\left(\frac{\phi(Y)^k}{w(X, Y)^2} \mid X = x\right) \text{ a.s}.$$

The ϵ_i's are martingale difference with respect to the sigma-field $\mathcal{A}_i := \sigma\left((\epsilon_1, X_1), \ldots, (\epsilon_i, X_i), X_{i+1}\right)$, such that $\mathbb{E}_w(\epsilon^2 \mid \mathcal{A}_{i-1}) = 1$ a.s.

We will check the condition A(5)-(iv) only. We have $\mathcal{G}_i \equiv \mathcal{A}_i$, therefore

$$\begin{aligned}
\mathbb{E}_w\left(\left[\frac{\phi(Y_i) - m(x)}{w(X_i, Y_i)}\right]^2 \mid \mathcal{G}_{i-1}\right) &= \mathbb{E}\left(\left(\frac{\phi(Y_i)}{w(X_i, Y_i)}\right)^2 \mid \mathcal{G}_{i-1}\right) \\
&\quad - 2m(x)\mathbb{E}\left(\frac{\phi(Y_i)}{w^2(X_i, Y_i)} \mid \mathcal{G}_{i-1}\right) + m^2(x)\mathbb{E}\left(\frac{1}{w^2(X_i, Y_i)} \mid \mathcal{G}_{i-1}\right) \\
&= s^2(X_i) + \vartheta^2(X_i) - 2m(x)h_{1.2}(X_i) + m^2(x)h_{0,2}(X_i),
\end{aligned}$$

which is a function of X_i.

3 Main Results

From now on set, for $x \in \mathbb{R}^+$,

$$\overline{r}_{n,j}(x) := n^{-1} \sum_{i=1}^{n} \mathbb{E}_w \left[\phi(Y_i)^{j-1} w(X_i, Y_i)^{-1} \Delta_i(x + \epsilon) \,|\mathcal{F}_{i-1}\right], \quad \text{for} \quad j = 1, 2, \qquad (3.1)$$

where $\mathbb{E}_w(X \mid \mathcal{F})$ is the conditional expectation of X given the sigma-field $\mathcal{F}$. Define the conditional bias of the regression estimator $r_n(x)$ as

$$B_n(x) := C_n(x) - m(x), \quad \text{where} \quad C_n(x) := \frac{\overline{r}_{n,2}(x)}{\overline{r}_{n,1}(x)}. \qquad (3.2)$$

3.1 Consistency with rate

The following theorem gives the consistency in probability with rate of the estimate $r_n(x)$ at interior and boundaries of the support.

Theorem 1 . *In addition to conditions (A1)-(A6) and (A7), assume that*

$$\frac{\max(v_n^{5/2}, \epsilon^{5/4})}{v_n(x + \epsilon)\sqrt{n \log \log n}} \longrightarrow 0, \quad \frac{n \max(v_n^5, \epsilon^{5/2})}{\log \log n} \longrightarrow 0, \quad \frac{\max(v_n, \epsilon^{1/2})}{(x + \epsilon) v_n \log \log n} \longrightarrow 0 \qquad (3.3)$$

as $n \to \infty$. *Then for,*

$$\frac{n \max(v_n, \epsilon^{1/2})}{\log \log n} \to \infty \qquad (3.4)$$

we have for any $x \geq 0$

$$\sqrt{\frac{n \max(v_n, \epsilon^{1/2})}{\log \log n}} \, (r_n(x) - m(x)) \xrightarrow{\mathbb{P}} 0 \qquad as \quad n \to \infty.$$

3.2 Asymptotic Normality

Before state our result we introduce further notation. Let, for any $x \in \mathbb{R}^+$, whenever $f(x) > 0$

$$\psi(x) := \mathbb{E} \left(\frac{(\phi(Y) - m(x))^2}{w(X, Y)} \,\Big|\, X = x \right) \quad \text{and} \quad \sigma^2(x) := \frac{\mu \psi(x)}{2\sqrt{\pi}\, x\, \kappa(x) f(x)}. \qquad (3.5)$$

Theorem below gives the asymptotic normality of the estimator $r_n(x)$ at interior and boundaries of the support as well as the form of its asymptotic variance in both cases.

7

Theorem 2 . *Assume that conditions (A1)-(A6) and (A7)(i) hold.*

(i) We have, for any fixed $x > 0$ such that $f(x) > 0$

$$\sqrt{n\bar{v}_n}(r_n(x) - m(x) - B_n(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \quad \sigma^2(x)\right).$$

(ii) Suppose in addition that the condition (A7)(ii) is satisfied and

$$\sqrt{n\bar{v}_n} \max(v_n^2, \epsilon) \longrightarrow 0 \qquad as \quad n \to \infty. \tag{3.6}$$

Then, we have

$$\sqrt{n\bar{v}_n}(r_n(x) - m(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \quad \sigma^2(x)\right).$$

(iii) Assume, moreover that

$$\sqrt{n\bar{v}_n\epsilon} \max(v_n^2, \epsilon) \longrightarrow 0 \quad as \quad n \longrightarrow \infty. \tag{3.7}$$

Then we have, whenever $f(0) > 0$

$$\sqrt{n\bar{v}_n\epsilon_n}(r_n(0) - m(0)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \quad \rho^2(0)\right),$$

where

$$\rho^2(0) := \frac{\mu\phi(0)}{2\sqrt{\pi}\,\kappa(0)f(0)}. \tag{3.8}$$

Remark 1. The choose of $w(x, y) = \phi(y) = y$ correspond to the case of length biased data frequently studied in literature. In this case, the functions $\psi(x)$, $\gamma(x)$ and the mean μ being

$$\psi_b(x) = m(x)\left[m(x)\mathbb{E}(Y^{-1}|X = x) - 1\right], \quad \gamma_b(x) = m(x) \quad \text{and} \quad \mu_b = \mathbb{E}(Y), \tag{3.9}$$

and the estimator defined in (2.4) takes the form

$$r_{n,b}(x) = \frac{\sum_{i=1}^n \Delta_i(x + \epsilon)}{\sum_{i=1}^n Y_i^{-1}\Delta_i(x + \epsilon)}, \tag{3.10}$$

which may be interpreted as a local harmonic mean for estimating the unconditional mean of length biased data. In this case, the limiting variance functions given in (3.5) and (3.8) become

$$\sigma_b^2(x) := \frac{\mu_b\psi_b(x)}{2\sqrt{\pi}\,x\,\kappa(x)f(x)} \quad \text{and} \quad \rho_b^2(0) := \frac{\mu_b\psi_b(0)}{2\sqrt{\pi}\,\kappa(0)f(0)}.$$

8

3.3 Asymptotic Bias

The following results give the asymptotic order of the conditional bias $B_n(x)$ and specified the convergence rate of the errors. Before stat our result, assume the following additional conditions

(B1) There exist a constant $c > 0$ such that $w(x, y) \geq c$.

(B2) The functions f_w, m, s and t admit derivatives up to order 3 and their third derivatives are bounded.

Proposition 1 *Under assumptions (A1), (A2), (A3), (A4)(ii), (A5)(ii), (B1) and (B2) and the condition*

$$\lim_{n \to \infty} \frac{(x + \epsilon) n v_n}{\log n} = \infty. \tag{3.11}$$

We have for any $x \geq 0$

$$
\begin{aligned}
B_n(x) = C_n(x) - m(x) &= m'(x)\epsilon + \left(m''(x) + m'(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right) A(v, \epsilon) + o\left(A(v, \epsilon) \right) \\
&+ O_{a.s.}\left(\sqrt{\frac{\max(v_n^4, \epsilon^2) \log n}{n(x + \epsilon) v_n}} \right) + O_{a.s.}\left(\frac{\max(v_n^2, \epsilon) \log n}{n(x + \epsilon) v_n} \right),
\end{aligned}
$$

where

$$A(v, \epsilon) := x^2 v_n^2 + 2x\epsilon v_n^2 + \epsilon^2(v_n^2 + 1). \tag{3.12}$$

Remark. The condition (B1) permits to apply an exponential inequality for bounded martingale difference random variable to obtain the convergence rate of the errors. It can be relaxed easily by applying a more general exponential inequality for unbounded martingale difference random variables given in Laïb and Louani (2010).

3.4 Mean square error (MSE)

The $\mathrm{MSE}(r_n(x)) := \mathbb{E}_w(r_n(x) - r(x))^2$, which measure the average of the square of the error, is a quantity used to quantify the amount by which the estimator $r_n(x)$ differs from the true value $m(x)$ that being estimate. It may used for comparative purposes of two or more estimators.

The following Proposition gives the MSE of the estimator $r_n(x)$ in the interior as well as in the boundary of the interval, which allows us to determine the optimal (in the sense of minimizing the quantity (MSE)) rates of convergence of $v_n \to 0$ and $\epsilon_n \to 0$.

Proposition 2 *Assume that Conditions (A1)-(A3), (A5)(ii)-(iv) and (B2) are satisfied. Then we have for any $x \geq 0$ whenever $fx) > 0$ that*

$$
\begin{aligned}
MSE(r_n(x) \quad \simeq \quad & (m'(x)\epsilon)^2 + \epsilon A(v_n, \epsilon) m'(x) \left\{ m''(x) + m'(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right\} \\
+ \quad & \frac{1}{4} A^2(v, \epsilon) \left\{ m''(x) + m'(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right\}^2 \\
+ \quad & \frac{\mu \psi(x)}{2\sqrt{\pi}(x+\epsilon) n v_n \; \kappa(x) f(x)} \\
+ \quad & \frac{\mu}{2\sqrt{\pi} \; \kappa(x) f(x)} \left\{ \psi'(x) + \left(\frac{1}{\gamma(x)} - 1 \right) \frac{\gamma'(x)}{\gamma(x)} \psi(x) + \right. \\
& \left. \psi(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right\} \left\{ \frac{1}{nv} - \frac{x}{(x+\epsilon)nv} \right\}.
\end{aligned}
$$

The following Corollary gives the form of the MSE for $x > 0$ as well as for $x = 0$.

Corollary 1 *i) If $x > 0$, we have*

$$
\begin{aligned}
MSE(r_n(x) \quad \simeq \quad & \frac{\sigma^2(x)}{nv} + (m'(x)\epsilon)^2 + (x+2\epsilon)\epsilon x v^2 m'(x) \left\{ m''(x) + m'(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right\} \\
+ \quad & \frac{x^3 v^4}{4}(x + 4\epsilon) \left\{ m''(x) + m'(x) \left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)} \right] \right\}^2.
\end{aligned}
$$

ii) If $x = 0$, we have

$$
\begin{aligned}
MSE(r_n(0) \quad \simeq \quad & \frac{\rho^2(0)}{nv\epsilon} + (m'(0)\epsilon)^2 \\
+ \quad & \frac{\mu}{2\sqrt{\pi} \; nv \; \kappa(0) f(0)} \left[\psi'(0) + \psi(0) \left[\left(\frac{1}{\gamma(0)} - 1 \right) \frac{\gamma'(0)}{\gamma(0)} + \frac{\kappa'(0)}{\kappa(0)} + \frac{f'(0)}{f(0)} \right] \right].
\end{aligned}
$$

Remarks.

1) Corollary 1 shows that, when $x > 0$, the optimal choice of ϵ_n is 0 which gives the optimal choice of v_n to be $v_n = O(n^{-1/5})$ and the optimal order of MSE is then $O(n^{-4/5})$. Note also that if we take $\epsilon_n = O(v_n^2)$ we obtain the same optimum. When $x = 0$, the optimal choice of ϵ_n is $O((nv_n)^{-1/})$, however there is no optimal choice for $v_n > 0$.

2) If we choose $\kappa(x) = w(x, y) = 1$, then $\mu = 1$, $t(x) = 1$ and $f_w \equiv f$, which corresponds to the non sampled data. In this case, the MSE is the same of that obtained in Chaubey *et al.* (2010). If in addition that $\epsilon = 0$, we get the classical MSE for the kernel Gamma regression, that is

$$
MSE(x) = \frac{\sigma^2(x)}{nv} + \frac{x^4 v^4}{4} \left\{ m''(x) + m'(x) \frac{f'(x)}{f(x)} \right\}^2.
$$

3) If we take $k(x) = 1$, $w(x,y) = \phi(y) = y$, which corresponds to the usual length biased data, then the functions $\psi(x)$, $\gamma(x)$ and the mean μ take the forme as in (3.9) and the MSE being, whenever $x > 0$,

$$
\begin{aligned}
MSE(r_{b,n}(x)) \simeq{} & \frac{\sigma_b^2(x)}{nv} + (m'(x)\epsilon)^2 + (x + 2\epsilon)\epsilon x v^2 m'(x)\left\{m''(x) + m'(x)\frac{f'(x)}{f(x)}\right\} \\
& + \frac{x^3 v^4}{4}(x + 4\epsilon)\left\{m''(x) + m'(x)\frac{f'(x)}{f(x)}\right\}^2 .
\end{aligned}
$$

And if $x = 0$, it being

$$
\begin{aligned}
MSE(r_{n,b}(0)) \simeq{} & \frac{\sigma_b^2(0)}{nv\epsilon} + (m'(0)\epsilon)^2 \\
& + \frac{\mu_b}{2\sqrt{\pi}\, nv\, f(0)}\left[\psi_b'(0) + \psi_b(0)\left[\left(\frac{1}{m(0)} - 1\right)\frac{m'(0)}{m(0)} + \frac{f'(0)}{f(0)}\right]\right] .
\end{aligned}
$$

4) Compare our results to that obtained in Sköld (1999), one may observe that the bias of the classical kernel estimate given in Theorem 1 of Sköld (1999) is affected by the scale parameter $\kappa(x)$, which is comparable to our result given Corollary 1. However, the bias for the local linear given in Theorem 2 in Sköld (1999) is free from $\kappa(x)$, this is due to the power of the local linear smoothing, which makes the leading bias term free of the first derivative of m and f. Note however that, if we choose, for example, the scale parameter $\kappa(x) = c/f(x)$ for some positive constant c, then the model (1.1) may be written as

$$
f_w(x,y) = \frac{w(x,y)f(y|x)}{\mu}, \quad \text{where} \quad \mu = \mathbb{E}(f^{-1}(X)w(X,Y)) = \int y f(y|x)dxdy. \quad (3.13)
$$

In this case the variance $\sigma^2(x)$ being independent of the marginal density f and the expression $\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)}$ involved in the bias term is 0. In this case the bias square term given in Corollary 1 is comparable to the local linear estimate, which is small than the usual case when, for example, the function m is close to linear. Note moreover that despite of the nice properties of the local linear approach, this method uses a fixed symmetric kernel with support compact with mean 0 and variance 1, thus its support does not matches the support of the regression curve. It has also a drawback as the variance is unbounded in finite sample as point out by Seifert and Gasser (1996). While the family of Gamma kernels used here posses the following properties: (a) It provide an asymmetric density with mean x and variance $(xv_n)^2$. Here the quantity $h = h(x) = xv_n$ may be interpreted as the bandwidth parameter, that depends on x,), (b) It has varying shapes and varying degree of smoothness. (c) It's support matches the support of the regression curve, which leads to increase the effective sample size, and therefore the finite sample variance of the estimator may be reduced.

The perturbed version of the estimator introduced here appears a very useful new idea to deal with boundary bias in the case of nonnegative data, which also avoids the complication

11

of some of the rigorous boundary correction methods in the literature. The estimator defined here may be generalized easily to higher dimension. Note finally that, the numerical study has carried out recently in Chaubey *et al.* (2010) showed that our methods is competitive to the local linear smoothing. We would like to investigate such properties for the length biased case in this paper.

4 Simulation Studies

4.1 Selection of parameters

The method for selecting parameter we use here is given by Hengartner and Wegkamp (2002). We randomly select m pairs of samples from the whole data set with size $m + l$. Then we use m pairs of samples $\{Y_i, X_i\}_{i=1}^m$ as testing sample and the rest l pairs of samples to build regression estimator, say $m_{l,h(x)}$ where h represents smoothing parameters and might be a vector. To find optimal solutions of parameter, we minimize quadratic risk function

$$Q_m(h) = \frac{1}{m} \sum_{i=1}^m \left[Y_i - m_{l,h}(X_i) \right]^2. \tag{4.1}$$

To reduce variability caused by random partition of the data, Hengartner and Wegkamp (2002) suggested dividing the same data set 21 times and using the median of the obtained 21 samples of optimal solution as the final selection of parameter. They also suggested that $m = \lceil n^\beta \rceil$ ($\beta = 0.75, 0.8, 0.85, 0.9, 0.95$). Their simulation showed that under the choice $\beta = 0.85$ the results are more satisfactory in most cases. For sample size $n = 200$, we take $m = 90$. Under the chosen parameters, we obtain 1000 samples of $m_n^i(x)(i = 1, \ldots, 1000)$ at a grid of interval. Based on those samples, we compute average bias

$$AB(m_n(x), m(x)) = \frac{1}{1000} \sum_{i=1}^{1000} \left[m_n^i(x) - m(x) \right],$$

sample variance

$$SV(m_n(x)) = \frac{1}{1000} \sum_{i=1}^{1000} \left[m_n^i(x) - \bar{m}_n(x) \right]^2$$

where $\bar{m}_n(x) = \frac{1}{1000} \sum_{i=1}^{1000} m_n^i(x)$ and averaged squared error

$$ASE(m_n(x), m(x)) = \frac{1}{1000} \sum_{i=1}^{1000} \left[m_n^i(x) - m(x) \right]^2$$

for numerical comparison.

4.2   i.i.d case

We generate the length biased data for $f_w(y/x)$ given by a $Gamma(2, \ m(x)/2)$ density and $f(y/x)$ is given by $\Gamma(3, \ m(x)/2)$ density. We consider two examples here for $m(x)$.

- This example is taken from Sköld (1999). Here we consider X to be uniformly distributed on $[0, 1]$.

$$m(x) = \frac{\cos(3x)}{2} + 3x^4 - 2.2x^8.$$

- This example is taken from Cristóbal and Alcalà (2001), where $m(x)$ is given by

$$m(x) = 15 + 2x + 50[\exp(-(x - 45)^2/16) - \exp(-(x - 55)^2/16)]$$

and we take $X \hookrightarrow U[20, \ 80]$.

Figures 1 and 3 display Average Bias, Sample Variance and Average Squared Error for generalized kernel weighted (GKW) non-parametric estimators with asymmetric and symmetric kernels along with local linear estimators with asymmetric and symmetric kernels. Figures 2 and 4 present the graph of different estimators for a typical sample. We draw similar conclusions as in Chaubey *et al.* (2010) regarding the new estimator for the case of non-weighted data. Namely, the generalized asymmetric kernel estimator handles the boundary bias well and presents it self a good competitor to the local linear estimator with respect to bias and mean squared error.
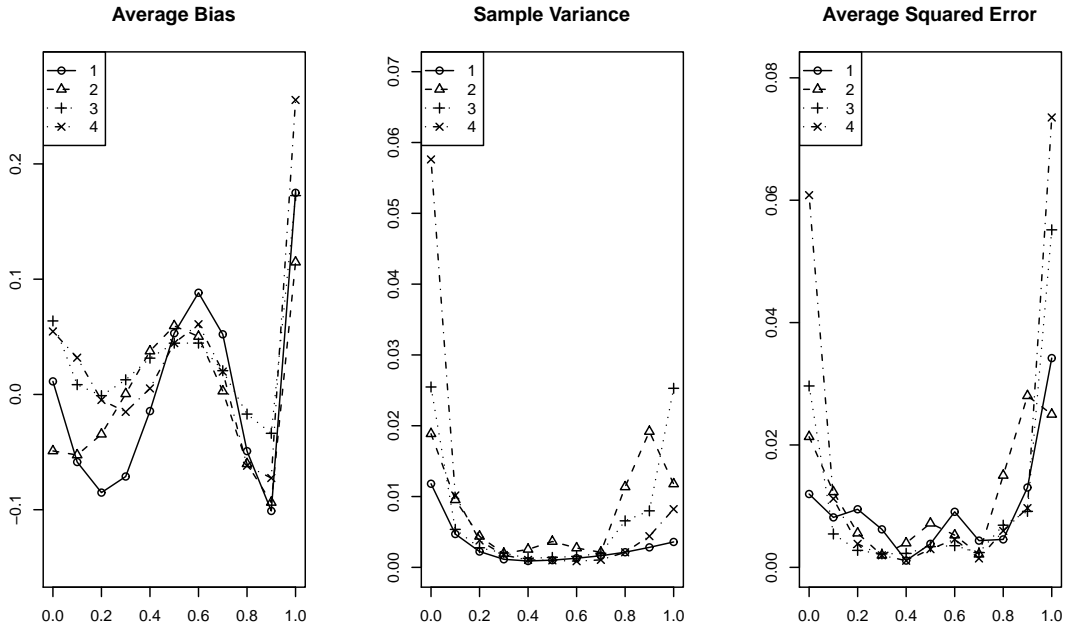
Figure 1: 1: GKW estimator with asymmetric kernel; 2: GKW estimator with symmetric kernel ; 3: LLE with symmetric kernel; 4: LLE with asymmetric kernel.
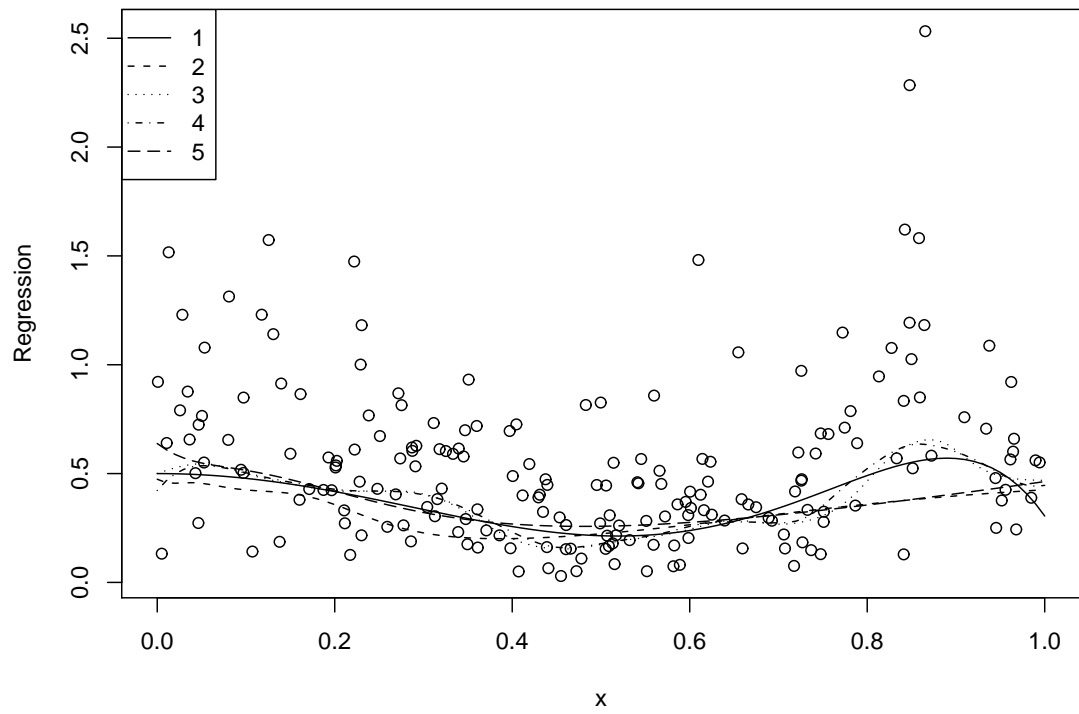
Figure 2: 1: True regression function; 2: GKW estimator with asymmetric kernel; 3: GKW estimator with symmetric kernel ; 4: LLE with symmetric kernel; 5: LLE with asymmetric kernel.
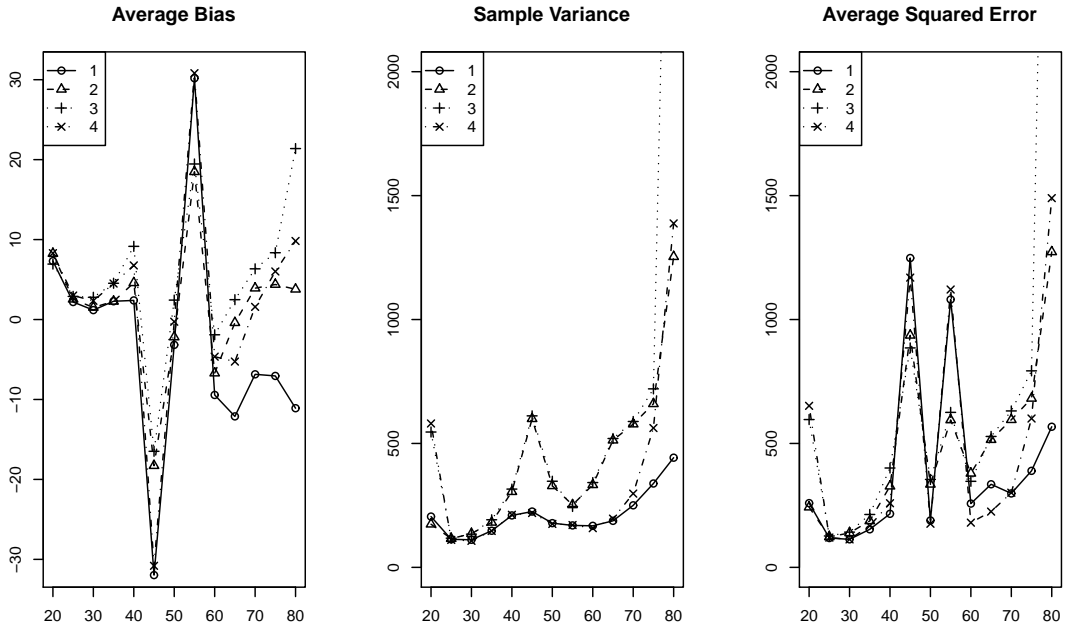
Figure 3: 1: GKW estimator with asymmetric kernel; 2: GKW estimator with symmetric kernel ; 3: LLE with symmetric kernel; 4: LLE with asymmetric kernel.
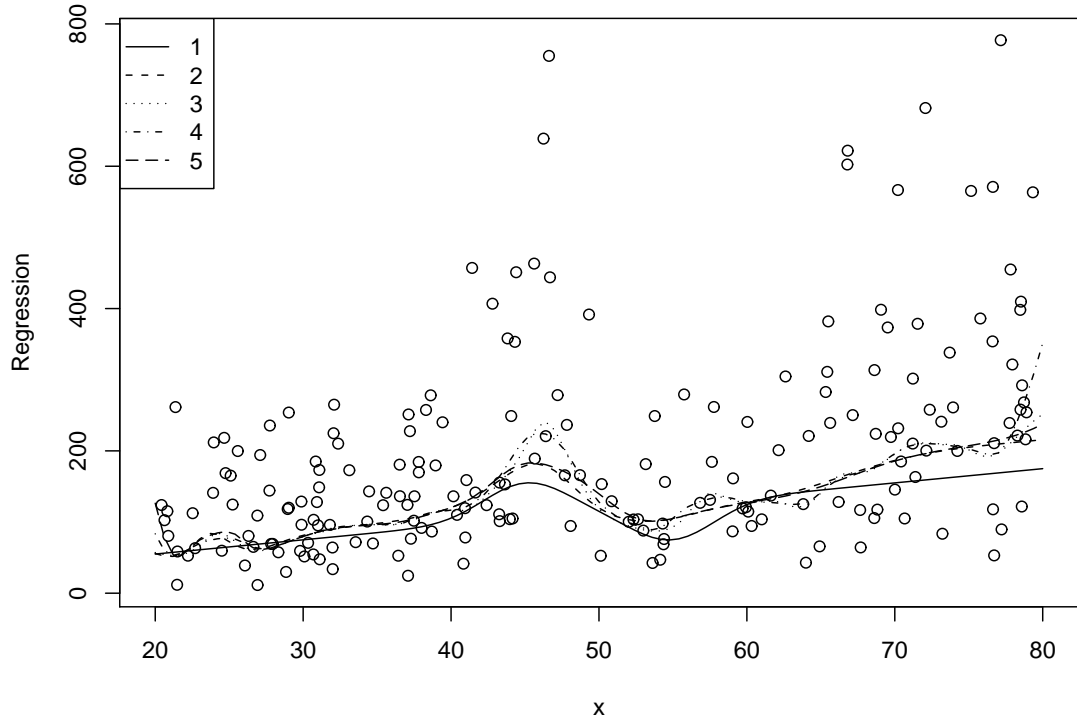
Figure 4: 1: True regression function; 2: GKW estimator with asymmetric kernel; 3: GKW estimator with symmetric kernel ; 4: LLE with symmetric kernel; 5: LLE with asymmetric kernel.

4.3 dependent data case

In this case we consider the two regressions considered above, however,

- In the first example, X_i are generated by

$$X_i = 0.5X_{i-1} + \left(\sqrt{0.2 + 0.1X_{i-1}^2} \right) \eta_i$$

with $X_0 \sim [0, \ 1]$ and $\eta_i \sim [1, 1/(2\sqrt{0.3}]$. Y_i^w are generated by

$$Y_i^w = m(X_i)(1 + 0.35\epsilon_i)$$

where $\epsilon_i \sim [-\sqrt{3}, \sqrt{3}]$

- In the second example, X_i are generated by

17

- In the first example, X_i are generated by

$$X_i = 0.5X_{i-1} + \left(\sqrt{96 + 0.01X_{i-1}^2}\right)\eta_i$$

with $X_0 \sim [20, 80]$ and $\eta_i \sim [1, \sqrt{10}]$. Y_i^w are generated by

$$Y_i^w = m(X_i)(1 + 0.35\epsilon_i)$$

where $\epsilon_i \sim [-\sqrt{3}, \sqrt{3}]$

Figures 5 and 7 display Average Bias, Sample Variance and Average Squared Error for generalized kernel weighted (GKW) non-parametric estimators with asymmetric and symmetric kernels along with local linear (LL) estimators with asymmetric and symmetric kernels. Figures 6 and 8 present the graph of different estimators for a typical sample. We draw similar conclusions here as in the case of *i.i.d.* data.
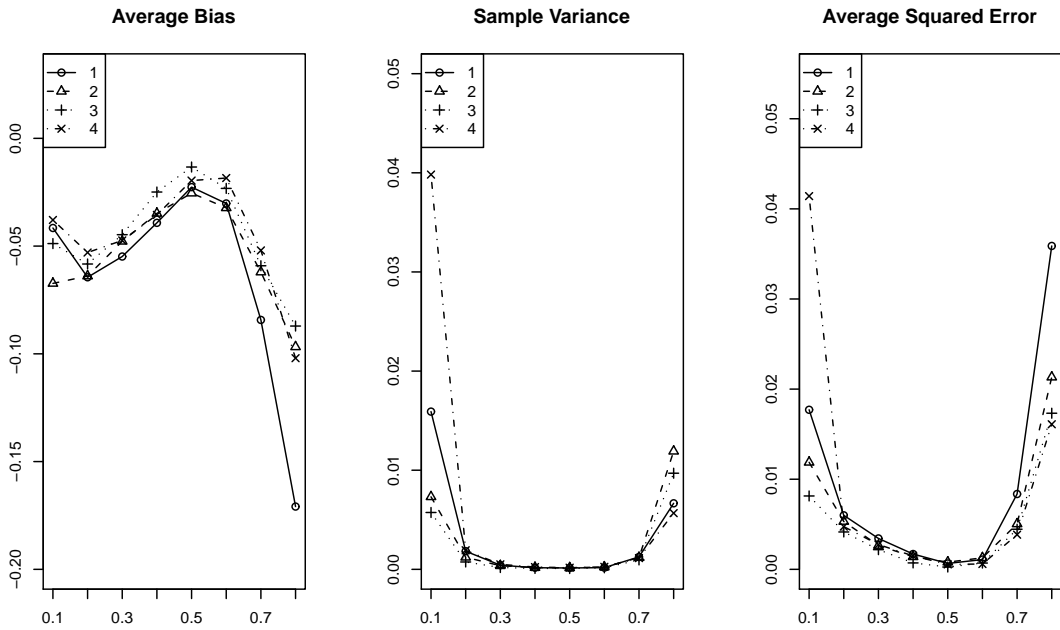

Figure 5: 1: GKW estimator with asymmetric kernel; 2: GKW estimator with symmetric kernel ; 3: LLE with symmetric kernel; 4: LLE with asymmetric kernel.
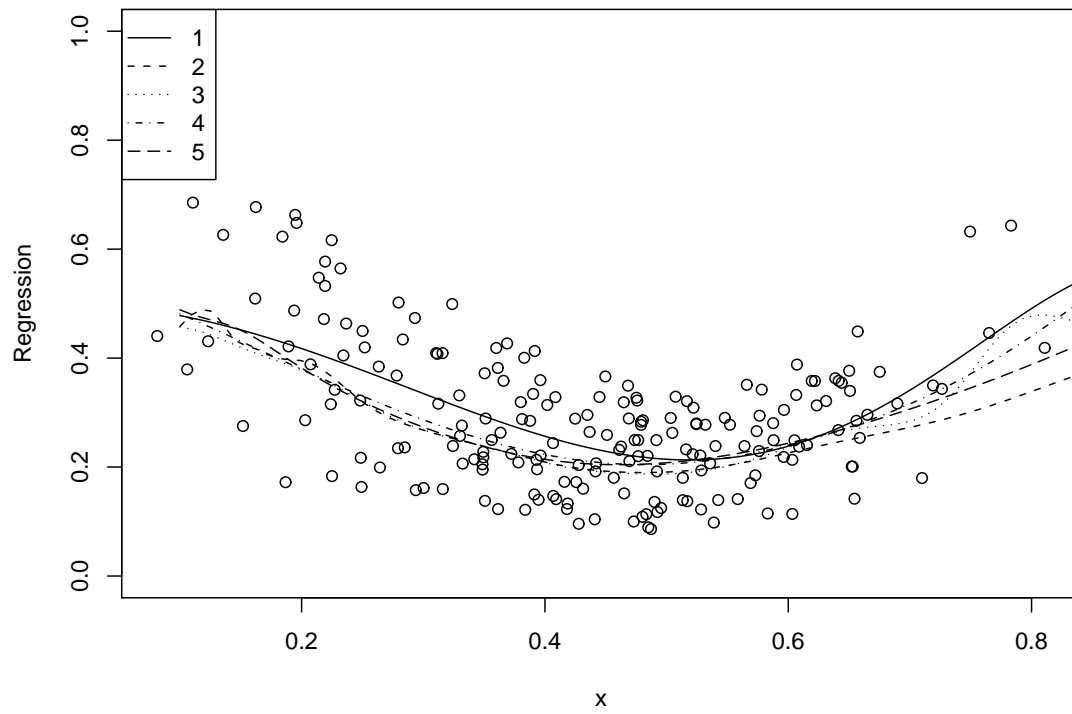
Figure 6: 1: True regression function; 2: GKW estimator with asymmetric kernel; 3: GKW estimator with symmetric kernel ; 4: LLE with symmetric kernel; 5: LLE with asymmetric kernel.
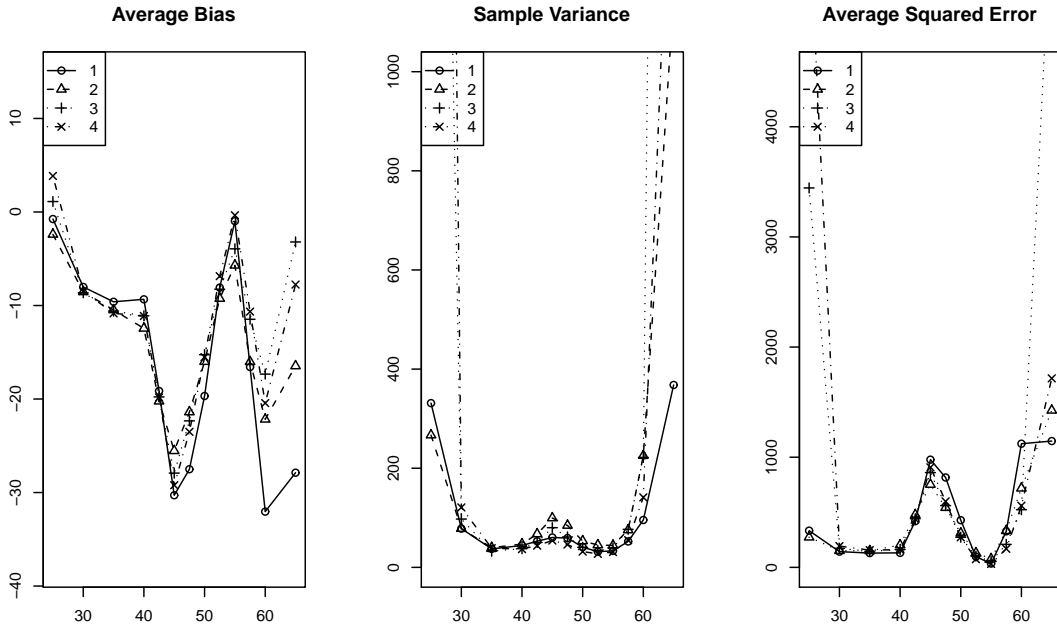
Figure 7: 1: GKW estimator with asymmetric kernel; 2: GKW estimator with symmetric kernel ; 3: LLE with symmetric kernel; 4: LLE with asymmetric kernel.
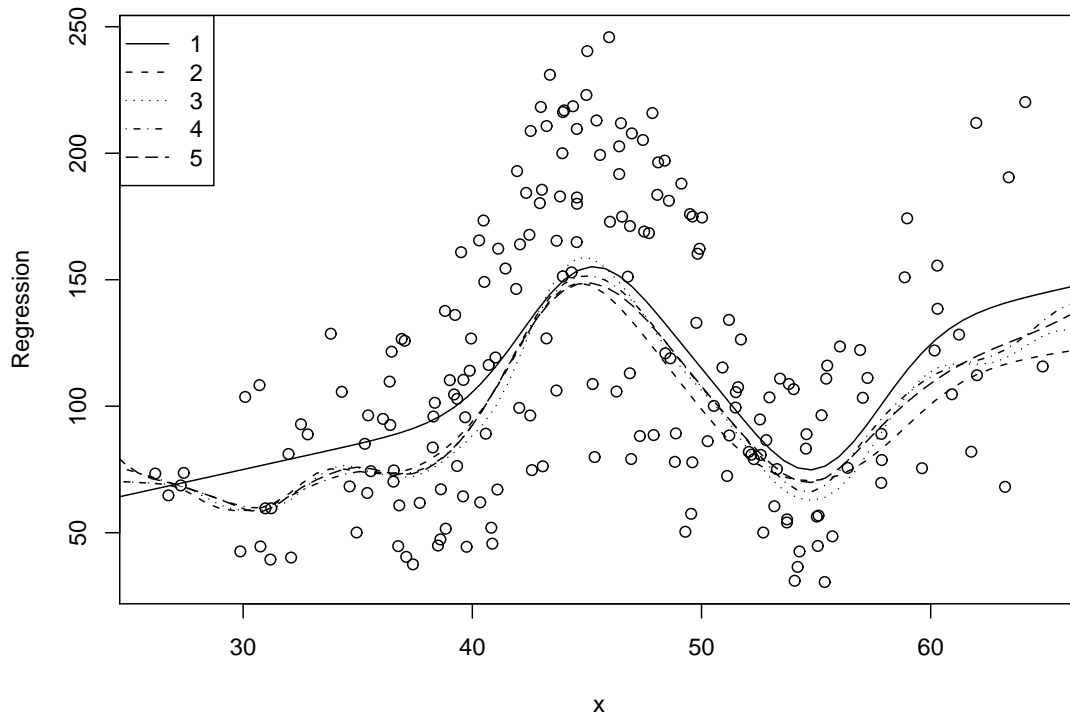
Figure 8: 1: True regression function; 2: GKW estimator with asymmetric kernel; 3: GKW estimator with symmetric kernel ; 4: LLE with symmetric kernel; 5: LLE with asymmetric kernel.

5   Proofs

In order to establish our results, introduce some additional notations. Set, for $x \in \mathbb{R}^+$,

$$S_n(x) := (r_{n,2}(x) - \bar{r}_{n,2}(x)) - m(x)(r_{n,1}(x) - \bar{r}_{n,1}(x)) \tag{5.1}$$

and

$$R_n(x) := -B_n(x)(r_{n,1}(x) - \bar{r}_{n,1}(x)). \tag{5.2}$$

Clearly, we have

$$r_n(x) - C_n(x) = \frac{S_n(x) + R_n(x)}{r_{n,1}(x)}. \tag{5.3}$$

The proof of Theorem 1 is split up into several lemmas establishing respectively the convergence in probability of $r_{n,1}(x)$ to $t(x)f_w(x)$, the fact that $R_n(x)$, suitably normalized, is

21

actually equal to $o_{\mathbb{P}}(1)$ and the asymptotic normality of $S_n(x)$. We start with some technical lemmas.

The following Lemma plays the same role as the classical Böchner's Lemma for the kernel estimate.

Lemma 1 *Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be any bounded continuous function and f be the density function of X that assumed to be bounded and continuous. Moreover assume that conditions (A1)-(A2) and (A3) hold true. Then we have, almost surely that*

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_w \left(\varphi(X_i)\Delta_i(x+\epsilon)|\mathcal{F}_{i-1}\right) \longrightarrow \varphi(x)f(x) \quad as \quad n \to \infty. \tag{5.4}$$

The convergence is uniformly in x whenever φ and f are uniform continuous functions.

Proof of Lemma 1. The proof is similar of that of the Proposition 1 in Chaubey *et al.* (2010). $\square$

Lemma 2 .

a) We have for any $p \geq 0$ and $m \geq 1$ that

$$\int_0^\infty t^p Q^m_{x+\epsilon,v_n}(t)dt = \frac{\left(\frac{1}{v_n^2(x+\epsilon)}\right)^{m/v_n^2}}{\left(\frac{m}{v_n^2(x+\epsilon)}\right)^{((m/v_n^2)+p+1-m)}} \cdot \frac{\Gamma(m/v_n^2 + p + 1 - m)}{\Gamma^m(1/v_n^2)}$$

$$\approx \frac{1}{\sqrt{m(2\pi)^{m-1}}}\frac{1}{v_n^{m-1}(x+\epsilon)^{m-p-1}}\frac{1}{\sqrt{1 - v_n^2\left(\frac{m-p-1}{m}\right)}}, \quad as \quad (\epsilon, v_n) \to (0,0). \tag{5.5}$$

b) For any fixed x and any $t \geq 0$, there exists $c > 0$ such that

$$Q_{x+\epsilon_n,v_n}(t) \leq \frac{c}{\sqrt{2\pi}(x+\epsilon_n)v_n}, \quad whenever \quad v_n \to 0. \tag{5.6}$$

Proof of Lemma 2. The proofs uses routine calculations and the Sterling Lemma. $\square$.

The following lemma describes the asymptotic behavior of the term $r_{n,1}(x)$.

Lemma 3 *Assume that hypotheses (A1)-(A4), (A5)(ii) are satisfied and the function $t(\cdot)$ is bounded. . Then we have, for any $x \geq 0$, that*

$$r_{n,1}(x) \xrightarrow{\mathbb{P}} t(x)f(x).$$

Proof of Lemma 3. Observe that

$$\hat{r}_{n,1}(x) = R_{1,n}(x) + R_{2,n}(x), \tag{5.7}$$

where

$$R_{1,n}(x) := n^{-1} \sum_{i=1}^{n} \left(w^{-1}(X_i, Y_i)\Delta_i(x) - \mathbb{E}_w[w^{-1}(X_i, Y_i)\Delta_i(x) \,|\mathcal{F}_{i-1}] \right),$$

$$R_{2,n}(x) := n^{-1} \sum_{i=1}^{n} \mathbb{E}_w[w^{-1}(X_i, Y_i)\Delta_i(x) \,|\mathcal{F}_{i-1}]. \tag{5.8}$$

Combined Lemma 1 with conditions (A4)-(ii) and (A5)-(ii), and using the property of the conditional expectation, one may see that $R_{n,2}(x)$ converges almost surely to $t(x)f(x)$ as n goes to infinity.

To handle the first term, observe that $R_{n,1}(x) = \sum_{i=1}^{n} L_{ni}(x)$ where $\{L_{ni}(x)\}$ is a triangular array of martingale differences with respect to the σ-field $\mathcal{F}_{i-1}$. Combined Burkholder with Jensen inequalities we obtain for any $\lambda > 0$ that there exists a constant $c > 0$ such that

$$\mathbb{P}(|R_{n,1}(x)| > \lambda) \leq c \frac{\mathbb{E}_w(\Delta_i^2(x)w(X_i, Y_i)^{-2})}{n\lambda^2}$$

$$= c \frac{\mathbb{E}_w\left(\Delta_i^2(x)\tilde{s}(X_i)\right)}{n\lambda^2}$$

One may write then in view of the statement (5.5) and the fact that $\tilde{s}(\cdot)$ is bounded

$$\mathbb{E}_w\left(\Delta_1^2(x)\tilde{S}(X_i)\right) = \int_0^\infty Q_{x+\epsilon}^2(t)\tilde{s}(t)f_w(t)dt < \sup_{t\in\mathbb{R}^+}(f(t)\tilde{s}(t)) \int_0^\infty Q_{x+\epsilon}^2(t)dt = O\left(\frac{1}{v_n(x+\epsilon)}\right).$$

It follows that $|R_{n,1}(x)| = o_{\mathbb{P}}(1)$ because of condition (A1), $nv_n \to \infty$ and $nv_n\epsilon_n \to \infty$ as $n \to \infty$. $\square$

Lemma 4 *i) In addition to (A1)-(A4)(i), (A5)(ii)-(iii), assume that the functions $s(\cdot)$ and $t(\cdot)$ are bounded continuous. Then, one has for any $x \geq 0$,*

$$B_n(x) = o_{a.s.}(1)$$

ii) If moreover condition (A4)(ii) is satisfied, one get for any $x \geq 0$

$$R_n(x) = o_{\mathbb{P}}\left(\frac{1}{nv_n(x+\epsilon)}\right). \tag{5.9}$$

iii) If in addition that the conditions (A7) are satisfied, then we have for any $x \geq 0$, that

$$B_n(x) = O_{\mathbb{P}}\left(\max(v_n^2, \epsilon)\right) \quad and \quad R_n(x) = O_{\mathbb{P}}\left(\frac{\max(v_n^2, \epsilon)}{nv_n(x+\epsilon)}\right). \tag{5.10}$$

23

Proof of Lemma 4. Recall that

$$B_n(x) = \frac{\bar{r}_{n,2}(x) - m(x)\bar{r}_{n,1}(x)}{\bar{r}_{n,1}(x)} := \frac{\tilde{B}_n(x)}{\bar{r}_{n,1}(x)}.$$

Making use of condition (A5)(ii) and Lemma 1, we obtain $\bar{r}_{n,1}(x)$ converges almost surely to $t(x)f_w(x)$. Thus, $\bar{r}_{n,1}(x) = O_{a.s}(1)$ since $t(x)f(x)$ is bounded.

To prove the first part of proposition, using conditions (A5)(ii)-(iii) and the fact that $s(x) = m(x)t(x)$, one can easily see that

$$\begin{aligned}
\tilde{B}_n(x) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_w[\Delta_i(x+\epsilon)(r(X_i) - m(x))(t(X_i) - t(x)) \,|\mathcal{F}_{i-1}] \\
&+ \frac{t(x)}{n}\sum_{i=1}^{n} \mathbb{E}_w[\Delta_i(x+\epsilon)(r(X_i) - m(x))|\mathcal{F}_{i-1}] =: \tilde{B}_{1,n}(x) + \tilde{B}_{2,n}(x). \quad (5.11)
\end{aligned}$$

It follows then by Lemma 1 that $\tilde{B}_{j,n}(x) = o(1)$ (j=1,2). This leads to $B_n(x) = o_{a.s.}(1)$.

To prove the second part of Lemma 4, observe

$$R_n(x) = -B_n(x)R_{1,n}(x),$$

where $R_{1,n}(x)$ is defined in (5.8). One can then obtained, whenever condition (A5)-(ii) is satisfied, that $R_n(x) = o(1)O_{\mathbb{P}}(\frac{1}{nv_n(x+\epsilon)})$, which completes the proof of part ii).

To give now an estimate of the rate of $\tilde{B}_{2,n}(x)$, observe that $Q_{x+\epsilon}(u) = \frac{1}{x+\epsilon}q(\frac{u}{x+\epsilon})$ where $q(\cdot)$ is gamma density with mean 1 and variance v_n. Using condition (A3), which permits to interchange the sum with the integral, we obtain for n large enough that

$$\begin{aligned}
\tilde{B}_{2,n}(x) &= \frac{t(x)}{n}\sum_{i=1}^{n}\int_0^{\infty} [r((x+\epsilon)s) - m(x)]\, q(s)f_{w,i}((x+\epsilon)s)ds \\
&= t(x)\int_0^{\infty} [r((x+\epsilon)s) - m(x)]\, q(s)\left[\frac{1}{n}\sum_{i=1}^{n} f_{w,i}((x+\epsilon)s)\right] ds \\
&\simeq t(x)\int_0^{\infty} [r((x+\epsilon)s) - m(x)]\, q(s)f_w((x+\epsilon)s)ds. \quad (5.12)
\end{aligned}$$

Making use of a Taylor expansion of the functions $r(\cdot)$ and $f_{w,i}(\cdot)$ around x, we obtain by conditions (A7) that $\tilde{B}_{2,n}(x) = O_{a.s.}(\max(v_n^2, \epsilon))$. Similarly, we have $\tilde{B}_{1,n}(x) = O_{a.s.}(\max(v_n^2, \epsilon))$. This completes the proof of part iii) and therefore that of Lemma 4. $\qquad\square$

The following Proposition establish the asymptotic normality of $S_n(x)$.

24

Proposition 3 *Assume conditions (A1)-(A3), (A5)(ii)-(iii)-(iv), (A6) and (A7)(i) are satisfied.*

i) We have, for any a given $x > 0$, that

$$\sqrt{nv_n}S_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \quad \tau^2(x)\right), \quad where \quad \tau^2(x) = \frac{g(x)f_w(x)}{2\sqrt{\pi}x}. \tag{5.13}$$

ii) Moreover we have, whenever $nv_n\epsilon \to \infty$ as $n \to \infty$, that

$$\sqrt{nv_n\epsilon}\,S_n(0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \quad \varrho^2(0)\right), \quad where \quad \varrho^2(0) = \frac{g(0)f_w(0)}{2\sqrt{\pi}}. \tag{5.14}$$

Proof of Proposition 3. Part i). Let

$$\eta_{ni} = \left(\frac{v_n}{n}\right)^{1/2}\left(\phi(Y_i) - m(x)\right)w(X_i, Y_i)^{-1}\Delta_i(x + \epsilon) \quad \text{and } \xi_{ni} = \eta_{ni} - \mathbb{E}_w\left(\eta_{ni}|\mathcal{F}_{i-1}\right).$$

Then

$$\sqrt{nv_n}S_n(x) = \sum_{i=1}^{n}\xi_{ni}, \qquad \text{where} \qquad \mathbb{E}_w\left(\xi_{ni}|\mathcal{F}_{i-1}\right) = 0 \text{ a.s.}$$

Thus, for any fixed $x \geq 0$, $\sqrt{nv_n}S_n(x)$ form a triangular array stationary martingale with respect the sigma field $\mathcal{F}_{i-1}$. To prove the statement (5.14), it suffices then to show that:

a) $\sum_{i=1}^{n}\mathbb{E}_w\left[\xi_{ni}^2|\mathcal{F}_{i-1}\right] \xrightarrow{\mathbb{P}} \tau^2(x)$

b) $nE\left[\xi_{ni}^2I_{[|\xi_{ni}|>\lambda]}\right] = o(1)$ for any $\lambda > 0$.

Using conditions (A1), (A2) and (A5)(ii)-(iii), and proceeding as in the proof of Proposition 1 in Chaubey *et al.* (2010), one can see that the statement a) will be proved if

$$\sum_{i=1}^{n}\mathbb{E}_w\left[\eta_{ni}^2|\mathcal{F}_{i-1}\right] \xrightarrow{\mathbb{P}} \tau^2(x). \tag{5.15}$$

Making use of condition (A5)(iv), one may write

$$\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}_w\left[\eta_{ni}^2|\mathcal{F}_{i-1}\right] &= \frac{v_n}{n}\sum_{i=1}^{n}\mathbb{E}_w\left[(g(X_i) - g(u))\Delta_i^2(x + \epsilon)\,|\mathcal{F}_{i-1}\right] \\
&\quad + \frac{v_n g(x)}{n}\sum_{i=1}^{n}\mathbb{E}_w\left[\Delta_i^2(x + \epsilon_n)\,|\mathcal{F}_{i-1}\right] \\
&=: \quad J_{1n}(x) + J_{2n}(x).
\end{aligned} \tag{5.16}$$

25

It is easily seen that $J_{1n} = o_{a.s.}(1)$ in view of conditions (A3) and the fact that the functions $g(\cdot)$ and $Q_{x+\epsilon}(\cdot)$ are bounded. Moreover, using Lemma 2 twice combined with Cauchez shwartz inequality, we obtain

$$
\begin{aligned}
J_{2n}(x) &\simeq \frac{g(x)f_w(x)}{2\sqrt{\pi}(x+\epsilon)} + v_n g(x) \int_0^\infty Q_{x+\epsilon}^2(u)[f_w(u) - f_w(x)]du \\
&\simeq \frac{g(x)f_w(x)}{2\sqrt{\pi}(x+\epsilon)} + O\left(\frac{1}{x+\epsilon}\right) \left(\int_0^\infty Q_{x+\epsilon}(u)[f(u) - f(x)]^2 du\right)^{1/2}. \quad (5.17)
\end{aligned}
$$

It follows then by the mean value theorem combined with condition (A7)-(i) and the fact that the function $q(\cdot)$ is gamma density with mean 1 and variance v_n^2

$$
\begin{aligned}
\int_0^\infty Q_{x+\epsilon}(u)[f(u) - f(x)]^2 du &\leq \sup_t (f_w'(t))^2 \int_0^\infty q(s)[x(s-1) + s\epsilon]^2 ds \\
&= O\left(\max(\epsilon^2, v_n^4)\right). \quad (5.18)
\end{aligned}
$$

Therefore, we have for any $x > 0$

$$
\begin{aligned}
\sum_{i=1}^n \mathbb{E}\left[\eta_{ni}^2 | \mathcal{F}_{i-1}\right] &= \frac{g(x)f_w(x)}{2\sqrt{\pi}x} + O\left(\frac{\max(\epsilon, v_n^2)}{x+\epsilon}\right) \\
&= \frac{g(x)f_w(x)}{2\sqrt{\pi}x} + o(1) := \tau^2(x) + o(1) \quad \text{as} \quad (\epsilon, v_n) \to (0,0). \quad (5.19)
\end{aligned}
$$

Making use of Hölder and Markov inequalities and condition (A6), one may easily see that the condition b) follows if

$$
n\mathbb{E}_w|\eta_{ni}|^{2+\delta} = o(1) \qquad \text{fro some} \quad \delta > 0.
$$

Combined conditions (A1) and (A6) and Lemma 2 we obtain for any $x > 0$

$$
\begin{aligned}
n\mathbb{E}_w(|\eta_{ni}|^{2+\delta}) &= O(v.(\frac{v}{n})^{\delta/2}) \int_0^\infty g_\delta(u)Q_{x+\epsilon}(u)^{2+\delta}(u)f_w(u)du \\
&= O\left(\frac{1}{(nv)^{\delta/2}(x+\epsilon)^{1+\delta}}\right) = o(1). \quad (5.20)
\end{aligned}
$$

This completes the proof of the first part of Proposition 1. The proof of the second part is similar by replacing v_n by $v_n\epsilon$ and x by 0 in the different steps of the proof of part one. $\square$

Proof of Theorem 1. Part (i). We have from (5.9), whenever $x > 0$, that

$$
\sqrt{nv_n}R_n(x) = o_\mathbb{P}\left(\frac{1}{\sqrt{nv}}\right) = o_\mathbb{P}(1).
$$

26

It follows then from (5.3), Lemma 3 and the first part of proposition 3 that $\sqrt{nv_n}(r_n(x) - B_n(x) - m(x))$ converges in distribution to a centered Gaussian random variable with variance $\sigma^2(x) = \frac{\tau^2(x)}{t^2(x)f_w^2(x)}$. Therefor, we obtain from (2.1), (2.2) and (2.3), whenever $x > 0$ and $f(x) > 0$, that

$$
\begin{aligned}
\sigma^2(x) &= \frac{\mathbb{E}(\kappa(X)w(X,Y))}{2\sqrt{\pi}\, x\, \kappa(x)f(x)} \mathbb{E}\left[w(X,Y)\, X = x\right] = \mathbb{E}_w\left[\left(\frac{\phi(Y) - m(x)}{w(X,Y)}\right)^2 |X = x\right] \\
&= \frac{\mathbb{E}(\kappa(X)w(X,Y))}{2\sqrt{\pi}\, x\, \kappa(x)f(x)} \mathbb{E}\left[\frac{(\phi(Y) - m(x))^2}{w(X,Y)} |X = x\right].
\end{aligned} \tag{5.21}
$$

Part ii). The second part of Theorem follows from the third part of Lemma 3 combined with the second part of proposition 3 and conditions 3.6.

Part iii). The proof of the third part of Theorem is similar . $\qquad\square$

Proof of Theorem 1. Following the decomposition (5.3), Lemma 3 and part iii) of Lemma 4, we have for any $x \geq 0$ that

$$
\sqrt{\frac{n\max(v, \epsilon^{1/2})}{\log\log n}} R_n(x) = O_{\mathbb{P}}\left(\frac{\max(v^{5/2}, \epsilon^{5/4})}{v_n(x + \epsilon)\sqrt{n\log\log n}}\right) = o_{\mathbb{P}}(1) \tag{5.22}
$$

in view of the first condition of (3.3) and

$$
\sqrt{\frac{n\max(v, \epsilon^{1/2})}{\log n}} B_n(x) = O_{\mathbb{P}}\left(\left(\frac{n\max(v^5, \epsilon^{5/2})}{\log\log n}\right)^{1/2}\right) = o_{\mathbb{P}}(1) \tag{5.23}
$$

in view of the second condition of (3.3).

To end the proof of theorem we have to give an estimate of the convergence rate of the quantity $S_n(x)$. Since $S_n(x)$ is a centered martingale with respect $\mathcal{F}_{i-1}$, one may then use successively the Burkholder inequality and the Jensen inequality to conclude that

$$
\mathrm{Var}(S_n(x)) = O\left(n^{-1}\mathbb{E}_w\left(g(X_i)\Delta_i^2(x + \epsilon)\right)\right) = O\left(\frac{1}{nv_n(x + \epsilon)}\right)
$$

in view of Lemma 2 and the fact that $g(\cdot)$ is bonded. The result follows then by application of the Tchebycheff inequality combined with the third condition of (3.3). This completes the proof. $\qquad\square$

Proof of Proposition 1. As a first step, observe that the conditional bias may be decomposed as follows

$$B_n(x) = \frac{\overline{r}_{n,2}(x) - m(x)\overline{r}_{n,1}(x)}{r_{n,1}(x)} + \frac{(\overline{r}_{n,2}(x) - m(x)r_{n,1}(x))(r_{n,1}(x) - \overline{r}_{n,1}(x)) + (r_{n,1}(x) - \overline{r}_{n,1}(x))^2 C_n(x)}{(r_{n,1}(x))^2}$$
$$:= B_n^*(x) + U_n(x)$$

where $B_n^*(x)$ stands as the main term while $U_n(x)$ is the residual one. We have first show that the numerator $r_{n,1}(x)$ of $B_n^*(x)$ converges almost surely to $t(x)f_w(x)$ and also to give an estimate of the convergence rite of the quantity $r_{n,1}(x) - \overline{r}_{n,1}(x)$.

Lemma 5 *Assuming conditions (B1) and (3.11) are satisfied.*

(i) We have

$$r_{n,1}(x) - \overline{r}_{n,1}(x) = O_{a.s.}\left(\frac{\log n}{nv_n(x+\epsilon)}\right).$$

(ii) If in addition that the conditions (A1)-(A4)-(ii) and (A5)-(ii) are satisfying, we gate

$$r_{n,1}(x) \longrightarrow t(x)f_w(x) \quad almost\ surely \quad as \quad n \to \infty.$$

Proof of Lemma 5. (i) First part: Proceeding as in the proof of Lemma 3, one can see that $r_{n,1}(x) - \overline{r}_{n,1}(x) = R_{1,n}(x) = \sum_{i=1}^{n} L_{ni}(x)$, where

$$|L_{ni}(x)| \leq \frac{C}{(x+\epsilon)v_n} := M_n$$

in view of condition (B1) and by application of the second part of Lemma 2. Moreover, we have by Jensen inequality and condition $(B1)$ combined with the first part of Lemma 2 that

$$\mathbb{E}\left(L_{in}^2(x) \mid \mathcal{F}_{i-1}\right) \leq C\mathbb{E}\left(|\Delta_i^2(x+\epsilon)\ \mathcal{F}_{i-1}\right) \leq \frac{C}{(x+\epsilon)v_n} = b.$$

Choosing in Corollary 3.1 of Laïb (1999)

$$\lambda = \lambda_n = \alpha\sqrt{\frac{\log n}{(x+\epsilon)nv_n}} \to 0 \quad as \quad n \to \infty,$$

where α is a large positive constant, it follows that

$$\mathbb{P}\left(|r_{n,1}(x) - \overline{r}_{n,1}(x)| \geq \lambda_n\right) \leq C_1 n^{-C\alpha}.$$

The desired result follows then by application of Borell Cantelli Lemma.

Proof of Part (ii). Write $r_{n,1}(x) = (r_{n,1}(x) - \overline{r}_{n,1}(x)) + \overline{r}_{n,1}(x)$. The first part of this decomposition goes to 0 almost surely in view of Lemma 5, whereas the second one is studied

in Lemma 4, which converges almost surely to $t(x)f)w(x)$ in view of conditions (A4)(ii) and (A5)(ii).

<div style="text-align: right">□</div>

Considering now $N_n(x)$ as the numerator in the form of $B_n^*(x)$, we have by condition (A3), for n large enough, that almost surely

$$
\begin{aligned}
N_n(x) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^\infty \psi_1(u)f_{i,w}(u)Q_{x+\epsilon}(u)du \\
&\simeq \int_0^\infty \psi_1(u)f_w(u)Q_{x+\epsilon}(u)du := \int_0^\infty \psi_2(s(x+\epsilon))q(s)ds,
\end{aligned}
$$

where

$$
\psi_1(u) := \mathbb{E}_w\left[\frac{\psi(Y)-m(x)}{w(X,Y)}\mid X=u\right] = s(u)-m(x)t(u) \quad \text{and } \psi_2(u)=\psi_1(u)f_w(u)
$$

Since $\psi_2(x)=0$, using the second part of Lemma 5, the condition (B2) and Taylor series expansion of the function ψ_2 to the order three in the neighborhood of x, we obtain

$$
N_n(x) = \epsilon\psi_2'(x)+\frac{1}{2}\psi_2''(x)[x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2]+o\left([x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2]\right).
$$

Observe now from (2.3) that $s(u)=m(u)t(u)$, thus $\psi_1'(x)=t(x)m'(x)$. It follows then that

$$
\begin{aligned}
B_n^*(x) &= \frac{N_n(x)}{t(x)f_w(x)} = \epsilon m'(x)+\frac{1}{2}\left\{m''(x)+m'(x)\left[\frac{t'(x)}{t(x)}+\frac{f_w'(x)}{f_w(x)}\right]\right\}[x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2] \\
&\quad + o\left([x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2]\right) \\
&= m'(x)\epsilon+\left(m''(x)+m'(x)\left[\frac{\kappa'(x)}{\kappa(x)}+\frac{f'(x)}{f(x)}\right]\right)[x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2] \\
&\quad + o\left([x^2v_n^2+2x\epsilon v_n^2+\epsilon_n^2 v_n^2+\epsilon^2]\right) \qquad\qquad (5.24)
\end{aligned}
$$

because $t(x)=\frac{1}{\gamma(x)}$ and $\frac{f_w'(x)}{f_w(x)}=\frac{\kappa'(x)}{\kappa(x)}+\frac{\gamma'(x)}{\gamma(x)}+\frac{f'(x)}{f(x)}$.

Considering finally the residual term, it is easily seen that

$$
U_n(x) = N_n(x)\frac{\hat{r}_{n,1}(x)-\overline{r}_{n,1}(x)}{\hat{r}_{n,1}^2(x)}+\frac{B_n^*(x)}{\hat{r}_{n,1}(x)\overline{r}_{n,1}(x)}(\hat{r}_{n,1}(x)-\overline{r}_{n,1}(x))^2.
$$

Combined Lemma 5 with the statement (5.24), one may see that

<div style="text-align: center">29</div>

$$U_n(x) = \frac{B_n^*(x)}{t(x)f_w(x)}O_{a.s.}\left(\sqrt{\frac{\log n}{(x+\epsilon)v_n}}\right) + \frac{B_n^*(x)}{t^2(x)f_w^2(x)}O_{a.s}\left(\frac{\log n}{(x+\epsilon)v_n}\right)$$

$$= \frac{\mu B_n^*(x)}{\kappa(x)f(x)}O_{a.s.}\left(\sqrt{\frac{\log n}{(x+\epsilon)v_n}}\right) + \frac{\mu^2 B_n^*(x)}{\kappa^2(x)f^2(x)}O_{a.s.}\left(\frac{\log n}{(x+\epsilon)v_n}\right)$$

$$= O_{a.s.}\left(\sqrt{\frac{\max(v_n^4,\epsilon^2)\log n}{(x+\epsilon)v_n}}\right) + O_{a.s.}\left(\frac{\max(v_n^2,\epsilon)\log n}{(x+\epsilon)v_n}\right),$$

which completes the proof of Proposition. $\qquad\square$

Proof of Proposition 2. Making use of the decomposition (5.3), it is easily seen that

$$(r_n(x) - m(x))^2 = \frac{(\overline{r}_{n,1}(x)B_n(x))^2 + S_n^2(x) + 2S_n(x)B_n(x)\overline{r}_{n,1}(x)}{(\overline{r}_{n,1}(x))^2}$$

$$= \frac{N_n^2(x) + S_n^2(x) + 2S_n(x)N_n(x)}{(\overline{r}_{n,1}(x))^2}$$

Since $\overline{r}_{n,1}(x) \to t(x)f_w(x)$ as $n \to \infty$, in view of Lemma 1, then we have approximately

$$\mathbb{E}_w(r_n(x) - m(x))^2 \simeq \frac{\mathbb{E}_w(N_n^2(x)) + \mathbb{E}_w(S_n^2(x))}{(t(x)f_w(x))^2} \tag{5.25}$$

ignoring the product term , because by the Cauchy-Schwartz inequality we have

$$\mathbb{E}_w(S_n(x)N_n(x)) \le \sqrt{\mathbb{E}_w(S_n^2(x))}\sqrt{\mathbb{E}_w(N_n^2(x))} \le \max\left(\mathbb{E}_w(S_n^2(x)),\ \mathbb{E}_w(N_n^2(x))\right).$$

We have to evaluate each term in (5.25). We begin by the bias term. To do that, let

$$K(u) := \mathbb{E}_w\left[\frac{\phi(Y_i) - m(x)}{w(X,Y)}|X = u\right] = s(u) - m(x)t(u) \qquad \text{and} \quad \psi_2(u) = K(u)f_w(u)$$

Combined Conditions (A5)(ii)-(iii) and (A3), one may approximate $N(x) \simeq \int_0^\infty \psi_2(u)Q_{x+\epsilon}(u)du$. By Taylor's expansion up to order three of $\psi_2(\cdot)$ around x, with higher order terms denoted by $\theta_n(x)$, we obtain (since $\psi_2(x) = 0$) that

$$N(x) \simeq \psi_2'(x)\int_0^\infty [x(s-1) + s\epsilon]q(s)ds + \frac{1}{2}\psi_2''(x)\int_0^\infty [x(s-1) + s\epsilon]^2 q(s)ds + \theta_n(x)$$

$$\simeq \epsilon\psi_2'(x) + \frac{1}{2}\psi_2''(x)A(v,\epsilon). \tag{5.26}$$

We have from the statements (2.1), (2.2) and (2.3) that

$$\frac{N^2(x)}{(t(x)f_w(x))^2} \simeq \frac{(\epsilon\psi_2'(x))^2 + \epsilon A(\epsilon,v)\psi_2'(x)\psi_2''(x) + \frac{1}{4}(\psi_2''A(v,\epsilon))^2}{(t(x)f_w(x))^2}$$

$$= (m'(x)\epsilon)^2 + \epsilon A(v,\epsilon)m'(x)\left\{m''(x) + m'(x)\left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)}\right]\right\}$$

$$+ \frac{1}{4}A^2(v,\epsilon)\left\{m''(x) + m'(x)\left[\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)}\right]\right\}. \tag{5.27}$$

We have to evaluate now the quantity $\mathbb{E}_w(S_n^2(x))$. Recall that

$$S_n(x) = \frac{1}{n}\sum_{i=1}^{n}[Z_{n,i}(x) - \mathbb{E}(Z_{n,i})], \quad \text{where} \quad Z_{n,i}(x) = \frac{\phi(Y_i) - m(x)}{w(X_i,Y_i)}\Delta_i(x+\epsilon).$$

Since for any $(i,j) \in \mathbb{N}^2$, we have $\mathbb{E}_w[Z_{n,i}(x)Z_{n,j}(x)] = 0$, whenever $i \neq j$. It follows then that

$$\mathbb{E}_w(S_n^2(x)) = \frac{1}{n^2}\sum_{i=1}^{n}\left[\mathbb{E}_w(Z_{n,i}^2) - \mathbb{E}_w\left((\mathbb{E}_w\left(Z_{n,i}\mid\mathcal{F}_{i-1})\right)^2\right)\right] \simeq \frac{1}{n}\mathbb{E}_w(Z_{n,1}^2). \tag{5.28}$$

Making use of Condition (A5)(iv), one may write

$$\frac{1}{n}\mathbb{E}_w(Z_{n,1}^2) = n^{-1}\int_0^{\infty} h(u)Q_{x+\epsilon}^2(u)du, \quad \text{where} \quad h(u) = g(u)f_w(u)$$

$$\simeq n^{-1}h(x)\int_0^{\infty} Q_{x+\epsilon}^2(u)du + n^{-1}h'(x)\int_0^{\infty}(u-x)Q_{x+\epsilon}^2(u)du$$

$$\simeq \frac{h(x)}{2\sqrt{\pi}}\frac{1}{(x+\epsilon)nv} + \frac{h'(x)}{2\sqrt{\pi}}\left[\frac{1}{nv} - \frac{x}{(x+\epsilon)nv}\right]. \tag{5.29}$$

Using the relation

$$g(x) = \frac{\psi(x)}{\mathbb{E}[w(X,Y)|X=x]} \quad \text{with} \quad \psi(x) = \mathbb{E}_w\left[\frac{(\phi(Y) - m(x))^2}{w(X,Y)}\Big|X=x\right],$$

we obtain, in view of the statements (2.1), (2.2) and (2.3) that $g(x) = \frac{\psi(x)}{\gamma(x)} = \psi(x)t(x)$ and

therefore

$$
\begin{aligned}
\frac{\mathbb{E}_w(S_n^2(x))}{(t(x)f_w(x))^2} &= \frac{g(x)}{2\sqrt{\pi}\ t^2(x)f_w(x)(x+\epsilon)nv} \\
&+ \frac{1}{2\sqrt{\pi}}\left[\frac{g'(x)}{t^2(x)f_w(x)} + \frac{g(x)f_w'(x)}{t^2(x)f_w^2(x)}\right]\left[\frac{1}{nv} - \frac{x}{(x+\epsilon)nv}\right] \\
&= \frac{\mu\psi(x)}{2\sqrt{\pi}(x+\epsilon)\kappa(x)f(x)nv} \\
&+ \frac{\mu}{2\sqrt{\pi}\ \kappa(x)f(x)}\left[\psi'(x) + \left(\frac{1}{\gamma(x)} - 1\right)\frac{\gamma'(x)}{\gamma(x)}\psi(x)\right. \\
&\left. +\psi(x)\left\{\frac{\kappa'(x)}{\kappa(x)} + \frac{f'(x)}{f(x)}\right\}\right]\left[\frac{1}{nv} - \frac{x}{(x+\epsilon)nv}\right].
\end{aligned}
$$

References

Ahmad, I.A. (1995). On multivariate kernel estimation for samples from weighted distribution. *Statist. Probab. Lett.* **22**, 121-129.

Chaubey, Y. P., Laïb, N. and Sen, A. (2010). Generalised kernel smoothing for non-negative stationary ergodic processes. *Journal of Nonparametric Statistics,* **22,** 973-997.

Cristóbal, J.A. and Alcalá, J.T. (2000). Nonparametric regression estimation for length biased data. *Journal of Statistical Planning and Inference*, **89**, 145–168.

Cristóbal, J.A. and Alcalá, J.T. (2001). An overview of nonparametric contributions to the problem of functional estimation form biased data. *Test*, **10,** 309–332.

Cox, D.R. (1969). Some sampling problems in technology. In: Johnson, N.L. (Eds.), *Encyclopedia of Statistical Sciences*, Vol. **9**. Wiley, New York, 506-527.

Hengartner, N. W. and Wegkamp, M. H. (2002). Bandwidth selection for local linear regression smoothers. *J. R. Statist. Soc. B*, **64**, 791-804.

Jones, M.C. (1991). Kernel density estimation for length biased data. *Biometrika,* **78**, 511-519.

Laïb, N. (1999). Exponential-type inequalities for martingale difference sequences: Application to nonparametric regression estimation. *Communication in Statistics, Theory and Methods, Series A,* **28,** 1565–1576.

Laïb, N. and Louani, D. (2010). Nonparametric Kernel regression estimate for functional stationary ergodic data: asymptotic properties. *Journal of Multivariate analysis,* **101,** 2266-2281.

Muttlak, H.A. and McDonald, L.L. (1990). Ranked set sampling with size-biased probabilities of selection. *Biometrics,* **46**, 435-446.

Ojeda, J.L., Cristóbal, J.A. and Alcalá, J.T. (2004). Nonparametric confidence bonds for GLM models with length biased data. *Journal of Nonparametric Statistics,* **16,** 421-441.

Olave, P., Salvador, M., Miguel, J.A. and Muńoz, L. (1998). The fecundity and the time of unemployment on the Spanish woman: A nonparametric study. *Minist. Trabajo y Asuntos Sociales. Espańa.*

Patil, G.P. and Rao, C.R. (1978). Weighted distribution and size-biased sampling with application to wildlife population and human families. *Biometrics,* **34,** 179–189.

Patil, G.P., Rao, C.R. and Zelen, M. (1988). *Weighted distribution.* In: Kotz, S., Johnson, N.L. (Eds.), Encyclopedia of Statistical Sciences, Vol. **9**. Wiley, New York, 565-571.

Seifert, B. and Gasser, T. (1996). Finite sample variance of local polynomials: Analysis and solutions, *J. Amer. Statist. Assoc.,* **91,** 267–275.

Sköld, M. (1999). Kernel regression in the presence of size-bias. *Journal of nonparametric statistics,* **12,** 41-51.

Vardi, Y. (1982). Nonparametric estimation in the presence of length bias. *Ann. Statist.* **10,** 616–620.

Wu, C.O. and Mao, A.Q. (1996). A cross-validation bandwidth choice for kernel density estimates with selection biased data. *J. Multivariate Anal.,* **61,** 38-60.