The optimal risk of estimator of conditional distribution function in a model of heteroscedastic regression with weakly dependent observations

Farkhad Abdikalikov
Karakalpak State University, abdikalikovf@mail.ru

Laura Reymova
Karakalpak State University, reymova1998@gmail.com

Follow this and additional works at: https://bulletin.nuu.uz/journal

Part of the Statistics and Probability Commons

Recommended Citation
DOI: https://doi.org/10.56017/2181-1318.1041

This Article is brought to you for free and open access by Bulletin of National University of Uzbekistan: Mathematics and Natural Sciences. It has been accepted for inclusion in Bulletin of National University of Uzbekistan: Mathematics and Natural Sciences by an authorized editor of Bulletin of National University of Uzbekistan: Mathematics and Natural Sciences. For more information, please contact karimovja@mail.ru.
THE OPTIMAL RISK OF ESTIMATOR OF CONDITIONAL DISTRIBUTION FUNCTION IN A MODEL OF HETEROSCEDASTIC REGRESSION WITH WEAKLY DEPENDENT OBSERVATIONS

ABDIKALIKOV F. A., REYMOVA L.
Karakalpak State University, Nukus, Uzbekistan
e-mail: abdikalikovf@mail.ru, reymova1998@gmail.com

Abstract

Paper is devoted to estimation of conditional distribution function in heteroscedastical regression model, in which responses are $\alpha$-mixing random variables. It is found the expression for mean square deviation of estimator and optimal window width sequence.

Keywords: model of heteroscedastical regression, $\alpha$-mixing, kernel estimate.

Mathematics Subject Classification (2010): 46N30, 62H12.

Introduction

The model of heteroscedastical regression of the variable of response $Z$ on $X$ is determined by the formula [1]:

$$Z = m(X) + \sigma(X)\varepsilon,$$

where $X$ - is a random covariate, random error $\varepsilon$ - is independent of $X$. The functionals $m(x) = M(Z/X = x)$ and $\sigma^2(x) = D(Z/X = x)$ respectively, are conditionally mean and variance functions of regression $Z$ on $X$ that is, possible heterossedality. Define the conditional distribution function (d.f.) $F_x(t) = P(Z \leq t/X = x)$, $(t; x) \in R^+ \times D_X$. By definition, functionals $m(x) = T(F_x(\cdot))$ and $\sigma(x) = S(F_x(\cdot))$ for all $a \geq 0$ and $b \in R$ satisfy the equalities (see, [1]):

$$T(Q_{aZ+b}(\cdot/x)) = aT(Q_Z(\cdot/x)) + b = am(x) + b,$$

$$S(Q_{aZ+b}(\cdot/x)) = aS(Q_Z(\cdot/x)) = a\sigma(x),$$

where $Q_{aZ+b}(t/x) = P(aZ + b \leq t/X = x) = F_x((t-b)/a)$. For them, it is true representations

$$m(x) = \int_0^1 F_x^{-1}(s)J(s)ds, \quad \sigma^2(x) = \int_0^1 (F_x^{-1}(s))^2J(s)ds - m^2(x), \quad (1)$$

where $F_x^{-1}(s) = \inf\{y : F_x(y) \geq s\}$, $0 \leq s \leq 1$, quantile function of a random variable (r.v.) $Z$ for a given $X = x$ and $J(s)$ - is a given score function such that $J(s) \geq 0$ and $\int_0^1 J(s)ds = 1$. As $J(s)$ one can take a function $J(s) = I(0 \leq s \leq 1)$. 
1 Preliminaries

The basic problem is to estimate the conditional d.f. $F_x(t)$ for a sample of observations on a pair $(Z, X): (Z_1, X_1), (Z_2, X_2), \ldots, (Z_n, X_n)$ where the subsample $\{Z_1, \ldots, Z_n\}$ satisfies the following condition of $\alpha$-mixing at $n \to \infty$:

$$
\alpha(n) = \sup_{k \geq 1} \left\{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^k(Z), B \in \mathcal{F}_k^n(Z) \right\} \to 0,
$$

where $\mathcal{F}_1^k(Z)$ is the algebra of events generated by a set of a r.v. $Z_i$, $i \leq j \leq k$. Among most classes of dependent r.v. $\alpha$-mixed r.v. are most common in practice (see, [2]). It should be noted that indicators $\{I(Z_i \leq t)\}$ also form a sequence $\alpha$-mixing r.v.-s.

As an estimator, for $F_x(t)$, we consider the following Nadarya-Watson statistics (see, [3, 4]):

$$
F_{xh}(t) = \sum_{i=1}^{n} \Psi_{ni}(x; h_n) I(Z_i \leq t), \quad (t; x) \in R^+ \times D_X, \tag{2}
$$

where weights

$$
\Psi_{ni}(x; h_n) = \left( \sum_{j=1}^{n} k \left( \frac{x - X_j}{h_n} \right) \right)^{-1} k \left( \frac{x - X_i}{h_n} \right), \quad i = 1, \ldots, n,
$$

are given by the sequence $h_n \downarrow 0$ as $n \to \infty$ and by the kernel $k(\cdot)$. It is easy to see by direct calculation that the conditional expectation and variance of the estimate (2) for given $X_1, \ldots, X_n$ are equal

$$
M^*F_{xh}(t) = M \left[ F_{xh}(t) / X_1, \ldots, X_n \right] = \sum_{i=1}^{n} \Psi_{ni}(x; h_n) F_{X_i}(t),
$$

$$
D^*F_{xh}(t) = D \left[ F_{xh}(t) / X_1, \ldots, X_n \right] = \sum_{i=1}^{n} \Psi_{ni}^2(x; h_n) F_{X_i}(t) (1 - F_{X_i}(t)). \tag{3}
$$

Since for sufficiently large $n$ and $x \in (h_n, 1 - h_n)$, $F_{X_i}(t) \approx F_x(t)$ and $\sum_{i=1}^{n} \Psi_{ni}(x; h_n) = 1$, it is natural to expect that the right sides of formulas (3) are asymptotically unbiased estimates of the corresponding expectations, that is, for $n \to \infty$ and $(t, x) \in R^+ \times D_X$:

$$
MF_{xh}(t) = F_x(t) + o(1), \quad DF_{xh}(t) = \frac{1}{nh_n} F_x(t) (1 - F_x(t)) + o(1). \tag{4}
$$

In this paper, we will estimate the quadratic risk of the estimate (2). We denote $\hat{F}_x(t) = \frac{\partial F_x(t)}{\partial x}$, $\hat{F}_x(t) = \frac{\partial^2 F_x(t)}{\partial x^2}$. For a sequence $\{h_n, n \geq 1\}$, a kernel $k(\cdot)$ and a conditional distribution function $F_x$, we introduce the conditions (see, [3]):

(C1) At $n \to \infty, h_n \downarrow 0$ and $nh_n \to \infty$;
(C2) The kernel \( k(\cdot) \) is a continuous, bounded and symmetric density with a compact support \([-M, M]\), for \( M > 0 \);

(C3) For a fixed \( t \in R^+ \) derivative \( \tilde{F}_x(t) \) exists and is bounded in a neighborhood \( U_x \) of \( x \).

Further, we also need the following statement.
Suppose \( \xi \) and \( \eta \) are two r.v. that measurable with respect to \( \mathcal{F}_1^k(Z) \) and \( \mathcal{F}_k^\infty(Z) \), respectively, and

\[
\|\xi\|_\infty = \text{ess} - \sup \|\xi\| = \inf\{t \in R^+ \cup \{+\infty\} : P(|\xi| > t) = 0\}.
\]

**Lemma 1 ([2])**. It is true the following inequality

\[
|\text{Cov}(\xi, \eta)| \leq 4\alpha(n)\|\xi\|_\infty \cdot \|\eta\|_\infty.
\]

For example, for an indicator \( I(Z_i \leq t) \):

\[
\text{ess} - \sup |I(Z_i \leq t)| = 1.
\]

## 2 Main part

The following statement gives an estimate for the standard deviation of \( F_{xh}(t) \) from \( F_x(t) \).

**Theorem 1.** Let the conditions (C1)-(C3) and the sequence \( \{Z_i, i \geq 1\} \) are satisfied the condition of \( \alpha \)-mixing so that, for \( n \to \infty \), \( \alpha(n) \to 0 \) and

\[
\frac{1}{nh^2} \sum_{i=1}^{n-1} \alpha(i) \to 0.
\]

Then for \( n \to \infty \)

\[
M[F_{xh}(t) - F_x(t)]^2 = \left[ \frac{h^2}{2} \tilde{F}_x(t) \int u^2 k(u)du \right]^2 + 
\frac{1}{nh_n} \int k^2(u)du \cdot F_x(t)(1 - F_x(t)) + o\left( h_n^4 + \frac{1}{nh_n} \right).
\]

**Proof.** Since

\[
M[F_{xh}(t) - F_x(t)]^2 = DF_{xh}(t) + [MF_{xh}(t) - F_x(t)]^2,
\]

we estimate the terms on the right side of the last equality separately. Denote \( \xi_i = \frac{x - X_i}{h_n} \). Then, by the Taylor expansion, we easily have the equalities

\begin{align*}
F_{X_i}(t) &= F_{x-h_n\xi_i}(t) = F_x(t) - h_n\xi_i \hat{F}_x(t) + \frac{h_n^2}{2} \xi_i^2 \tilde{F}_x(t) + o(h_n^2), \\
F_{X_i}^2(t) &= F_{x-h_n\xi_i}^2(t) = F_x^2(t) - 2h_n\xi_i F_x(t) \hat{F}_x(t) + \\
&\quad + \frac{h_n^2}{2} \xi_i^2 \left( \hat{F}_x(t) \right)^2 + \frac{h_n^2}{2} \xi_i^2 F_x(t) \tilde{F}_x(t) + o(h_n^2).
\end{align*}

164
Taking into account (6) and the symmetry of the kernel $k(\cdot)$ (in condition (C2)) for the expectation, we have

$$MF_{xh}(t) = F_x(t) + \frac{h_n^2}{2} \ddot{F}_x(t) \int_{\mathbb{R}^n} k(u) du + o(h_n^2) =$$

$$= F_x(t) + \frac{h_n^2}{2} \ddot{F}_x(t) \int u^2 k(u) du + o(1),$$

where we used following equalities

$$R_{n}^{(0)} = \frac{1}{nh_n} \sum_{i=1}^{n} k(\xi_i) = \int k(u) du + o(1) = 1 + o(1), \quad n \to \infty,$$

$$R_{n}^{(1)} = \frac{1}{nh_n} \sum_{i=1}^{n} \xi_i k(\xi_i) = \int uk(u) du + o(1) = o(1), \quad n \to \infty,$$

$$R_{n}^{(2)} = \frac{1}{nh_n} \sum_{i=1}^{n} \xi_i^2 k(\xi_i) = \int u^2 k(u) du + o(1), \quad n \to \infty. \quad (9)$$

From (9), in particular, when $n \to \infty$ follows the asymptotic unbiasedness of estimate $MF_{xh}(t) = F_x(t) + O(h_n^2)$. Now we calculate the conditional variance of the estimate:

$$DF_{xh}(t) = \sum_{i=1}^{n} k^2(\xi_i) \left[ F_{x_i}(t) - F_{\xi_i}(t) \right] +$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{nh_n} \sum_{i=1}^{n} k(\xi_i) k(\xi_j) \text{Cov} (I(Z_i \leq t), I(Z_j \leq t)) \left[ \int_{\mathbb{R}^n} k(u) du + o(1) \right] = D_{1n} + D_{2n}. \quad (10)$$

We study the terms on the right-hand side of (10). Hence, we calculate

$$Q_{n}^{(0)} = \frac{1}{nh_n} \sum_{i=1}^{n} k^2(\xi_i) = \int k^2(u) du + o(1), \quad n \to \infty, \quad (11)$$

and also due to the symmetry of $k(\cdot)$,

$$Q_{n}^{(1)} = \frac{1}{nh_n} \sum_{i=1}^{n} \xi_i k^2(\xi_i) = \int uk^2(u) du + o(1) = o(1), \quad n \to \infty. \quad (12)$$

Using relations (6), (7), (9), (11) and (12), we have

$$D_{1n} = \left[ \sum_{i=1}^{n} k(\xi_i) \right]^{-2} \left\{ [F_{x}(t) - F_{x}^2(t)] \sum_{i=1}^{n} k^2(\xi_i) +
+ \frac{h_n^2}{2} \left[ \ddot{F}_x(t) - \left( \ddot{F}_x(t) \right)^2 - F_x(t) \dddot{F}_x(t) \right] \sum_{i=1}^{n} \xi_i^2 k^2(\xi_i) + o(1) \right\} =$$

$$= \frac{1}{nh_n} \left[ F_{x}(t) - F_{x}^2(t) \right] \int k^2(u) du + o(1). \quad (13)$$
Now we evaluate $D_{2n}$. Let be $c(j - i) = \text{Cov}(I(Z_i \leq t), I(Z_j \leq t))$. Then, in view of the first equality in (9),

$$|D_{2n}| \leq \left[ R_n^{(0)} \right]^{-2} \frac{1}{n^2 h_n^2} \left| \sum_{i=1}^{n} \sum_{j \neq i}^{n} k(\xi_i)k(\xi_j)c(j - i) \right| \leq \frac{2k_0^2}{nh_n^2} \sum_{i=1}^{n} (n - i)|c(i)| \leq \frac{2k_0^2}{nh_n^2} \sum_{i=1}^{n} |c(i)| \leq \frac{8k_0^2}{nh_n^2} \sum_{i=1}^{n-1} \alpha(i) = o(1), \ n \to \infty,$$

where $k_0 = \sup_{u \in R} k(u) \in (0, \infty)$, the lemma and conditions of the theorem are used. Now the statement of the theorem follows from (5), (8), (10), and (14). Theorem 1 is proved.

**Corollary 1.** Under the conditions of Theorem 1, for sufficiently large $n$, we have the following asymptotic representation for the variance of the estimate $F_{zh}(t)$:

$$DF_{zh}(t) = \frac{1}{nh_n} F_x(t)(1 - F_x(t)) \int k^2(u)du + o(1), \ x \in (h_n, 1 - h_n). \quad (15)$$

From Theorem 1 also follows for given covariates $X_1, \ldots, X_n$ the mean square consistency of the estimate $F_{zh}(t)$. Under the conditions of Theorem 1, we can write the following asymptotic representation for the quadratic risk of an estimate $F_{zh}(t)$ with a given weight function $w(\cdot)$ for a fixed $t \in R^+$:

$$\int_a^b M[F_{zh}(t) - F_x(t)]^2 w(x)dx \approx I(h_n),$$

where

$$I(h_n) = \frac{RA}{4} h_n^4 + \frac{QB}{nh_n}, \ Q = \int_{-M}^{M} k^2(u)du, \ R = \int_{-M}^{M} u^2 k(u)du, \ A = \int_a^b \tilde{F}_x(t)w(x)dx,$$

$$B = \int_a^b F_x(t)(1 - F_x(t))w(x)dx.$$

In order to find the optimal sequence $\{h_n, n \geq 1\}$ that gives the least value to risk, we solve the equation:

$$\frac{\partial I(h_n)}{\partial h} = RAh_n^3 - \frac{QB}{nh_n^2} = 0,$$

from where we find $h_{n, \text{opt}} = C n^{-1/5}$, where $C = \left( \frac{QB}{RA} \right)^{1/5}$.

For example, if $k(\cdot)$ is the uniform distribution density on $[-1, 1]$, then $Q = 1/2$ and $R = 1/3$ and the value $C$ also depends only on the degree of smoothness of the function $F_x(t)$ with respect to $x \in D_X$. 

166
References


