THE ASYMPTOTIC NORMALITY OF THE KOENKER–BASSETT ESTIMATORS IN NONLINEAR REGRESSION MODELS

UDC 519.21

O. V. IVANOV AND I. V. ORLOVSKIĬ

ABSTRACT. The asymptotic normality of the Koenker–Bassett estimators of parameters of a nonlinear regression model with discrete time and independent errors of observations is studied in the paper.

INTRODUCTION

We study the asymptotic normality of the Koenker–Bassett estimators [1] (in other words, the generalized least modules estimators) of parameters of a nonlinear regression model that are generalizations of the usual least modules estimators to the case where the errors of observations are nonsymmetric random variables (in the sense that the distribution function of errors at zero does not equal $\frac{1}{2}$).

The generalized least modules estimators belong to the class of M -estimators [2]. The most studied among M -estimators are the least squares estimators and least modules estimators [3].

The consistency of generalized least modules estimators of parameters of a nonlinear regression model is considered in [4]. There is another paper [5] devoted to the Koenker–Bassett estimators, but the results of [5] are wrong.

1. MAIN ASSUMPTIONS AND RESULTS

Assume that the observations X_j are random variables assuming values in $(\mathbf{R}^1, \mathcal{B}^1)$ and having the distribution P_j (here $\mathbf{R}^1$ stands for the set of real numbers, $\mathcal{B}^1$ is the σ -algebra of Borel subsets of $\mathbf{R}^1$). Assume also that the unknown distribution P_j belongs to some parametric family $\{P_{j\theta}, \theta \in \Theta\}$. The triple $\mathcal{E}_j = \{\mathbf{R}^1, \mathcal{B}^1, P_{j\theta}, \theta \in \Theta\}$ is called the statistical experiment generated by an observation X_j .

We say that a statistical experiment

$$\mathcal{E}^n = \{\mathbf{R}^n, \mathcal{B}^n, P_\theta^n, \theta \in \Theta\}$$

is a product of statistical experiments $\mathcal{E}_i$, $i = 1, \dots, n$, if $P_\theta^n = P_{1\theta} \times \dots \times P_{n\theta}$ ($\mathbf{R}^n$ is n -dimensional Euclidean space and $\mathcal{B}^n$ is the σ -algebra of its Borel subsets). We say in this case that a statistical experiment $\mathcal{E}^n$ is generated by n independent observations $X = (X_1, \dots, X_n)$.

Consider a nonlinear regression model

$$(1.1) \quad X_j = g(j, \theta) + \varepsilon_j, \quad j = 1, \dots, n,$$

where $g(j, \theta)$ is a nonrandom function defined on Θ^c . Here Θ^c is the closure of an open convex set $\Theta \subset \mathbf{R}^q$. We also assume that

2000 *Mathematics Subject Classification.* Primary 62J02; Secondary 62J99.

A1. The ε_j are independent identically distributed random variables with zero mean, and the distribution function $\mathcal{P}$ is continuous at zero. Let

$$(1.2) \quad \mathcal{P}(0) = \beta, \quad \beta \in (0, 1).$$

Definition. Any random vector

$$\widehat{\theta}_n = \widehat{\theta}_n(X_j, j = 1, \dots, n) \in \Theta^c$$

such that

$$(1.3) \quad S_\beta(\widehat{\theta}_n) = \inf_{\tau \in \Theta^c} S_\beta(\tau), \quad S_\beta(\tau) = \sum \rho_\beta(X_j - g(j, \tau))$$

is called the generalized least modules estimator of the unknown parameter $\theta \in \Theta$ constructed from observations X_j , $j = 1, \dots, n$, of the form (1.1) where $\sum = \sum_{j=1}^n$ (this notation is used throughout the paper) and

$$(1.4) \quad \rho_\beta(x) = \begin{cases} \beta x, & x \geq 0, \\ (\beta - 1)x, & x < 0, \end{cases} \quad \beta \in (0, 1).$$

The function $\rho_\beta(x)$ is called the risk function of the generalized least modules estimator.

Below we list some assumptions posed on the random variables ε_j :

A2. $\mu_s = \mathbb{E}|\varepsilon_j|^s < \infty$ for some integer positive s .

Note that $\mathbb{E}\rho_\beta^s(\varepsilon_j) < \infty$ if condition **A2** holds.

A3. Random variables ε_j have bounded density $p(x) = \mathcal{P}'(x)$ such that

$$|p(x) - p(0)| \leq H|x|, \quad p(0) > 0,$$

where $H < \infty$ is some constant.

Example. The random variable $\xi = \chi_{2m}^2 - 2m$ where χ_{2m}^2 is a χ^2 random variable with an even degree of freedom is an example for which conditions **A1–A3** hold and whose distribution function at zero differs from $\frac{1}{2}$.

To prove that $\mathcal{P}(0) \neq \frac{1}{2}$ it is sufficient to show that the equation

$$\int_0^m x^{m-1} e^{-x} dx = \frac{1}{2} \Gamma(m)$$

has no solutions for a positive integer m . This is, indeed, the case, since the left-hand side of the equation

$$\int_0^m x^{m-1} e^{-x} dx = (m-1)! - e^{-m} \left(\sum_{i=0}^{m-1} \frac{(m-1)!}{i!} m^i \right)$$

is an irrational number, while its right-hand side is a rational number.

By $\mathcal{C}^q \subset \mathcal{B}^q$ we denote the class of all convex Borel subsets in $\mathbf{R}^q$. Let $T \subset \Theta$ be a compact set.

We introduce the following notation:

$$\begin{aligned} g_i(j, \tau) &= \frac{\partial}{\partial \tau^i} g(j, \tau), \quad g_{il}(j, \tau) = \frac{\partial^2}{\partial \tau^i \partial \tau^l} g(j, \tau), \quad i, l = 1, \dots, q; \\ d_n^2(\theta) &= \text{diag}(d_{in}^2(\theta))_{i=1}^q = \text{diag} \left(\sum g_i^2(j, \theta) \right)_{i=1}^q; \\ d_{il,n}^2(\tau) &= \sum g_{il}^2(j, \tau), \quad \tau \in \Theta^c, \quad i, l = 1, \dots, q. \end{aligned}$$

We change the variable $u = n^{-1/2} d_n(\theta)(\tau - \theta)$ in the regression function; that is,

$$g(j, \tau) = g \left(j, \theta + n^{1/2} d_n^{-1}(\theta) u \right) = f(j, u)$$

assuming that θ is the true value of the parameter. The parametric set Θ is transformed into $\tilde{U}_n(\theta) = n^{-1/2}U_n(\theta)$ under this change where $U_n(\theta) = d_n(\theta)(\Theta - \theta)$. The main advantage of this change is that the generalized least modules estimator $\hat{\theta}_n$ is transformed into the normalized random vector

$$\hat{u}_n = n^{-1/2}d_n(\theta)(\hat{\theta}_n - \theta)$$

whose consistency is proved in [4]. Below we study the asymptotic normality of the vector $n^{1/2}\hat{u}_n$.

By k we denote positive constants. Assume that

B1. The functions $g(j, \theta)$, $j \geq 1$, are continuous on Θ^c together with all their first partial derivatives and, in addition, $g_i(j, \theta)$, $i = 1, \dots, q$, $j \geq 1$, are continuously differentiable in Θ ; moreover,

$$(1.5) \quad (i) \quad \sup_{\theta \in T} \sup_{u \in v(R) \cap \tilde{U}_n^c(\theta)} \max_{1 \leq j \leq n} \frac{|f_i(j, u)|}{d_{in}(\theta)} \leq k^i(R)n^{-1/2}, \quad i = 1, \dots, q,$$

$$(1.6) \quad (ii) \quad \sup_{\theta \in T} \sup_{u \in v(R) \cap \tilde{U}_n^c(\theta)} \frac{d_{il,n}(\theta + n^{1/2}d_n^{-1/2}(\theta)u)}{d_{in}(\theta)d_{ln}(\theta)} \leq k^{il}(R)n^{-1/2}, \quad i, l = 1, \dots, q,$$

for all $R \geq 0$.

It follows from (1.5) that

$$(1.7) \quad \sup_{\theta \in T} \sup_{u_1, u_2 \in v^c(R) \cap \tilde{U}_n^c(\theta)} n^{-1} \frac{\Phi_n(u_1, u_2)}{|u_1 - u_2|^2} \leq k(R),$$

where $\Phi_n(u_1, u_2) = \sum (f(j, u_1) - f(j, u_2))^2$.

Condition (1.6) implies that

$$(1.8) \quad \sup_{\theta \in T} \sup_{u_1, u_2 \in v^c(R) \cap \tilde{U}_n^c(\theta)} \frac{\Phi_n^{(i)}(u_1, u_2)}{d_{in}^2(\theta)|u_1 - u_2|^2} \leq \tilde{k}^{(i)}(R),$$

where $\Phi_n^{(i)}(u_1, u_2) = \sum ((f_i(j, u_1) - f_i(j, u_2))^2)$, $i = 1, \dots, q$.

Assume that the generalized least modules estimator is consistent, that is,

C. For all $r > 0$,

$$\sup_{\theta \in T} P_\theta^n \left\{ |n^{-1/2}d_n(\theta)(\hat{\theta}_n - \theta)| \geq r \right\} = z_n(s),$$

where

$$z_n(s) = \begin{cases} O(n^{-s+1}), & s \geq 2, \\ o(1), & s = 1. \end{cases}$$

Some sufficient conditions for **C** are given in [4].

Let

$$I(\theta) = \left(d_{in}^{-1}(\theta)d_{ln}^{-1}(\theta) \sum g_i(j, \theta)g_l(j, \theta) \right)_{i,l=1}^q, \quad \theta \in \Theta.$$

The matrix $I(\theta)$ is symmetric and nonnegative definite. Let $\lambda_{\min}(I(\theta))$ be the minimal eigenvalue of the matrix $I(\theta)$.

Below we make use of the following condition.

B2. For $n > n_0$,

$$\inf_{\theta \in T} \lambda_{\min}(I(\theta)) \geq \lambda_0 > 0.$$

Let l be an arbitrary direction in $\mathbb{R}^q$ and let $\tau \in \Theta$. Then

$$\frac{\partial}{\partial l} S_\beta(\tau) = \sum \langle \nabla g(j, \tau), l \rangle (\chi\{X_j * g(j, \tau) - \beta\}),$$

where “ $*$ ” stands for “ $\leq$ ” if $\langle \nabla g(j, \tau), l \rangle \geq 0$ or for “ $<$ ” if $\langle \nabla g(j, \tau), l \rangle < 0$.

Let r_0 be the distance between T and $\mathbb{R}^q \setminus \Theta$. If the event $\{|\hat{\theta}_n - \theta| < r\}$ occurs for $\theta \in T$ and $r < r_0$, then

$$\frac{\partial}{\partial l} S_\beta(\hat{\theta}_n) \geq 0$$

for an arbitrary direction l . We use this observation in the proof of the following result.

Theorem. *If conditions **A1–A3**, **B1**, **B2**, and **C** hold, then*

$$(1.9) \quad \sup_{\theta \in T} \sup_{C \in \mathcal{C}^q} \left| P_\theta^n \left\{ \frac{p(0)}{\sqrt{\beta(1-\beta)}} I^{1/2}(\theta) d_n(\theta) (\hat{\theta}_n - \theta) \in C \right\} - \Phi(C) \right| \xrightarrow{n \rightarrow \infty} 0,$$

where

$$\Phi(C) = \int_C \frac{1}{(2\pi)^{q/2}} e^{-|x|^2/2} dx.$$

In other words, this result claims that the normal distribution

$$N\left(0, \frac{\beta(1-\beta)}{p^2(0)} I^{-1}(\theta)\right)$$

is the accompanying law for the distribution of the normalized estimator $d_n(\theta)(\hat{\theta}_n - \theta)$.

2. AUXILIARY RESULTS

Our proof below is based on the idea of the proof of the asymptotic normality of least modules estimators [3] and uses the method of partitioning the parametric set due to Huber [2, 6].

Let $l_1, \dots, l_q$ be positive directions of coordinate axes. Consider the vectors $S_\beta^\pm(\tau)$ with the coordinates

$$S_{i\beta}^\pm(\tau) = d_{in}^{-1}(\theta) \left(\frac{\partial}{\partial(\pm l_i)} \right) S_\beta(\tau), \quad i = 1, \dots, q,$$

and vectors $E_\theta^n S_\beta^\pm(\theta)$ with coordinates

$$E_\theta^n S_{i\beta}^\pm(\tau) = \pm d_{in}^{-1}(\theta) \sum g_i(j, \tau) [\mathcal{P}(g(j, \tau) - g(j, \theta)) - \beta], \quad i = 1, \dots, q.$$

It is clear that

$$E_\theta^n S_\beta^\pm(\theta) = 0$$

in view of condition **A1**. Let $S_\beta^{*\pm}(u) = S_\beta^\pm(\theta + n^{1/2} d_n^{-1}(\theta) u)$ and

$$z_n^\pm(\theta, u) = \frac{|S_\beta^{*\pm}(u) - S_\beta^{*\pm}(0) - E_\theta^n S_\beta^{*\pm}(u)|}{1 + |E_\theta^n S_\beta^{*\pm}(u)|}.$$

Lemma 1. *If all the assumptions of the theorem hold, then*

$$(2.1) \quad \sup_{\theta \in T} P_\theta^n \left\{ \sup_{u \in v^c(r) \cap \bar{U}_n^c(\theta)} z_n^\pm(\theta, u) > \epsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

for all $\epsilon > 0$ and sufficiently small $r > 0$.

Proof. We prove the result for $z_n^+(\theta, u)$. For simplicity, we consider the case of $r = 1$ and $T = C_0$ in (2.1), where

$$C_0 = \left\{ u : |u|_0 = \max_{1 \leq i \leq q} |u_i| \leq 1 \right\} \supset v(1).$$

Using $N_0 = O(\ln n)$ cubes $C_{(1)}, \dots, C_{(N_0)}$ we cover C_0 as follows. Let $t \in (0, 1)$ be a certain number. Let

$$C^{(m)} = \{u: |u|_0 \in [(1-t)^{m+1}, (1-t)^m]\}, \quad m = 0, \dots, m_0 - 1,$$

$$C^{(m_0)} = \{u: |u|_0 \leq (1-t)^{m_0}\}.$$

We cover every one of the sets $C^{(m)}$ by cubes whose edges are of length

$$a_m = (1-t)^m - (1-t)^{m+1} = t(1-t)^m$$

and then assign a number to every cube. The system of these cubes forms a necessary covering

$$C_{(1)}, \dots, C_{(N_0-1)}, \quad C_{(N_0)} \stackrel{\text{def}}{=} C^{(m_0)}.$$

Let $m_0 = m_0(n)$ be such that

$$(1-t)^{\tilde{m}_0} = n^{-\gamma}, \quad m_0 = [\tilde{m}_0], \quad \gamma \in \left(\frac{1}{2}, 1\right).$$

Note that the $|\cdot|_0$ -distance between $C_{(j)}$ and 0 is equal to

$$r(j) = (1-t)n^{-\gamma m/\tilde{m}_0},$$

while the $|\cdot|_0$ -diameter of $C_{(j)}$ is equal to

$$a(j) = tn^{-\gamma m/\tilde{m}_0}$$

for some $m = m(j)$, $j = 1, \dots, N_0 - 1$. Moreover, if a cube $C_{(j)}$ is an element of the covering of the set $C^{(m)}$, then

$$a(j) = a_m, \quad r(j) = t(1-t)^{m+1} + \dots + t(1-t)^{m_0-1} + (1-t)^{m_0}.$$

Without loss of generality, we assume that the number of cubes $C_{(j)}$ of the covering of every set $C^{(m)}$ does not depend on m , thus also on n . Indeed, consider an arbitrary octant in $\mathbf{R}^q$. The volume of the common part of the set $C^{(m)}$ and this octant is

$$(1-t)^{mq} - (1-t)^{(m+1)q},$$

while the volume of the set $C_{(j)}$ is

$$a^q(j) = t^q(1-t)^{mq}.$$

Therefore, the number of cubes $C_{(j)}$ that belong to the common part of $C^{(m)}$ and the octant does not exceed

$$\frac{(1-t)^{mq} - (1-t)^{(m+1)q}}{t^q(1-t)^{mq}} = \frac{1 - (1-t)^q}{t^q}.$$

Since $m_0 = O(\ln n)$, we have $N_0 = O(\ln n)$. Fix $\theta \in T$. Then

$$(2.2) \quad P_\theta^n \left\{ \sup_{u \in C_0} z_n^+(\theta, u) > \epsilon \right\} \leq \sum_{j=1}^{N_0} P_\theta^n \left\{ \sup_{u \in C_{(j)}} z_n^+(\theta, u) > \epsilon \right\}.$$

We estimate every term in (2.2) separately. The general entry of the matrix of derivatives $D_n(u)$ of the mapping

$$u \longrightarrow E_\theta^n S_\beta^{*+}(u)$$

is of the form

$$\begin{aligned}
D_n^{il}(u) &= \frac{\partial}{\partial u_l} E_\theta^n S_{i\beta}^{*+}(u) \\
&= n^{1/2} d_{in}^{-1}(\theta) d_{ln}^{-1}(\theta) \sum f_{il}(j, u) [\mathcal{P}(g(j, \tau) - g(j, \theta)) - \beta] \\
&\quad + n^{1/2} d_{in}^{-1}(\theta) d_{ln}^{-1}(\theta) \sum f_i(j, u) f_l(j, u) p(g(j, \tau) - g(j, \theta)) \\
&= {}_1 D_n^{il}(u) + {}_2 D_n^{il}(u).
\end{aligned}$$

Considering (1.6), (1.7), and the inequality

$$\sup_{x \in \mathbf{R}^1} p(x) = p_0 < \infty$$

we obtain for $|u| < r$ that

$$\begin{aligned}
(2.3) \quad n^{-1/2} |{}_1 D_n^{il}(u)| &\leq n^{1/2} d_{in}^{-1}(\theta) d_{ln}^{-1}(\theta) d_{il,n}(\theta + n^{1/2} d_n^{-1}(\theta) u) \\
&\quad \times \left(n^{-1} \sum (\mathcal{P}(f(j, u) - f(j, 0)) - \mathcal{P}(0))^2 \right)^{1/2} \\
&\leq k^{(il)}(r) k^{1/2}(r) p_0 |u|.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
(2.4) \quad &\left| n_2^{-1/2} D_n^{il}(u) - p(0) I_{il}(\theta) \right| \\
&\leq p_0 \left[d_{in}^{-1}(\theta) d_{in}(\theta + n^{1/2} d_n^{-1}(\theta) u) d_{ln}^{-1}(\theta) (\Phi_n^{(l)}(u, 0))^{1/2} + d_{in}^{-1}(\theta) (\Phi_n^{(i)}(u, 0))^{1/2} \right] \\
&\quad + d_{in}^{-1}(\theta) d_{ln}^{-1}(\theta) \left| \sum g_i(j, \theta) g_l(j, \theta) (p(f(j, u) - f(j, 0)) - p(0)) \right|.
\end{aligned}$$

It follows from (1.5) and (1.8) that the terms in the square brackets are bounded by

$$p_0 \left((\tilde{k}^{(i)})^{1/2} + k^{(i)}(r) (\tilde{k}^{(l)})^{1/2} \right) |u|.$$

We find a bound for the second term on the right-hand side of (2.4) by using condition **A3** and inequalities (1.5):

$$\begin{aligned}
(2.5) \quad &n^{1/2} d_{in}^{-1}(\theta) \max_{1 \leq j \leq n} |g_i(j, \theta)| \left(n^{-1} \sum (p(f(j, u) - f(j, 0)) - p(0))^2 \right)^{1/2} \\
&\leq k^{(i)}(r) H k^{1/2}(r) |u|.
\end{aligned}$$

It follows from condition **B2** that the matrix

$$n^{-1/2} D_n(0) = p(0) I(\theta)$$

is positive definite. Thus the above reasoning implies that

$$(2.6) \quad \inf_{\theta \in T} \left| E_\theta^n S_\beta^+(\theta + n^{1/2} d_n^{-1}(\theta) u) \right| \geq k_0 n^{1/2} |u|_0$$

for sufficiently small u (for simplicity, we assume that $u \in C_0$) and some $k_0 > 0$.

Let $l \neq N_0$ and $v \in C_{(l)}$ be an arbitrary point. Using (2.6) we obtain

$$\begin{aligned} \sup_{u \in C_{(l)}} z_n^+(\theta, u) &\leq \left(\sup_{u \in C_{(l)}} M_n^{(l)}(\theta, u, v) + L_n^{(l)}(\theta, v) \right) (1 + k_0 n^{1/2} r(l))^{-1}, \\ M_n^{(l)}(\theta, u, v) &= \sum_{\lambda=1}^4 M_{\lambda n}^{(l)}(\theta, u, v) \pmod{P_\theta^n}, \\ M_{1n}^{(l)}(\theta, u, v) &= \left| d_n^{-1}(\theta) \sum \nabla f(j, u) (\chi\{X_j * f(j, u)\} - \chi\{X_j < f(j, v)\}) \right|, \\ M_{2n}^{(l)}(\theta, u, v) &= \left| d_n^{-1}(\theta) \sum (\nabla f(j, u) - \nabla f(j, v)) (\chi\{X_j < f(j, v)\} - \beta) \right|, \\ M_{3n}^{(l)}(\theta, u, v) &= \left| d_n^{-1}(\theta) \sum \nabla f(j, u) (\mathcal{P}(f(j, u) - f(j, 0)) - \mathcal{P}(f(j, v) - f(j, 0))) \right|, \\ M_{4n}^{(l)}(\theta, u, v) &= \left| d_n^{-1}(\theta) \sum (\nabla f(j, u) - \nabla f(j, v)) (\mathcal{P}(f(j, v) - f(j, 0)) - \beta) \right|, \\ L_n^{(l)}(\theta, v) &= \left| d_n^{-1}(\theta) \sum \left(\nabla f(j, v) (\chi\{X_j < f(j, v)\} - \beta) - \nabla f(j, 0) (\chi\{\varepsilon_j * 0\} - \beta) \right. \right. \\ &\quad \left. \left. - \nabla f(j, v) (\mathcal{P}(f(j, v) - f(j, 0)) - \beta) \right) \right| \pmod{P_\theta^n}. \end{aligned}$$

It follows from (1.8) that

$$(2.7) \quad n^{-1/2} M_{2n}^{(l)}(\theta, u, v) \leq \beta' \left(\sum_{i=1}^q d_{in}^{-2}(\theta) \Phi_n^{(i)}(u, v) \right)^{1/2} \leq k_1 a(l),$$

where

$$k_1 = \beta' \left(\sum_{i=1}^q \tilde{k}^{(i)}(1) \right)^{1/2}.$$

Note also that (1.5), (1.7), and **A3** imply that

$$(2.8) \quad \begin{aligned} n^{-1/2} M_{3n}^{(l)}(\theta, u, v) &\leq p_0 n^{-1/2} \Phi_{2n}^{1/2}(u, v) \left(\sum_{i=1}^q \frac{d_{in}^2(\theta + n^{1/2} d_n^{-1}(\theta) u)}{d_{in}^2(\theta)} \right)^{1/2} \\ &\leq k_2 a(l), \end{aligned}$$

where

$$k_2 = p_0 k^{1/2}(1) \left(\sum_{i=1}^q (k^{(i)}(1))^2 \right)^{1/2}.$$

Similarly,

$$(2.9) \quad \begin{aligned} n^{-1/2} M_{4n}^{(l)}(\theta, u, v) &\leq p_0 n^{-1/2} \Phi_{2n}^{1/2}(v, 0) \left(\sum_{i=1}^q d_{in}^{-2}(\theta) \Phi_n^{(i)}(u, v) \right)^{1/2} \\ &\leq k_3 a(l), \end{aligned}$$

where

$$k_3 = p_0 k^{1/2}(1) \left(\sum_{i=1}^q \tilde{k}^{(i)}(1) \right)^{1/2}.$$

Now we estimate $M_{1n}^{(l)}(\theta, u, v)$. For arbitrary $u, v \in C_{(l)}$ we have

$$\begin{aligned} &|\chi\{X_j * f(j, u)\} - \chi\{X_j < f(j, v)\}| \\ &\leq \chi \left\{ \inf_{u \in C_{(l)}} f(j, u) - f(j, 0) \leq \varepsilon_j \leq \sup_{u \in C_{(l)}} f(j, u) - f(j, 0) \right\} = \chi_j \pmod{P_\theta^n}. \end{aligned}$$

Therefore

$$(2.10) \quad n^{-1/2} M_{1n}^{(l)}(\theta, u, v) \leq n^{-1/2} \left(\sum_{i=1}^q \left(d_{in}^{-1}(\theta) \max_{1 \leq j \leq n} |f_i(j, u)| \right)^2 \right)^{1/2} \sum \chi_j \\ \leq k_4 n^{-1} \sum \chi_j$$

by (1.5), where

$$k_4 = \left(\sum_{i=1}^q (k^{(i)}(1))^2 \right)^{1/2}.$$

By the mean value theorem,

$$(2.11) \quad n^{-1} \sum E_{\theta}^n \chi_j = n^{-1} \sum \left(\mathcal{P} \left(\sup_{u \in C_{(l)}} f(j, u) - f(j, 0) \right) \right. \\ \left. - \mathcal{P} \left(\inf_{u \in C_{(l)}} f(j, u) - f(j, 0) \right) \right) \\ \leq p_0 n^{-1} \sum \sup_{u_1, u_2 \in C_{(l)}} |f(j, u_1) - f(j, u_2)| \\ \leq p_0 q^{1/2} \left(\sum_{i=1}^q \left(n^{1/2} d_{in}^{-1}(\theta) \sup_{u \in C_{(l)}} \max_{1 \leq j \leq n} |f_i(j, u)| \right)^2 \right)^{1/2} a(l) \\ \leq k_5 a(l),$$

where

$$k_5 = p_0 q^{1/2} \left(\sum_{i=1}^q (k^{(i)}(1))^2 \right)^{1/2}.$$

Bounds (2.7)–(2.11) show that there are constants k_6 and k_7 such that

$$(2.12) \quad P_{\theta}^n \left\{ \sup_{u \in C_{(l)}} M_n^{(k)}(\theta, u, v) (1 + k_0 n^{1/2} r(l))^{-1} > \frac{\epsilon}{2} \right\} \\ \leq P_{\theta}^n \left\{ k_6 n^{-1} \sum (\chi_j - E_{\theta}^n \chi_j) > \frac{\epsilon}{2} r(l) - k_7 a(l) \right\}.$$

Note that

$$\frac{\epsilon}{2} r(l) - k_7 a(l) = \left(\frac{\epsilon}{2} (1 - t) - k_7 t \right) n^{-\gamma m / \tilde{m}_0} > 0$$

if t is sufficiently small. Using the Chebyshev inequality and (2.11) we estimate the probability on the right-hand side of (2.12) by

$$(2.13) \quad \frac{4k_6^2}{(\epsilon(1 - t) - 2k_7 t)^2} n^{-2+2\gamma m / \tilde{m}_0} \sum E_{\theta}^n \chi_j \leq k_8 n^{-1+\gamma m / \tilde{m}_0}.$$

Let

$$L_{1i}(j) = (f_i(j, v) - f_i(j, 0))(\chi\{X_j < f(j, v)\} - \beta), \\ L_{2i}(j) = f_i(j, 0)(\chi\{X_j < f(j, v)\} - \chi\{\varepsilon_j * 0\}), \\ i = 1, \dots, q.$$

Then

$$(2.14) \quad \begin{aligned} P_1 &= P_\theta^n \left\{ L_n^{(k)}(\theta, v) (1 + k_0 n^{1/2} r(l))^{-1} > \frac{\epsilon}{2} \right\} \\ &\leq \frac{4}{n(k_0 \epsilon)^2 r^2(l)} \sum_{i=1}^q d_{in}^{-2}(\theta) \sum_{\lambda=1}^2 E_\theta^n \left(\sum (L_{\lambda i}(j) - E_\theta^n L_{\lambda i}(j)) \right)^2, \end{aligned}$$

$$(2.15) \quad D_\theta^n \left(\sum L_{1i}(j) \right) \leq \Phi_{2n}^{(i)}(v, 0),$$

$$(2.16) \quad \begin{aligned} D_\theta^n \left(\sum L_{2i}(j) \right) &\leq \sum f_i^2(j, 0) |\mathcal{P}(f(j, v) - f(j, 0)) - \mathcal{P}(0)| \\ &\leq p_0 \max_{1 \leq j \leq n} |g_i(j, \theta)| d_{in}(\theta) \Phi_{2n}^{1/2}(v, 0). \end{aligned}$$

Relations (2.14)–(2.16) and the assumptions of the theorem imply that

$$(2.17) \quad \begin{aligned} P_1 &\leq \frac{4n^{-1}}{(k_0 \epsilon)^2} \left[\frac{(r(l) + a(l))^2}{r^2(l)} \sum_{i=1}^q \tilde{k}^{(i)}(1) + \frac{r(l) + a(l)}{r^2(l)} p_0 k^{1/2}(1) \sum_{i=1}^q k^{(i)}(1) \right] \\ &\leq k_9 n^{-1} \left[(1-t)^{-2} + (1-t)^{-2} n^{\gamma m / \tilde{m}_0} \right] = O \left(n^{-1 + \gamma m / \tilde{m}_0} \right). \end{aligned}$$

Inequalities (2.13) and (2.17) yield for $l = 1, \dots, N_0 - 1$ and some

$$m = m(l) < m_0$$

that

$$(2.18) \quad \sup_{\theta \in T} P_\theta^n \left\{ \sup_{u \in C_{(l)}} z_n^+(\theta, u) > \epsilon \right\} = O \left(n^{-1 + \gamma m / \tilde{m}_0} \right).$$

Consider the case of $l = N_0$. It is obvious that

$$(2.19) \quad \begin{aligned} &P_\theta^n \left\{ \sup_{u \in C_{(N_0)}} z_n^+(\theta, u) > \epsilon \right\} \\ &\leq P_\theta^n \left\{ \sup_{|u|_0 < n^{-\gamma m_0 / \tilde{m}_0}} \left| S_\beta^{*+}(u) - S_\beta^{*+}(0) - E_\theta^n S_\beta^{*+}(u) \right| > \epsilon \right\}. \end{aligned}$$

The random variable whose absolute value stays in the probability on the right-hand side of (2.19) can be represented as the sum of vectors

$$\nu_1(\theta, u) + \nu_2(\theta, u) + \nu_3(\theta, u),$$

where

$$(2.20) \quad \nu_1(\theta, u) = d_n^{-1}(\theta) \sum (\nabla f(j, u) - \nabla f(j, 0)) (\chi\{X_j * f(j, u)\} - \beta),$$

$$(2.21) \quad \nu_2(\theta, u) = d_n^{-1}(\theta) \sum \nabla f(j, 0) (\chi\{X_j * f(j, u)\} - \chi\{\varepsilon_j * 0\}),$$

$$(2.22) \quad \nu_3(\theta, u) = d_n^{-1}(\theta) \sum \nabla f(j, u) (\mathcal{P}(f(j, u) - f(j, 0)) - \beta).$$

It is easy to see that

$$(2.23) \quad |\nu_1(\theta, u)| \leq \beta' n^{1/2} \left(\sum_{i=1}^q d_{in}^{-2}(\theta) \Phi_{2n}^{(i)}(u, 0) \right)^{1/2} \leq k_1 n^{1/2 - \gamma m_0 / \tilde{m}_0},$$

$$(2.24) \quad |\nu_3(\theta, u)| \leq p_0 \Phi_{2n}^{1/2}(u, 0) \left(\sum_{i=1}^q \frac{d_{in}^2(\theta + n^{1/2} d_n^{-1}(\theta) u)}{d_{in}^2(\theta)} \right)^{1/2} \leq k_2 n^{1/2 - \gamma m_0 / \tilde{m}_0}$$

for $|u|_0 < n^{-\gamma m / \tilde{m}_0}$, where k_1 and k_2 are the same numbers as in (2.7) and (2.8), respectively.

If $\gamma > \frac{1}{2}$, then the exponents in (2.20) and (2.21) are negative for $n > n_0$. Thus it remains to estimate the probability

$$\begin{aligned}
 (2.25) \quad & P_\theta^n \left\{ \sup_{|u|_0 < n^{-\gamma m_0/\tilde{m}_0}} |\nu_2(\theta, u)| > \epsilon' \right\} \\
 & \leq P_\theta^n \left\{ \left(\sum_{i=1}^q \left(d_{in}^{-1}(\theta) \max_{1 \leq j \leq n} |g_i(j, \theta)| \right)^2 \right)^{1/2} \sum \tilde{\chi}_j > \epsilon' \right\} \\
 & \leq P_\theta^n \left\{ k_4 n^{-1/2} \sum \tilde{\chi}_j > \epsilon' \right\}, \\
 & \tilde{\chi}_j = \chi \left\{ \inf_{|u|_0 \leq n^{-\gamma m_0/\tilde{m}_0}} f(j, u) - f(j, 0) \leq \varepsilon_j \leq \sup_{|u|_0 \leq n^{-\gamma m_0/\tilde{m}_0}} f(j, u) - f(j, 0) \right\}
 \end{aligned}$$

for $\epsilon' < \epsilon$. It follows from the assumptions of the theorem that

$$\sum E_\theta^n \tilde{\chi}_j \leq k_5 n^{-\gamma m_0/\tilde{m}_0}, \quad j = 1, \dots, n;$$

thus it is sufficient to estimate the probability

$$P_\theta^n \left\{ n^{-1/2} \sum (\tilde{\chi}_j - E_\theta^n \tilde{\chi}_j) > \epsilon'' \right\} \leq (\epsilon'')^2 k_5 n^{-\gamma m_0/\tilde{m}_0}$$

for arbitrary $\epsilon'' > 0$.

Since all the estimates hold uniformly in $\theta \in T$, the lemma is proved for $z_n^+(\theta, u)$. The case of $z_n^-(\theta, u)$ is considered in the same way. $\square$

Put

$$E_\theta^n S_\beta^\pm(\hat{\theta}_n) = (E_\theta^n S_\beta^\pm(\tau))_{\tau=\hat{\theta}_n}.$$

Lemma 2. *If the assumptions of the theorem hold, then*

$$(2.26) \quad \sup_{\theta \in T} P_\theta^n \left\{ |S_\beta^\pm(\theta) + E_\theta^n S_\beta^\pm(\hat{\theta}_n)| > \epsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

for all $\epsilon > 0$.

Proof. Consider the events

$$A_i^\pm(\theta) = \left\{ S_{i\beta}^\pm(\theta) + E_\theta^n S_{i\beta}^\pm(\hat{\theta}_n) - S_{i\beta}^\pm(\hat{\theta}_n) \geq -\epsilon(1 + |E_\theta^n S_\beta^\pm(\hat{\theta}_n)|) \right\}, \quad i = 1, \dots, q.$$

Relation (1.11) and Lemma 1 imply that

$$(2.27) \quad \inf_{\theta \in T} P_\theta^n \{ A_i^\pm(\theta) \} \xrightarrow{n \rightarrow \infty} 1, \quad i = 1, \dots, q.$$

Note that

$$S_\beta^\pm(\hat{\theta}_n) \geq 0$$

for the event $\{|\hat{\theta}_n - \theta| < r\}$, $r < r_0$, whence we deduce that relation (2.24) holds also for the events

$$B_i^\pm(\theta) = \left\{ S_{i\beta}^\pm(\theta) + E_\theta^n S_{i\beta}^\pm(\hat{\theta}_n) \geq -\epsilon(1 + |E_\theta^n S_\beta^\pm(\hat{\theta}_n)|) \right\} \supset A_i^\pm(\theta).$$

On the other hand,

$$S_{i\beta}^+(\theta) + S_{i\beta}^-(\theta) = \sum |g_i(j, \theta)| \chi\{\varepsilon_j = 0\} = 0 \pmod{P_\theta^n}$$

and the events $B_i^-(\theta)$ have the same probabilities as

$$C_i^+(\theta) = \left\{ S_{i\beta}^+(\theta) + E_\theta^n S_{i\beta}^+(\hat{\theta}_n) \leq \epsilon(1 + |E_\theta^n S_\beta^+(\hat{\theta}_n)|) \right\}.$$

Moreover

$$\begin{aligned}
 (2.28) \quad B_i^+(\theta) \cap C_i^+(\theta) &= D_i^+(\theta) = \left\{ \left| S_{i\beta}^+(\theta) + E_{\theta}^n S_{i\beta}^+(\hat{\theta}_n) \right| \leq \epsilon (1 + |E_{\theta}^n S_{i\beta}^+(\hat{\theta}_n)|) \right\}, \\
 &\quad i = 1, \dots, q, \\
 \bigcap_{i=1}^q D_i^+(\theta) &\subseteq \left\{ \left| S_{\beta}^+(\theta) + E_{\theta}^n S_{\beta}^+(\hat{\theta}_n) \right| \leq q\epsilon (1 + |E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)|) \right\} \\
 &\subseteq \left\{ \left| E_{\theta}^n S_{\beta}^+(\hat{\theta}_n) \right| \leq (1 - q\epsilon)^{-1} (q\epsilon + |S_{\beta}^+(\theta)|) \right\} \\
 &= X^+(\theta)
 \end{aligned}$$

for $\epsilon < q^{-1}$, that is,

$$(2.29) \quad \inf_{\theta \in T} P_{\theta}^n \{X^+(\theta)\} \xrightarrow{n \rightarrow \infty} 1.$$

Note that

$$(2.30) \quad P_{\theta}^n \{|E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)| > M\} \leq P_{\theta}^n \{\overline{X^+(\theta)}\} + P_{\theta}^n \{|S_{\beta}^+(\theta)| > M(1 - q\epsilon) - q\epsilon\},$$

where $\overline{X^+(\theta)}$ is the complement of the event $X^+(\theta)$. Let

$$\begin{aligned}
 \eta_j &= \chi\{\varepsilon_j < 0\} - \beta, \quad j \geq 1, \\
 I_{in}(\theta) &= \{1, \dots, n\} \cap \{j : g_i(j, \theta) > 0\}.
 \end{aligned}$$

Then

$$S_{\beta}^+(\theta) - d_{in}^{-1}(\theta) \sum g_i(j, \theta) \eta_j = d_{in}^{-1}(\theta) \sum_{j \in I_{in}(\theta)} g_i(j, \theta) \chi\{\varepsilon_j = 0\} = 0$$

almost surely with respect to P_{θ}^n . By the Chebyshev inequality,

$$P_{\theta}^n \{|S_{\beta}^+(\theta)| > M(1 - q\epsilon) - q\epsilon\} \leq q(M(1 - q\epsilon) - q\epsilon)^{-2} \xrightarrow{M \rightarrow \infty} 0.$$

In other words, the vector $S_{\beta}^+(\theta)$ is bounded in probability. Relations (2.26) and (2.27) mean that the vector $E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)$ also is uniformly in $\theta \in T$ bounded in probability.

It follows from (2.25) that

$$\sup_{\theta \in T} P_{\theta}^n \left\{ \left| S_{\beta}^+(\theta) + E_{\theta}^n S_{\beta}^+(\hat{\theta}_n) \right| > \epsilon \left(1 + |E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)| \right) \right\} \xrightarrow{n \rightarrow \infty} 0.$$

Therefore (2.23) holds.

Note that the boundedness in probability of the random variable $E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)$ can be proved in a different way, namely by using condition **C**, the explicit expression for $E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)$, and the assumptions of the theorem. $\square$

Lemma 3. *If the assumptions of the theorem hold, then*

$$(2.31) \quad P_{\theta}^n \left\{ \left| E_{\theta}^n S_{\beta}^+(\hat{\theta}_n) - p(0)I(\theta)d_n(\theta)(\hat{\theta}_n - \theta) \right| > \epsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

for all $\epsilon > 0$.

Proof. If $n^{-1/2}|d_n(\theta)(\hat{\theta}_n - \theta)|$ is small, then we apply inequality (2.6) and the boundedness in probability of the random variable $E_{\theta}^n S_{\beta}^+(\hat{\theta}_n)$ to prove that the norm of the vector $d_n(\theta)(\hat{\theta}_n - \theta)$ is bounded in probability. Now Lemma 3 follows from condition **C** and inequalities (2.3)–(2.5). $\square$

3. PROOF OF THE THEOREM

Proof. Relations (2.23) and (2.28) show that

$$(3.1) \quad P_\theta^n \left\{ \left| (p(0))^{-1} \Lambda(\theta) S_\beta^+(\theta) + d_n(\theta) (\hat{\theta}_n - \theta) \right| > \epsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

for all $\epsilon > 0$. As above,

$$S_\beta^+(\theta) = d_n^{-1}(\theta) \sum \nabla g(j, \theta) \eta_j \pmod{P_\theta^n}.$$

Now we apply Proposition 17.2 (p. 165 of [6]) to the random vectors

$$\xi_{jn} = n^{1/2} d_n^{-1}(\theta) \nabla g(j, \theta) \eta_j, \quad j = 1, \dots, n.$$

According to condition (1.5),

$$n^{-1} \sum E_\theta^n |\xi_{jn}|^3 \leq q^{1/2} \sum_{i=1}^q n^{-1} \sum d_{in}^{-3}(\theta) |g_i(j, \theta)|^3 n^{3/2} \leq k_{10} < \infty$$

uniformly in $\theta \in T$. Thus

$$(3.2) \quad \sup_{\theta \in T} \sup_{C \in \mathcal{C}^q} \left| P_\theta^n \left\{ \frac{I^{-1/2}(\theta) S_\beta^+(\theta)}{\sqrt{\beta(1-\beta)}} \in C \right\} - \Phi(C) \right| = O(n^{-1/2}).$$

Now we evaluate the correlation matrix of $S_\beta^+(\theta)$. It is obvious that

$$E S_\beta^+(\theta) = 0.$$

Taking into account **A1** we get

$$E_\theta^n S_{i\beta}^+(\theta) S_{l\beta}^+(\theta) = d_{in}^{-1}(\theta) d_{ln}^{-1}(\theta) \sum g_i(j, \theta) g_l(j, \theta) E \eta_j^2, \quad i, l = 1, \dots, q.$$

The definition of η_j yields

$$E \eta_j^2 = \beta(1 - \beta).$$

Then

$$(3.3) \quad E_\theta^n S_\beta^+(\theta) (S_\beta^+(\theta))^T = \beta(1 - \beta) I(\theta).$$

Relations (3.1)–(3.3) imply that

$$(3.4) \quad -\Delta_n + \Phi(C_{-\epsilon}) \leq P_\theta^n \left\{ \frac{p(0)}{\sqrt{\beta(1-\beta)}} I^{1/2}(\theta) d_n(\theta) (\hat{\theta}_n - \theta) \in C \right\} \leq \Delta_n + \Phi(C_\epsilon)$$

for all $\epsilon > 0$ and $C \in \mathcal{C}^q$ where $C_{-\epsilon}$ and C_ϵ are outer and inner sets, parallel to C . Moreover

$$\Delta_n \xrightarrow{n \rightarrow \infty} 0$$

uniformly in $\theta \in T$ and $C \in \mathcal{C}^q$. Now the theorem follows from (3.4) and a theorem in Chapter 3 of [7] saying that

$$\sup_{C \in \mathcal{C}^q} |\Phi(C_{\pm\epsilon}) - \Phi(C)| \leq k\epsilon$$

for all $\epsilon > 0$, where the constant k does not depend on ϵ . $\square$

BIBLIOGRAPHY

1. R. Koenker and G. Bassett, *Regression quantile*, *Econometrica* **46** (1978), 33–50. MR0474644 (57:14279)
2. P. J. Huber, *Robust Statistics*, Wiley, New York, 1981. MR0606374 (82i:62057)
3. A. V. Ivanov, *Asymptotic Theory of Nonlinear Regression*, Kluwer Academic Publishers, Dordrecht, 1997. MR1472234 (99h:62086)
4. I. V. Orlovskii, *Consistency of Koenker–Bassett estimators in nonlinear regression models*, *Naukovi Visti NTUU “KPI”* **3** (2004), 144–152. (Ukrainian)
5. Tae Soo Kim and Hae Kyung Kim, *Consistency of regression quantile estimators for the nonlinear time series regression model*, *Far East J. Theor. Stat.* **5**(1) (2001), 181–191. MR1848451 (2002g:62101)
6. P. J. Huber, *The behavior of maximum likelihood estimates under nonstandard conditions*, *Proc. Fifth Berkeley Sympos. Math. Statist. Probability. I*, Univ. California Press, Berkeley, 1967, 221–234. MR0216620 (35:7449)
7. R. N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansions*, Wiley, New York, 1976. MR0436272 (55:9219)

NATIONAL TECHNICAL UNIVERSITY OF UKRAINE (KPI), PEREMOGY AVENUE 37, KYIV, UKRAINE
E-mail address: ivanov@paligora.kiev.ua

NATIONAL TECHNICAL UNIVERSITY OF UKRAINE (KPI), PEREMOGY AVENUE 37, KYIV, UKRAINE

Received 19/MAR/2004

Translated by OLEG KLESOV