LIMIT THEOREMS FOR EXTREMAL RESIDUALS
IN A REGRESSION MODEL WITH HEAVY TAILS
OF OBSERVATION ERRORS

UDC 519.21

O. V. IVANOV AND I. K. MATSAK

ABSTRACT. Limit theorems for maximal residuals in a linear regression model with
observation errors having heavy tails are obtained.

We consider a linear regression model and study the asymptotic behavior of the max-
imum of deviations between the observations and empirical regression model for the case
where observation errors have heavy tails. We prove the convergence of the distribu-
tion of the normalized maximum to one of the limit max-distributions of independent
identically distributed random variables as the size of the sample tends to infinity.

V. V. Buldygin turned our attention to this type of problem with observations with
heavy tails at a seminar at National Technical University of Ukraine when discussed the
paper [1]. The main results of the current paper can be used in designing statistical tests
for making a decision about the distribution of observation errors and for a goodness of
fit test for a regression model.

Consider the following linear regression model

\( y_j = \sum_{i=1}^{q} \theta_i x_{ji} + \varepsilon_j, \quad j = 1, \ldots, n, \)

where \( \varepsilon_j \) are independent identically distributed random variables.

Let

\[
X = \begin{pmatrix}
  x_{11} & x_{12} & \cdots & x_{1q} \\
x_{21} & x_{22} & \cdots & x_{2q} \\
  \vdots & \vdots & \ddots & \vdots \\
x_{n1} & x_{n2} & \cdots & x_{nq}
\end{pmatrix}
\]

be a \( n \times q \) matrix of design of a regression experiment, \( \det(X^TX) \neq 0 \). Then the
least square estimator of unknown parameters \( \theta^T = (\theta_1, \ldots, \theta_q) \in \mathbb{R}^q \) constructed from
observations is given by

\[
\hat{\theta}_n = (X^TX)^{-1} X^T Y.
\]

The least square estimator is defined as a vector \( \hat{\theta}^T_n = (\hat{\theta}_{1n}, \ldots, \hat{\theta}_{qn}) \) that minimizes the functional

\[
Q(\theta) = (Y - X\theta)^T(Y - X\theta),
\]

2010 Mathematics Subject Classification. Primary 60G70, 62J05.
Key words and phrases. Regression model, extremal values, heavy tails.

©2014 American Mathematical Society
$Y^T = (y_1, \ldots, y_n)$. Put
\[ \hat{y}_j = \sum_{i=1}^{q} \hat{\theta}_{in} x_{ji}, \quad \hat{\varepsilon}_j = y_j - \hat{y}_j, \quad j = 1, \ldots, n, \]

(3)
\[ Z_n = \max_{1 \leq j \leq n} \varepsilon_j, \quad \hat{Z}_n = \max_{1 \leq j \leq n} \hat{\varepsilon}_j, \]
\[ d_{in}^2 = \sum_{j=1}^{n} x_{ji}^2, \quad d_n = \text{diag}(d_{in}, i = 1, \ldots, q). \]

In what follows we will use the following two standard assumptions about the design of a regression experiment. The first assumption is that
\[ (i) \quad d_{in}^{-1} \max_{1 \leq j \leq n} |x_{ji}| \leq k_i n^{-1/2}, \quad i = 1, \ldots, q, \]
for sufficiently large $n$. Put
\[ (ii) \quad \lambda_{\min}(J_n) \geq \lambda_0 > 0. \]

We are interested in studying the limit behavior of the extremal residuals $\hat{Z}_n$ in the model (1). This problem is considered by the authors in [1] for the classical case, namely for the case where $E \varepsilon_j = 0$ and $E \varepsilon_j^2 = \sigma^2 < \infty$ for all $j$. In the current paper, we consider the case of heavy tails, that is, the case of $\sigma^2 = \infty$. Moreover, if the expectation of random variables $\varepsilon_j$ does not exist as well, then the least square estimator $\hat{\theta}_n$ loses the usual statistical properties. Nevertheless, we will see below that some meaningful results follow if one uses formulas (2) in expressions (3).

Let $F(x)$ be the common distribution function of the random variables $(\varepsilon_j)$. Assume that
\[ b_n(Z_n - a_n) \xrightarrow{D} \zeta \]
as $n \to \infty$ for some sequences of nonrandom numbers $b_n > 0$ and $a_n$, where the random variable $\zeta$ possesses a nondegenerate distribution function $G(x) = P(\zeta < x)$. (The symbol $\xrightarrow{D}$ denotes the weak convergence of random variables.)

If relation (7) holds, then we say that a distribution function $F$ belongs to the domain of max-attraction of the law $G$ and write $F \in D(G)$.

It is known (see [2]–[4]) that if relation (7) holds, then the law $G$ belongs to one of the following three types of distributions:

Type I: $\Phi_\alpha(x) = \begin{cases} 0, & \text{for } x \leq 0, \\
\exp(-x^{-\alpha}), & \text{for } x > 0; \end{cases}$

(8)

Type II: $\Psi_\alpha(x) = \begin{cases} \exp(-(x)^\alpha), & \text{for } \alpha > 0, x \leq 0, \\
1, & \text{for } x > 0; \end{cases}$

Type III: $\Lambda(x) = \exp(-e^{-x}), \quad \text{for } -\infty < x < \infty.$

Let $x_F = \sup \{x : F(x) < 1\}$. If $F \in D(\Psi_\alpha)$, then $x_F < \infty$ (see [3], Chapter 4, §6]), whence
\[ E(\varepsilon_j)^m_+ < \infty, \quad m > 0. \]
If a distribution function $F$ belongs to the domain of max-attraction of a law $\Lambda$, then condition (9) holds as well (see [3, Theorem 2.7.3]).

Therefore, only Type I of distribution functions requires a closer look in the case of heavy tails. In order that $F \in D(\Phi_\alpha)$ it is necessary and sufficient that

\[(10)\quad x_F = \infty \text{ and } \lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \quad x > 0, \quad \alpha > 0.\]

In this case, the constants in relation (7) can be chosen as follows (see [4, Corollary 1.6.3]):

\[(11)\quad a_n = 0, \quad b_n = (\gamma_n)^{-1}, \quad \gamma_n = F^{-1} \left( 1 - \frac{1}{n} \right).\]

**Theorem 1.** Let the distribution of random variables $\varepsilon_j$ in the model (1) be symmetric and let their common distribution function $F(x)$ satisfy condition (10) for some $\alpha \in (0, 2]$. Also let conditions (i) and (ii) hold. Then

\[(12)\quad \lim_{n \to \infty} P \left\{ \frac{\hat{Z}_n}{\gamma_n} < x \right\} = \Phi_\alpha(x), \quad x \in \mathbb{R}.\]

The following auxiliary result is used in the proof of Theorem 1.

**Lemma 1.** Let $0 < \beta < \infty$ and let $\xi_j, j \geq 1$ be independent symmetric random variables such that $E|\xi_j|^\beta < +\infty$. Then there exists a constant $C_\beta < +\infty$ such that

\[(13)\quad E\left( \sum_{j=1}^{n} \xi_j \right)^\beta \leq C_\beta E \left( \sum_{j=1}^{n} |\xi_j|^\beta \right)^{\beta/2}.\]

**Proof of Lemma 1.** Inequality (13) is a partial case of the Marcinkiewicz–Zygmund inequality if $\beta \geq 1$ ([5, Chapter 3, §5, Proposition 14]). Moreover, relation (13) coincides with the well-known Khintchine inequality if $\chi_j$ are independent symmetric Bernoulli random variables, that is, if $P(\chi_j = \pm 1) = \frac{1}{2}$; $c_j, j \geq 1$, are real numbers; and $\xi_j = \chi_j c_j$ (see [6]).

The symmetric random variables $\xi_j$ can be represented as follows

\[\xi_j \overset{d}{=} \chi_{j}\xi'_j,\]

where $\xi_j \overset{d}{=} \xi'_j$ and $\chi_j$ and $\xi'_j$ are independent random variables. Thus the general case of inequality (13) follows from the Khintchine inequality with $\xi'_j$ instead of $c_j$. \(\square\)

**Proof of Theorem 1.** Below we follow some ideas introduced in [1].

Condition (6) implies that the inverse matrix $J_n^{-1} = \mathcal{L}_n = (\mathcal{L}^{kl}_n)_{k,l=1}^{q}$ (for $q > 0$) exists and moreover

\[\det J_n = \lambda_1 \cdots \lambda_q \geq \lambda_{\min}^q \geq \lambda_0^q > 0,\]

\[\det \mathcal{L}_n \leq \lambda_0^{-q}, \quad n > n_0,\]

where $\lambda_1, \ldots, \lambda_q$ are eigenvalues of the matrix $J_n$.

In the matrix constituted from cofactors related to the matrix $J_n$, each entry is a product of entries of $J_n$ and note also that each entry of $J_n$ is bounded from above by 1 in view of the Cauchy–Bunyakovskii inequality. Since each minor contains $(q - 1)!$ terms, condition (ii) implies that every entry of the matrix $\mathcal{L}_n$ is bounded by

\[(14)\quad |\mathcal{L}^{kl}_n| \leq \lambda_0^{-q}(q - 1)!, \quad k, l = 1, \ldots, q.\]

Let $0 < \beta < \alpha$. A distribution function $F(x)$ that satisfies condition (10) can be written as follows:

\[F(x) = 1 - x^{-\alpha}L(x), \quad x > 1,\]
where \( L(x) \) is a slowly varying function. Thus \( \mathbb{E} |\varepsilon_j|^\beta < +\infty \) (see, for example, [1] Chapter 8, §8).

Now the inequality

\[
\max_{1 \leq j \leq n} |c_j - \max_{1 \leq j \leq n} d_j| \leq \max_{1 \leq j \leq n} |c_j - d_j|
\]

implies that

\[
\mathbb{E} |\hat{Z}_n - Z_n|^\beta \leq \mathbb{E} \max_{1 \leq j \leq n} |\hat{\varepsilon}_j - \varepsilon_j|^\beta
\]

(15)

\[
\leq \mathbb{E} \left( \sum_{i=1}^{q} |\hat{\theta}_i - \theta_i| \max_{1 \leq j \leq n} |x_{ji}| \right)^\beta
\]

\[
\leq k(\beta) \sum_{i=1}^{q} \mathbb{E} |\hat{\theta}_i - \theta_i|^\beta \max_{1 \leq j \leq n} |x_{ji}|,
\]

where \( k(\beta) = 1 \) if \( 0 < \beta < \alpha \leq 1 \) and \( k(\beta) = q^{\beta-1} \) if \( 1 \leq \beta < \alpha \leq 2 \).

Equality (2) together with inequality (13) yields

\[
\mathbb{E} |\hat{\theta}_i - \theta_i|^\beta = \mathbb{E} \left[ \sum_{j=1}^{n} \left( d_{in}^{-1} \sum_{l=1}^{q} \mathcal{L}_{il} d_{ln}^{-1} x_{jl} \right) \varepsilon_j \right]^{\beta}
\]

(16)

\[
\leq \mathbb{E} \left[ \sum_{j=1}^{n} \left( d_{in}^{-1} \sum_{l=1}^{q} \mathcal{L}_{il} d_{ln}^{-1} x_{jl} \right)^2 \varepsilon_j^2 \right]^{\beta/2}
\]

\[
\leq \mathbb{E} \max_{1 \leq j \leq n} |\varepsilon_j|^{\beta} d_{in}^{-\beta} \left( \sum_{j=1}^{n} \left( \sum_{l=1}^{q} \mathcal{L}_{il} d_{ln}^{-1} x_{jl} \right)^2 \right)^{\beta/2}
\]

for \( i = 1, \ldots, q \).

Using bound (14) we get

\[
\sum_{j=1}^{n} \left( \sum_{l=1}^{q} \mathcal{L}_{il} d_{ln}^{-1} x_{jl} \right)^2 \leq q \sum_{j=1}^{n} \sum_{l=1}^{q} (\mathcal{L}_{il})^2 d_{ln}^{-2} x_{jl}^2 \leq (q!)^2 \lambda_0^{-2q}
\]

and thus we can continue the estimation in (16) as follows:

(17)

\[
\mathbb{E} |\hat{\theta}_i - \theta_i|^\beta \leq (q!)^\beta \lambda_0^{-\beta q} d_{in}^{-\beta} \mathbb{E} \max_{1 \leq j \leq n} |\varepsilon_j|^\beta.
\]

Combining this result with (15) and (4) we obtain

(18)

\[
\mathbb{E} |\hat{Z}_n - Z_n|^\beta \leq k(\beta)(q!)^\beta \lambda_0^{-\beta q} \left( \sum_{i=1}^{q} \theta_i^\beta \right) \left( \mathbb{E} \max_{1 \leq j \leq n} |\varepsilon_j|^\beta \right) n^{-\beta/2}.
\]

Put

\[
Z_n^* = \max_{1 \leq j \leq n} |\varepsilon_j|, \quad W_n = \min_{1 \leq j \leq n} \varepsilon_j.
\]

Then

\[
P\{Z_n^* > x\} = P\{Z_n > x, W_n < -x\} \leq P\{Z_n > x\} + P\{W_n < -x\}
\]

for \( x > 0 \). Since the random variables \( \varepsilon_j \) are symmetric, \( P\{W_n < -x\} = P\{Z_n > x\} \).
Hence
\begin{equation}
E(Z_n^+)^\beta = \beta \int_0^\infty x^{\beta-1} P\{Z_n^+ > x\} \, dx \leq 2\beta \int_0^\infty x^{\beta-1} P\{Z_n > x\} \, dx
\end{equation}
(19)
\begin{equation}
= 2\beta E(Z_n)^\beta_+.
\end{equation}
Here and in what follows
\begin{align*}
x_+ &= \max(x,0), \quad x_- = \max(-x,0)
\end{align*}
for all \( x \in \mathbb{R}^1 \).

If \( F \in D(\Phi_\alpha) \) and \( \beta < \alpha \), then
\begin{equation}
\lim_{n \to \infty} E \left( \frac{Z_n}{\gamma_n} \right)^\beta_+ \int_0^\infty x^{\beta} \, d\Phi_\alpha(x) < +\infty.
\end{equation}
(20)

Taking into account (19) and (20), we deduce from (18) that
\begin{equation}
E \left| \frac{\hat{Z}_n - Z_n}{\gamma_n} \right|^\beta = O \left( n^{-\beta/2} \right).
\end{equation}
(21)

Under assumptions of Theorem 1, the random variables \( Z_n \) satisfy equality (12). Therefore, relation (12) holds for \( \hat{Z}_n \), as well. \( \square \)

The asymptotic behavior of the normalizing sequence \( \gamma_n \) can easily be established from condition (10). Indeed, since \( 1 - F(x) = x^{-\alpha} L(x), \, x > 1 \), where \( L(x) \) is a slowly varying at infinity function, \( \gamma_n \) is a solution of the equation
\[ \gamma_n^{-\alpha} L(\gamma_n) = \frac{1}{n}, \quad \text{or} \quad \gamma_n^\alpha L_1(\gamma_n) = n, \]
where \( L_1(x) = L(x)^{-1} \) (see (11)).

Consider the function \( U_1(x) = x^\alpha L_1(x) \). It is known (see [9, Chapter 1, §5, Proposition 5]) that there exists a slowly varying function \( L_2(x) \) such that
\[ U_1(U_2(x)) \sim U_2(U_1(x)) \sim x, \quad x \to \infty, \]
where \( U_2(x) = x^1/\alpha L_2(x) \). Then
\begin{equation}
n^{1/\alpha} L_2(n) = U_2(n) = U_2(U_1(\gamma_n)) \sim \gamma_n
\end{equation}
as \( n \to \infty \).

**Example 1.** (see [3, Example 2.3.4] or [4, Example 1.7.8]) Let the random variables \( (\varepsilon_j) \) possess the Cauchy distribution, that is, \( F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x \). It can be easily checked that
\[ \lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1}. \]
Thus the distribution \( \Phi_1 \) defined by [3] is the limit distribution for \( \hat{Z}_n/\gamma_n \). Further,
\begin{align*}
1 - F(\gamma_n) &= \frac{1}{n}, \quad \gamma_n = \tan \left( \frac{\pi}{2} - \frac{\pi}{n} \right) = nL_2(n),
L_2(x) = \frac{1}{x \tan \left( \frac{\pi}{2} \right)} \sim \frac{1}{\pi}.
\end{align*}
This means that one can choose \( \gamma_n = n/\pi \).

Put \( \hat{Z}_n^* = \max_{1 \leq j \leq n} |\hat{\varepsilon}_j| \).
Theorem 2. Under the assumptions of Theorem 1,

\[
\lim_{n \to \infty} P \left\{ \frac{\hat{Z}_n^*}{\gamma_n} < x \right\} = \Phi_\alpha^2(x), \quad x \in \mathbb{R}^1.
\]

Proof of Theorem 2. Put \( u_n = \gamma_n x \). Then

\[
nF(-u_n) = n(1 - F(u_n)) = \frac{1 - F(u_n)}{1 - F(\gamma_n)} \to x^{-\alpha}, \quad n \to \infty.
\]

Here we used Theorem 1.8.2 of [4]:

\[
P\{ Z_n^* < u_n \} = P\{ Z_n < u_n, W_n > -u_n \} \to \Phi_\alpha^2(x)
\]
as \( n \to \infty \).

To complete the proof of (23) we write

\[
|Z_n^* - \hat{Z}_n^*| \leq \max_{1 \leq j \leq n} |\varepsilon_j - \hat{\varepsilon}_j|
\]
and repeat the same reasoning as in the proof of Theorem 1. \(\Box\)

Remark 1. It is clear that Theorems 1 and 2 remain true if a weaker condition

\[
(iii) \quad \lim_{n \to \infty} d_{in}^{-1} \max_{1 \leq j \leq n} |x_{ji}| = 0, \quad i = 1, \ldots, q
\]
substitutes condition (4). On the other hand, condition (i) allows one to estimate the rate of convergence to zero of the expectations of normalized deviations

\[
E \left( \frac{|\hat{Z}_n - Z_n|}{\gamma_n} \right)^\beta, \quad E \left( \frac{\hat{Z}_n^* - Z_n^*}{\gamma_n} \right)^\beta.
\]

Consider the simple linear regression model

\[
y_j = \theta_1 + \theta_2 x_j + \varepsilon_j, \quad j = 1, \ldots, n,
\]
where \( \varepsilon_j \) are independent identically distributed random variables.

Since (25) is a particular case of the model (11), Theorems 1 and 2 hold for the simple linear regression model, too. Of course, conditions (ii) and (iii) should be rewritten for model (25) in a specific form. For model (25), we have

\[
d_{1n}^2 = n, \quad d_{2n}^2 = \sum_{j=1}^n x_j^2, \quad v_n = n^{-1/2} d_{2n}^{-1} \sum_{j=1}^n x_j,
\]
\[
J_n = \begin{pmatrix} 1 & v_n \\ v_n & 1 \end{pmatrix},
\]
and thus (ii) holds if \( |v_n| \leq 1 - \delta \) for some \( \delta > 0 \). In turn, condition (iii) for the simple linear regression model is rewritten as follows

\[
\lim_{n \to \infty} d_{2n}^{-1} \max_{1 \leq j \leq n} |x_j| = 0.
\]

On the other hand, one can prove Theorems 1 and 2 for model (25) independently of the general case. Moreover, one can weaken to some extent the sufficient conditions in the case of a linear regression model. The least square estimator of unknown parameters \( (\theta_1, \theta_2) \) for this case is given by

\[
\hat{\theta}_{2n} = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^n (x_j - \bar{x})^2}, \quad \hat{\theta}_{1n} = \bar{y} - \hat{\theta}_{2n} \bar{x},
\]
\[
\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j, \quad \bar{y} = \frac{1}{n} \sum_{j=1}^n y_j.
\]
Note that \( \hat{y}_j = \hat{\theta}_{1n} + \hat{\theta}_{2n} x_j \). The other notation is not changed.

**Theorem 3.** Let random variables \( \varepsilon_j \) be symmetric in the model (25). Assume that their common distribution function \( F \) belongs to \( D(\Phi_{\alpha}) \) for some \( \alpha \in (0, 2] \) and

\[
\frac{\max_{1 \leq j \leq n} |x_j|}{\sqrt{\sum_{j=1}^{n} (x_j - \bar{x})^2}} \to 0, \quad n \to \infty.
\]

Then relations (12) and (23) hold.

**Proof of Theorem 3.** Since

\[
\hat{\varepsilon}_j = y_j - \hat{y}_j = (\theta_1 - \hat{\theta}_{1n}) + (\theta_2 - \hat{\theta}_{2n}) x_j + \varepsilon_j,
\]

we have

\[
|Z_n - \hat{Z}_n| \leq \max_{1 \leq j \leq n} |\varepsilon_j - \hat{\varepsilon}_j| \leq |\theta_1 - \hat{\theta}_{1n}| + |\theta_2 - \hat{\theta}_{2n}| \max_{1 \leq j \leq n} |x_j|.
\]

We are going to estimate the absolute value of the difference \( \theta_2 - \hat{\theta}_{2n} \). It is clear that

\[
y_j - \bar{y} = \theta_2 (x_j - \bar{x}) + \varepsilon_j - \bar{\varepsilon},
\]

where \( \bar{\varepsilon} \) is the average of errors,

\[
\bar{\varepsilon} = n^{-1} \sum_{j=1}^{n} \varepsilon_j.
\]

Thus

\[
\hat{\theta}_{2n} = \frac{\sum_{j=1}^{n} (x_j - \bar{x}) (\theta_2 (x_j - \bar{x}) + \varepsilon_j - \bar{\varepsilon})}{\sum_{j=1}^{n} (x_j - \bar{x})^2} = \theta_2 + \frac{\sum_{j=1}^{n} (x_j - \bar{x}) \varepsilon_j}{\sum_{j=1}^{n} (x_j - \bar{x})^2}.
\]

Let \( 0 < \beta < \alpha \). Equality (27) implies that

\[
E |\theta_2 - \hat{\theta}_{2n}|^\beta \leq E \left( \frac{\sum_{j=1}^{n} (x_j - \bar{x})^2 \varepsilon_j^2}{(\sum_{j=1}^{n} (x_j - \bar{x})^2)^2} \right)^{\beta/2} \leq \frac{\max_{1 \leq j \leq n} |\varepsilon_j|^\beta}{(\sum_{j=1}^{n} (x_j - \bar{x})^2)^{\beta/2}}
\]

according to Lemma 1. Taking into account relations (19), (20), (26), and (28), we obtain

\[
E \left| \frac{\theta_2 - \hat{\theta}_{2n}}{\gamma_n} \right|^\beta \max_{1 \leq j \leq n} |x_j|^\beta \leq C \left( \frac{\max_{1 \leq j \leq n} |x_j|}{\sqrt{\sum_{j=1}^{n} (x_j - \bar{x})^2}} \right)^{\beta} \to 0, \quad n \to \infty.
\]

It remains to estimate \( |\theta_1 - \hat{\theta}_{1n}| \):

\[
|\theta_1 - \hat{\theta}_{1n}| = |\theta_1 - (\bar{y} - \hat{\theta}_{2n} \bar{x})| \leq |\theta_2 - \hat{\theta}_{2n}| \cdot |\bar{x}| + |\bar{\varepsilon}|.
\]

Since \( |\bar{x}| \leq \max_{1 \leq j \leq n} |x_j| \), relation (29) implies that

\[
E \left| \frac{(\theta_2 - \hat{\theta}_{2n}) \bar{x}}{\gamma_n} \right|^\beta \to 0, \quad n \to \infty.
\]

On the other hand, if \( 0 < \beta < \alpha \leq 1 \), then relation (22) implies that

\[
E \left( \frac{|\bar{\varepsilon}|}{\gamma_n} \right)^\beta \leq \frac{n^{1-\beta}}{\gamma_n} \leq \frac{E |\varepsilon_1|^\beta}{\gamma_n} = \frac{1}{L_2^{\beta}(n)} \frac{1}{n^{\beta/\alpha+\beta-1}}.
\]

If \( \beta > \alpha/(1 + \alpha) \), then \( \beta/\alpha + \beta - 1 > 0 \) and thus

\[
E (|\bar{\varepsilon}|/\gamma_n)^\beta \to 0
\]

as \( n \to \infty \).
If $1 \leq \beta < \alpha \leq 2$, then $\bar{\varepsilon} \to 0$ almost surely as $n \to \infty$ by the strong law of large numbers for independent identically distributed random variables and thus relation (12) is proved.

Now the proof of (20) is straightforward. \(\Box\)

**Remark 2.** Let condition (iii) hold and $v_n^2 \geq 1 - \delta$, $\delta > 0$. Then

\[
\sum_{j=1}^{n}(x_j - \bar{x})^2 = \sum_{j=1}^{n}x_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n}x_j\right)^2 = d_{2n}^2(1 - v_n^2) \geq \delta d_{2n}^2,
\]

\[
\frac{\max_{1 \leq j \leq n}|x_j|}{\sqrt{\sum_{j=1}^{n}(x_j - \bar{x})^2}} \leq \delta^{-1/2}d_{2n}^{-1}\max_{1 \leq j \leq n}|x_j| \to 0, \quad n \to \infty.
\]

This shows that condition (26) is indeed weaker than (ii) or (iii).

Below we consider the case of a nonsymmetric distribution of observation errors ($\varepsilon_j$). It is clear that one needs to impose some additional conditions on the left tail of the distribution of random variables ($\varepsilon_j$) in this case.

**Theorem 4.** Let $F(x)$ be the distribution function of the random variables $\varepsilon_j$ in the model (1). Assume that $F$ satisfies condition (10) for some $\alpha \in [1, 2]$ and $E\varepsilon_j = 0$, $E\varepsilon_j^2 = \infty$. Also let conditions (i) and (ii) hold. Then equality (12) is satisfied if $1 \leq \alpha < 2$ or if $\alpha = 2$ and

\[
E(\varepsilon_j)^\beta < \infty
\]

for some $\beta > 1$.

**Proof of Theorem 4.** Examining the proof of bound (18) in Theorem 1, we see that the assumption that the distribution function of the random variables $\varepsilon_j$ is symmetric is needed only to make sure that Lemma 1 applies. Under the assumptions $E\varepsilon_j = 0$ and $\beta \geq 1$, Lemma 1 remains true for the nonsymmetric case, too (due to the Marcinkiewicz–Zygmund inequality). Thus bound (18) holds under the assumptions of Theorem 4 if $\beta \geq 1$.

Now assume that $1 \leq \alpha < 2$ and let $\beta = 1$. Then inequality (18) is rewritten as follows

\[
E|\hat{Z}_n - Z_n| \leq (q!)\lambda_0^{-q}\left(\sum_{i=1}^{q}k_i\right)\left(E\max_{1 \leq j \leq n}|\varepsilon_j|\right)n^{-1/2}.
\]

Considering (22) we get

\[
E\max_{1 \leq j \leq n}|\varepsilon_j| \leq \frac{nE|\varepsilon_1|}{\gamma_n} = \frac{E|\varepsilon_1|}{L_2(n)n^{1/\alpha-1}}.
\]

The latter result together with (31) implies that

\[
E|\hat{Z}_n - Z_n| \leq C\frac{E|\varepsilon_1|}{L_2(n)n^{1/\alpha+1/2-1}} \to 0
\]
as $n \to \infty$. Now we establish equality (12) in the same way as in the proof of Theorem 4.

It remains to consider the case of $\alpha = 2$. Choose a number $1 < \beta < 2$ such that condition (30) holds. Then (18) implies that

\[
E|\hat{Z}_n - Z_n|^\beta \leq C\frac{E\max_{1 \leq j \leq n}|\varepsilon_j|^\beta}{\gamma_n^\beta}n^{-\beta/2}
\]

(32)

\[
\leq C\frac{E\max_{1 \leq j \leq n}(\varepsilon_j)^\beta + E\max_{1 \leq j \leq n}(\varepsilon_j)^\beta}{\gamma_n^\beta}n^{-\beta/2}.
\]
According to equality (20),
\[
E \max_{1 \leq j \leq n} (\varepsilon_j)_{\beta n} \leq \beta \gamma n^{-1/2} \to 0
\]
as \( n \to \infty \). Further,
\[
E \max_{1 \leq j \leq n} (\varepsilon_j)_{\beta n} \leq \frac{n E(\varepsilon_1)_{\beta n}}{n^2 L_2^\beta(n)} \to 0.
\]
Combining relations (32)–(34) we complete the proof of Theorem 4.

Consider an application of the above results. Assume that all the assumptions imposed above hold for the simple linear regression model (25). We want to test the null hypothesis
\[
H_0 = \{ F \text{ is the Cauchy distribution function} \}
\]
under the alternatives
\[
H_1^+ = \{ F \in D(\Phi_\alpha), \alpha > 1 \}, \quad H_1^- = \{ F \in D(\Phi_\alpha), \alpha < 1 \}.
\]
If all the assumptions of Theorem 3 are valid for \((x_i)\), then we get the following rule to test the null hypothesis \(H_0\).

As a test statistic, we choose
\[
T_n = \pi Z_1^*/n.
\]
Given a level of significance \(\delta\), the rejection regions are given by
\[
K^+ = \{ x : 0 < x < U_{\delta/2} \}, \quad K^- = \{ x : U_{1-\delta/2} < x \}, \quad K = K^+ \cup K^-,
\]
where \(U_p\) are the \(p\)-quantiles for the distribution function \(\Phi_1(x)\). For example, if \(\delta = 0.05\), then
\[
K^+ = \left(0, \frac{2}{\ln(2/\delta)}\right) \approx (0, 0.54), \quad K^- = \left(\frac{2}{\ln(1/(1-\delta/2))}, \infty\right) \approx (79.0, \infty).
\]
At first glance, the complement region \(R_1 \setminus K\) seems to be too wide. On the other hand, one should take into account that the right tail of the distribution function \(\Phi_1^2(x)\) is heavy.

If \(T_n \in R_1 \setminus K\), then the hypothesis \(H_0\) is accepted; otherwise
\[
\begin{align*}
&\text{If } T_n \in K^+, \quad \text{then } H_1^+ \text{ is accepted;} \\
&\text{If } T_n \in K^-, \quad \text{then } H_1^- \text{ is accepted.}
\end{align*}
\]

Bibliography


National Technical University of Ukraine “Kyiv Polytechnic Institute”, Department of Mathematical Analysis and Probability Theory, Peremogy Square, 37, Kyiv 03056, Ukraine

E-mail address: alexntuu@gmail.com

Kyiv National Taras Shevchenko University, Faculty for Cybernetics, Glushkov Avenue 2, Building 6, Kyiv 03127, Ukraine

E-mail address: mik@unicyb.kiev.ua

Received 17/JUL/2012
Translated by S. KVASKO