Теорія Ймовір. та Матем. Статист.
Вип. 88, 2013

Theor. Probability and Math. Statist.
No. 88, 2014, Pages 99–108
S 0094-9000(2014)00921-7
Article electronically published on July 24, 2014

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 maximum of deviations between the observations and empirical regression model for the case where observation errors have heavy tails. We prove the convergence of the distribution 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

$$y_j = \sum_{i=1}^{q} \theta_i x_{ji} + \varepsilon_j, \qquad j = 1, \dots, n, \tag{1}$$

where ε_j are independent identically distributed random variables.

Let

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix}$$

be a $n \times q$ matrix of design of a regression experiment, $\det(X^T X) \neq 0$. Then the least square estimator of unknown parameters $\theta^T = (\theta_1, \dots, \theta_q) \in \mathbf{R}^q$ constructed from observations (1) is given by

$$\hat{\theta}_n = \left(X^T X\right)^{-1} X^T Y. \tag{2}$$

The least square estimator is defined as a vector $\hat{\theta}_n^T = (\hat{\theta}_{1n}, \dots, \hat{\theta}_{qn})$ 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.

$Y^T = (y_1, \dots, y_n)$. Put

$$\hat{y_j} = \sum_{i=1}^{q} \hat{\theta}_{in} x_{ji}, \qquad \hat{\varepsilon}_j = y_j - \hat{y_j}, \qquad j = 1, \dots, n,$$

$$Z_n = \max_{1 \le j \le n} \varepsilon_j, \qquad \hat{Z}_n = \max_{1 \le j \le n} \hat{\varepsilon}_j, \tag{3}$$

$$d_{in}^2 = \sum_{j=1}^{n} x_{ji}^2, \qquad d_n = \operatorname{diag}(d_{in}, i = 1, \dots, q).$$

In what follows we will use the following two standard assumptions about the design of a regression experiment. The first assumption is that

$$(i) \qquad d_{in}^{-1} \max_{1 \le j \le n} |x_{ji}| \le k_i n^{-1/2}, \qquad i = 1, \dots, q, \tag{4}$$

for sufficiently large n. Put

$$J_n = d_n^{-1} X^T X d_n^{-1} = \left(d_{kn}^{-1} d_{ln}^{-1} \sum_{j=1}^{n} x_{jk} x_{jl} \right)_{k,l=1}^{q}, \tag{5}$$

where $\lambda_{\min}(A)$ is the least eigenvalue of the positive definite matrix A. The second assumption is that

$$(ii) \qquad \lambda_{\min}(J_n) \ge \lambda_0 > 0. \tag{6}$$

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 $\mathsf{E}\,\varepsilon_j = 0$ and $\mathsf{E}\,\varepsilon_j^2 = \sigma^2 < \infty$ 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 ε_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 (ε_j). Assume that

$$b_n(Z_n - a_n) \xrightarrow{D} \zeta \tag{7}$$

as $n \to \infty$ for some sequences of nonrandom numbers $b_n > 0$ and a_n, where the random variable ζ possesses a nondegenerate distribution function $G(x) = \mathsf{P}(\zeta < x)$. (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:

$$\begin{aligned} &\text{Type I:} \quad \Phi_\alpha(x) = \begin{cases} 0, & \text{for } x \le 0, \\ \exp(-x^{-\alpha}), & \text{for } \alpha > 0,\ x > 0\,; \end{cases} \\ &\text{Type II:} \quad \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & \text{for } \alpha > 0,\ x \le 0, \\ 1, & \text{for } x > 0; \end{cases} \\ &\text{Type III:} \quad \Lambda(x) = \exp(-e^{-x}), \quad \text{for } -\infty < x < \infty. \end{aligned} \tag{8}$$

Let $x_F = \sup\{x \colon F(x) < 1\}$. If $F \in D(\Psi_\alpha)$, then $x_F < \infty$ (see [4, Chapter 4, §6]), whence

$$\mathsf{E}(\varepsilon_j)_+^m < \infty, \qquad m > 0. \tag{9}$$

If a distribution function F belongs to the domain of max-attraction of a law Λ, 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

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

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

$$a_n = 0, \qquad b_n = (\gamma_n)^{-1}, \qquad \gamma_n = F^{-1}\left(1 - \frac{1}{n}\right). \tag{11}$$

Theorem 1. *Let the distribution of random variables ε_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*

$$\lim_{n\to\infty} \mathsf{P}\left\{\frac{\hat{Z}_n}{\gamma_n} < x\right\} = \Phi_\alpha(x), \qquad x \in \mathbf{R}^1. \tag{12}$$

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

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

$$\mathsf{E}\left|\sum_{j=1}^n \xi_j\right|^\beta \leq C_\beta\, \mathsf{E}\left(\sum_{j=1}^n |\xi_j|^2\right)^{\beta/2}. \tag{13}$$

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 χ_j are independent symmetric Bernoulli random variables, that is, if $\mathsf{P}(\chi_j = \pm 1) = \frac{1}{2}$; c_j, $j \geq 1$, are real numbers; and $\xi_j = \chi_j c_j$ (see [6]).

The symmetric random variables ξ_j can be represented as follows

$$\xi_j \stackrel{d}{=} \chi_j \xi_j^{'},$$

where $\xi_j \stackrel{d}{=} \xi_j^{'}$ and χ_j and ξ_j' are independent random variables. Thus the general case of inequality (13) follows from the Khintchine inequality with $\xi_j^{'}$ instead of c_j. □

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

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

$$\det J_n = \lambda_1 \ldots \lambda_q \geq \lambda_{\min}^q \geq \lambda_0^q > 0,$$
$$\det \mathcal{L}_n \leq \lambda_0^{-q}, \qquad n > n_0,$$

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. Note also that each entry of J_n is bounded from above by 1 in view of the Cauchy–Bunyakovskiĭ inequality. Since each minor contains $(q-1)!$ terms, condition (ii) implies that every entry of the matrix $\mathcal{L}_n$ is bounded by

$$\left|\mathcal{L}_n^{kl}\right| \leq \lambda_0^{-q}(q-1)!, \qquad k, l = 1, \ldots, q. \tag{14}$$

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

$$F(x) = 1 - x^{-\alpha} L(x),\ x > 1,$$

where $L(x)$ is a slowly varying function. Thus $\mathsf{E}\,|\varepsilon_j|^\beta < +\infty$ (see, for example, [7, Chapter 8, §8]).

Now the inequality

$$\left|\max_{1\le j\le n} c_j - \max_{1\le j\le n} d_j\right| \le \max_{1\le j\le n} |c_j - d_j|$$

implies that

$$\begin{aligned}\mathsf{E}\,|\hat{Z}_n - Z_n|^\beta &\le \mathsf{E}\max_{1\le j\le n} |\hat{\varepsilon}_j - \varepsilon_j|^\beta \\ &\le \mathsf{E}\left(\sum_{i=1}^q |\hat{\theta}_{in} - \theta_i| \max_{1\le j\le n} |x_{ji}|\right)^\beta \\ &\le k(\beta)\sum_{i=1}^q \mathsf{E}\,|\hat{\theta}_{in} - \theta_i|^\beta \max_{1\le j\le n} |x_{ji}|^\beta,\end{aligned} \tag{15}$$

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

Equality (2) together with inequality (13) yields

$$\begin{aligned}\mathsf{E}\,|\hat{\theta}_i - \theta_i|^\beta &= \mathsf{E}\left|\sum_{j=1}^n \left(d_{in}^{-1}\sum_{l=1}^q \mathcal{L}_n^{il} d_{ln}^{-1} x_{jl}\right)\varepsilon_j\right|^\beta \\ &\le \mathsf{E}\left|\sum_{j=1}^n \left(d_{in}^{-1}\sum_{l=1}^q \mathcal{L}_n^{il} d_{ln}^{-1} x_{jl}\right)^2 \varepsilon_j^2\right|^{\beta/2} \\ &\le \mathsf{E}\max_{1\le j\le n} |\varepsilon_j|^\beta d_{in}^{-\beta}\left(\sum_{j=1}^n\left(\sum_{l=1}^q \mathcal{L}_n^{il} d_{ln}^{-1} x_{jl}\right)^2\right)^{\beta/2}\end{aligned} \tag{16}$$

for $i = 1, \dots, q$.

Using bound (14) we get

$$\sum_{j=1}^n\left(\sum_{l=1}^q \mathcal{L}_n^{il} d_{ln}^{-1} x_{jl}\right)^2 \le q\sum_{j=1}^n\sum_{l=1}^q (\mathcal{L}_n^{il})^2 d_{ln}^{-2} x_{jl}^2 \le (q!)^2\lambda_0^{-2q}$$

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

$$\mathsf{E}\,|\hat{\theta}_i - \theta_i|^\beta \le (q!)^\beta \lambda_0^{-\beta q} d_{in}^{-\beta}\,\mathsf{E}\max_{1\le j\le n} |\varepsilon_j|^\beta. \tag{17}$$

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

$$\mathsf{E}\,|\hat{Z}_n - Z_n|^\beta \le k(\beta)(q!)^\beta\lambda_0^{-\beta q}\left(\sum_{i=1}^q k_i^\beta\right)\left(\mathsf{E}\max_{1\le j\le n} |\varepsilon_j|^\beta\right) n^{-\beta/2}. \tag{18}$$

Put

$$Z_n^* = \max_{1\le j\le n} |\varepsilon_j|, \qquad W_n = \min_{1\le j\le n} \varepsilon_j.$$

Then

$$\mathsf{P}\{Z_n^* > x\} = \mathsf{P}\{Z_n > x, W_n < -x\} \le \mathsf{P}\{Z_n > x\} + \mathsf{P}\{W_n < -x\}$$

for $x > 0$. Since the random variables ε_j are symmetric, $\mathsf{P}\{W_n < -x\} = \mathsf{P}\{Z_n > x\}$.

Hence

$$\begin{aligned}\mathsf{E}(Z_n^*)^\beta &= \beta\int_0^\infty x^{\beta-1}\,\mathsf{P}\{Z_n^*>x\}\,dx \le 2\beta\int_0^\infty x^{\beta-1}\,\mathsf{P}\{Z_n>x\}\,dx\\ &= 2\beta\,\mathsf{E}(Z_n)_+^\beta. \end{aligned}\tag{19}$$

Here and in what follows

$$x_+ = \max(x,0), \qquad x_- = \max(-x,0)$$

for all $x \in \mathbf{R}^1$.

If $F \in D(\Phi_\alpha)$ and $\beta < \alpha$, then [8]

$$\lim_{n\to\infty}\mathsf{E}\left(\frac{Z_n}{\gamma_n}\right)_+^\beta = \int_0^\infty x^\beta\,d\Phi_\alpha(x) < +\infty. \tag{20}$$

Taking into account (19) and (20), we deduce from (18) that

$$\mathsf{E}\left|\frac{\hat Z_n - Z_n}{\gamma_n}\right|^\beta = O\left(n^{-\beta/2}\right). \tag{21}$$

Under assumptions of Theorem 1, the random variables Z_n satisfy equality (12). Therefore, relation (12) holds for $\hat Z_n$, as well. □

The asymptotic behavior of the normalizing sequence γ_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, γ_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

$$U_1(U_2(x)) \sim U_2(U_1(x)) \sim x, \qquad x\to\infty,$$

where $U_2(x) = x^{1/\alpha}L_2(x)$. Then

$$n^{1/\alpha}L_2(n) = U_2(n) = U_2(U_1(\gamma_n)) \sim \gamma_n \tag{22}$$

as $n\to\infty$.

Example 1. (see [3, Example 2.3.4] or [4, Example 1.7.8]) Let the random variables (ε_j) possess the Cauchy distribution, that is, $F(x) = \frac12 + \frac1\pi\arctan x$. It can be easily checked that

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)} = x^{-1}.$$

Thus the distribution Φ_1 defined by (8) is the limit distribution for $\hat Z_n/\gamma_n$. Further,

$$1-F(\gamma_n) = \frac1n, \qquad \gamma_n = \tan\left(\frac\pi2 - \frac\pi n\right) = nL_2(n),$$
$$L_2(x) = \frac{1}{x\tan\left(\frac\pi x\right)} \sim \frac1\pi.$$

This means that one can choose $\gamma_n = n/\pi$.

Put $\hat Z_n^* = \max_{1\le j\le n}|\hat\varepsilon_j|$.

Theorem 2. *Under the assumptions of Theorem* 1,

$$\lim_{n\to\infty} \mathsf{P}\left\{\frac{\hat{Z}_n^*}{\gamma_n} < x\right\} = \Phi_\alpha^2(x), \qquad x \in \mathbf{R}^1. \tag{23}$$

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

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

Here we used Theorem 1.8.2 of [4]:

$$\mathsf{P}\{Z_n^* < u_n\} = \mathsf{P}\{Z_n < u_n, W_n > -u_n\} \to \Phi_\alpha^2(x)$$

as $n \to \infty$.

To complete the proof of (23) we write

$$|Z_n^* - \hat{Z}_n^*| \le \max_{1\le j\le n} |\varepsilon_j - \hat{\varepsilon}_j|$$

and repeat the same reasoning as in the proof of Theorem 1. □

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

$$(iii) \qquad \lim_{n\to\infty} d_{in}^{-1} \max_{1\le j\le n} |x_{ji}| = 0, \qquad i = 1, \dots, q \tag{24}$$

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

$$\mathsf{E}\left(\frac{|\hat{Z}_n - Z_n|}{\gamma_n}\right)^\beta, \qquad \mathsf{E}\left(\frac{|\hat{Z}_n^* - Z_n^*|}{\gamma_n}\right)^\beta.$$

Consider the simple linear regression model

$$y_j = \theta_1 + \theta_2 x_j + \varepsilon_j, \qquad j = 1, \dots, n, \tag{25}$$

where ε_j are independent identically distributed random variables.

Since (25) is a particular case of the model (1), 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

$$d_{1n}^2 = n, \qquad d_{2n}^2 = \sum_{j=1}^n x_j^2, \qquad v_n = n^{-1/2} d_{2n}^{-1} \sum_{j=1}^n x_j,$$

$$J_n = \begin{pmatrix} 1 & v_n \\ v_n & 1 \end{pmatrix},$$

and thus (ii) holds if $|v_n| \le 1 - \delta$ for some $\delta > 0$. In turn, condition (iii) for the simple linear regression model is rewritten as follows

$$\lim_{n\to\infty} d_{2n}^{-1} \max_{1\le j\le n} |x_j| = 0.$$

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 (θ_1, θ_2) for this case is given by

$$\hat{\theta}_{2n} = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^n (x_j - \bar{x})^2}, \qquad \hat{\theta}_{1n} = \bar{y} - \hat{\theta}_{2n}\bar{x},$$

$$\bar{x} = \frac{1}{n}\sum_{j=1}^n x_j, \qquad \bar{y} = \frac{1}{n}\sum_{j=1}^n y_j.$$

Note that $\hat{y}_j = \hat{\theta}_{1n} + \hat{\theta}_{2n} x_j$. The other notation is not changed.

Theorem 3. *Let random variables ε_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*

$$\frac{\max_{1\le j\le n} |x_j|}{\sqrt{\sum_{j=1}^n (x_j - \bar{x})^2}} \to 0, \qquad n \to \infty. \tag{26}$$

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

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

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

$$\hat{\theta}_{2n} = \frac{\sum_{j=1}^n (x_j - \bar{x})(\theta_2 (x_j - \bar{x}) + \varepsilon_j - \bar{\varepsilon})}{\sum_{j=1}^n (x_j - \bar{x})^2} = \theta_2 + \frac{\sum_{j=1}^n (x_j - \bar{x}) \varepsilon_j}{\sum_{j=1}^n (x_j - \bar{x})^2}. \tag{27}$$

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

$$\mathsf{E}\, |\theta_2 - \hat{\theta}_{2n}|^\beta \le \mathsf{E} \left(\frac{\sum_{j=1}^n (x_j - \bar{x})^2 \varepsilon_j^2}{(\sum_{j=1}^n (x_j - \bar{x})^2)^2} \right)^{\beta/2} \le \frac{\mathsf{E} \max_{1\le j\le n} |\varepsilon_j|^\beta}{(\sum_{j=1}^n (x_j - \bar{x})^2)^{\beta/2}} \tag{28}$$

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

$$\mathsf{E} \left| \frac{\theta_2 - \hat{\theta}_{2n}}{\gamma_n} \right|^\beta \max_{1\le j\le n} |x_j|^\beta \le C \left(\frac{\max_{1\le j\le n} |(x_j|}{\sqrt{\sum_{j=1}^n (x_j - \bar{x})^2}} \right)^\beta \to 0, \qquad n \to \infty. \tag{29}$$

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

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

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

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

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

$$\mathsf{E} \left(\frac{|\bar{\varepsilon}|}{\gamma_n} \right)^\beta \le \frac{n^{1-\beta}\, \mathsf{E}\, |\varepsilon_1|^\beta}{\gamma_n^\beta} = \frac{\mathsf{E}\, |\varepsilon_1|^\beta}{L_2^\beta(n)} \frac{1}{n^{\beta/\alpha + \beta - 1}}.$$

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

$$\mathsf{E}\, (|\bar{\varepsilon}|/\gamma_n)^\beta \to 0$$

as $n \to \infty$.

If $1 \le \beta < \alpha \le 2$, 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 (23) is straightforward. □

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

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

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

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

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

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

$$\mathsf{E}(\varepsilon_j)_-^{\beta} < \infty \tag{30}$$

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 ε_j is symmetric is needed only to make sure that Lemma 1 applies. Under the assumptions $\mathsf{E}\,\varepsilon_j = 0$ and $\beta \ge 1$, 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 $\beta \ge 1$.

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

$$\mathsf{E}\,|\hat{Z}_n - Z_n| \le (q!)\lambda_0^{-q}\left(\sum_{i=1}^{q} k_i\right)\left(\mathsf{E}\max_{1\le j\le n}|\varepsilon_j|\right) n^{-1/2}. \tag{31}$$

Considering (22) we get

$$\frac{\mathsf{E}\max_{1\le j\le n}|\varepsilon_j|}{\gamma_n} \le \frac{n\,\mathsf{E}\,|\varepsilon_1|}{\gamma_n} = \frac{\mathsf{E}\,|\varepsilon_1|}{L_2(n)n^{1/\alpha-1}}.$$

The latter result together with (31) implies that

$$\frac{\mathsf{E}\,|\hat{Z}_n - Z_n|}{\gamma_n} \le C\frac{\mathsf{E}\,|\varepsilon_1|}{L_2(n)n^{1/\alpha+1/2-1}} \to 0$$

as $n \to \infty$. Now we establish equality (12) in the same way as in the proof of Theorem 1.

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

$$\begin{aligned}\frac{\mathsf{E}\,|\hat{Z}_n - Z_n|^{\beta}}{\gamma_n^{\beta}} &\le C\frac{\mathsf{E}\max_{1\le j\le n}|\varepsilon_j|^{\beta}}{\gamma_n^{\beta}} n^{-\beta/2} \\ &\le C\frac{\mathsf{E}\max_{1\le j\le n}(\varepsilon_j)_-^{\beta} + \mathsf{E}\max_{1\le j\le n}(\varepsilon_j)_+^{\beta}}{\gamma_n^{\beta}}\, n^{-\beta/2}.\end{aligned} \tag{32}$$

According to equality (20),

$$\frac{\mathsf{E}\max_{1\le j\le n}(\varepsilon_j)_+^\beta}{\gamma_n^\beta}\, n^{-\beta/2} \to 0 \tag{33}$$

as $n\to\infty$. Further,

$$\frac{\mathsf{E}\max_{1\le j\le n}(\varepsilon_j)_-^\beta}{\gamma_n^\beta}\, n^{-\beta/2} \le \frac{n\,\mathsf{E}(\varepsilon_1)_-^\beta}{n^\beta L_2^\beta(n)} \to 0. \tag{34}$$

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

$$H_1^+ = \{F\in D(\Phi_\alpha), \alpha>1\}, \qquad H_1^- = \{F\in D(\Phi_\alpha), \alpha<1\}.$$

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 $T_n = \pi\hat{Z}_n^*/n$. Given a level of significance δ, the rejection regions are given by

$$K^+ = \{x\colon 0<x<U_{\delta/2}\}, \qquad K^- = \{x\colon U_{1-\delta/2}<x\}, \qquad K = K^+\cup K^-,$$

where U_p are the p-quantiles for the distribution function $\Phi_1^2(x)$. For example, if $\delta=0.05$, then

$$K^+ = \left(0, \frac{2}{\ln(2/\delta)}\right) \approx (0, 0.54), \qquad K^- = \left(\frac{2}{\ln(1/(1-\delta/2)}, \infty\right) \approx (79.0, \infty).$$

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 $T_n\in\mathbf{R}^1\setminus K$, then the hypothesis H_0 is accepted; otherwise

$$\text{If } T_n\in K^+, \quad \text{then } H_1^+ \text{ is accepted.}$$
$$\text{If } T_n\in K^-, \quad \text{then } H_1^- \text{ is accepted.}$$

Bibliography

1. O. V. Ivanov and I. K. Matsak, *Limit theorems for extremal residuals in linear and nonlinear regression models*, Teor. Imovir. Mat. Stat. **86** (2012), 69–80; English transl. in Theory Probab. Math. Statist. **86** (2013), 79–91. MR2986451
2. B. V. Gnedenko, *Sur la distribution limit du terme maximum d'une série aléatoire*, Ann. Math. **44** (1943), 423–453. MR0008655 (5:41b)
3. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York–Chichester–Brisbane, 1978. MR489334 (80b:60040)
4. M. R. Leadbetter, Georg Lindgren, and Holger Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983. MR691492 (84h:60050)
5. V. V. Petrov, *Sums of Independent Random Variables*, "Nauka", Moscow, 1972; English transl., Springer-Verlag, New York–Heidelberg, 1975. MR0322927 (48:1288)
6. G. Peshkir and A. N. Shiryaev, *Khintchine inequalities and a martingale extension of the sphere of their action*, Uspekhi Matematicheskikh Nauk **50** (1995), no. 3, 3–62; English transl. in Russian Mathematical Surveys **50** (1995), no. 5, 849–904. MR1365047 (96k:60038)
7. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II, Second edition John Wiley & Sons Inc., New York–London–Sydney, 1971. MR0270403 (42:5292)

8. J. Pickands, *Moment convergence of sample extremes*, Ann. Math. Statist. **39** (1968), no. 3, 881–889. MR0224231 (36:7275)
9. Eugene Seneta, *Regularly Varying Functions*, Lecture Notes in Mathematics, Vol. 508, Springer-Verlag, Berlin, 1976. MR0453936 (56:12189)

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