S(e)MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports

http://semr.math.nsc.ru

*Том 18, №2, стр. 1058–1074 (2021)* УДК 519.233
DOI 10.33048/semi.2021.18.081 MSC 62F03

# REMARKS ON INVARIANCE PRINCIPLE FOR ONE-PARAMETRIC RECURSIVE RESIDUALS

A.I. SAKHANENKO, A.P. KOVALEVSKII, A.D. SHELEPOVA

Abstract. We investigate a linear regression model with one unknown parameter. The idea of recursive regression residuals is to estimate the regression parameter at each moment on the base of previous variables. Therefore the distribution of recursive residuals does not depend on the parameter. We investigate conditions for the weak convergence of the process of sums of recursive residuals, properly normalized, to a standard Wiener process. We obtain new conditions, which are better than ones in Sen (1982). The recursive residuals were introduced by Brown, Durbin and Evans (1975). Such residuals are the useful instrument for testing hypotheses about linear regression. Our results give opportunity to use correctly recursive residuals for a wide class of regression sequences, including sinusoidal and i.i.d. bounded.

**Keywords:** linear regression, recursive residuals, weak convergence, Wiener process.

## 1. Introduction

1.1. **Preliminary notions.** Consider the regression model

$$Y_i = x_i\beta_i + \varepsilon_i, \qquad i = 1, 2, \dots, \tag{1}$$

where, at time $i$, $Y_i$ is the dependent variate, $x_i$ is the regressor, $\beta_i$ is the (unknown) regression coefficient and $\varepsilon_i$ is the (unobservable) random error component.

---

Sakhanenko, A.I., Kovalevskii, A.P., Shelepova, A.D., Remarks on invariance principle for one-parametric recursive residuals.



The work is supported by Mathematical Center in Akademgorodok under agreement No. 075-15-2019-1675 with the Ministry of Science and Higher Education of the Russian Federation.

*Received August, 29, 2021, published October, 20, 2021.*

Introduce the following important notation:

$$v_n^2 := x_1^2 + \cdots + x_n^2, \qquad n = 1, 2, \ldots.$$

Denote by $H_0(r)$ the (null) hypothesis that

$$\beta_i = \beta \text{ (unknown)} \quad \text{for all } i \geq 1 \qquad \text{and}$$

$$\varepsilon_1, \varepsilon_2, \ldots \quad \text{are i.i.d. with} \quad \mathbb{E}\varepsilon_1 = 0 \quad \text{and} \quad 0 < \sigma^2 = \mathbb{E}\varepsilon_i^2 < \infty,$$

along with the assumption that

$$v_r^2 = x_1^2 + \cdots + x_r^2 > 0 \quad \text{for some integer} \quad r \geq 1. \tag{2}$$

Thus, we suppose in hypothesis $H_0(r)$ that the random variable $\varepsilon_1$ has an unknown (but fixed) distribution. Note, that under $H_0(r)$, representation (1) may be rewritten in the following way:

$$Y_i = x_i\beta + \varepsilon_i, \qquad i = 1, 2, \ldots. \tag{3}$$

Assuming $H_0(r)$ to be true, let $\widehat{\beta}_k$ be the least-squares estimate of $\beta$, based on the first $k$ observations. It is known that

$$\widehat{\beta}_k := \sum_{j=1}^{k} x_j Y_j \Big/ \sum_{j=1}^{k} x_j^2 = \beta + \sum_{j=1}^{k} x_j \varepsilon_j \Big/ v_k^2, \qquad \forall\, k \geq r. \tag{4}$$

The second equality is true by (3).

Underline that, in applications, only the first $n$ variables $Y_1, \ldots, Y_n$ are observed (for sufficiently large $n$), and only the first $n$ regressors $x_1, \ldots, x_n$ are known. Of course, in this case we have right to use estimates $\widehat{\beta}_k$ only for $r \leq k \leq n$.

1.2. **Ideas of Brown, Durbin and Evans (1975).** To describe these ideas we first introduce the following recursive regression residuals

$$\widehat{\varepsilon}_i := Y_i - x_i\widehat{\beta}_{i-1} = \varepsilon_i - x_i \sum_{j=1}^{i-1} x_j \varepsilon_j \Big/ v_{i-1}^2,, \quad \forall\, i > r, \tag{5}$$

where the second equality in (5) follows from (3) and (4). Underline, that in the definition (5) we use the estimate $\widehat{\beta}_{i-1}$ instead of the sharper estimate $\widehat{\beta}_n$ even in the case when we have $n$ observations with large $n > i - 1$. Now from (5) it is clear that

$$\mathbb{E}\,\widehat{\varepsilon}_i = 0 \quad \text{and} \quad 1 \leq \alpha_i^2 := \mathbb{E}\,\widehat{\varepsilon}_i^2/\sigma^2 = 1 + x_i^2/v_{i-1}^2 < \infty, \quad \forall\, i > r. \tag{6}$$

Second, we introduce the following weighted recursive regression residuals

$$w_i := \frac{\widehat{\varepsilon}_{r+i}}{\alpha_{r+i}} = \frac{Y_{r+i} - x_{r+i}\widehat{\beta}_{r+i-1}}{\sqrt{1 + x_i^2/v_{i-1}^2}}, \quad i = 1, 2, \ldots. \tag{7}$$

It is easy to see from (5) – (7) that

$$\mathbb{E}w_i = 0 \quad \text{and} \quad \mathbb{E}w_i^2 = \sigma^2, \quad i = 1, 2, \ldots. \tag{8}$$

**Property A.** *Suppose that the hypothesis $H_0(r)$ is fulfilled. Then equalities* (8) *are true and, moreover, random variables $w_1, w_2, \ldots$ are uncorrelated, i.e.*

$$\mathbb{E}w_k w_j = 0, \quad \textit{for all} \quad k > j \geq 1. \tag{9}$$

*If, in addition, random variable $\varepsilon_1$ has a normal distribution, then random variables $w_1, w_2, \ldots$ are independent and identically distributed with $\varepsilon_1$.*

This remarkable property follows from Lemma 1 in Brown, Durbin and Evans (1975).

1.3. **Invariance principle for recursive residuals.** For any sequence of (one-dimensional) random variables, say $\xi_1, \xi_2, \dots$, we will denote by $S_n(t; \xi_\bullet)$ random processes with the following properties:

$$\forall t \in [k, k+1) \quad S_1(t; \xi_\bullet) := \sum_{i=1}^{k} \xi_i + h(t-k)\xi_{k+1}, \quad k = 0, 1, 2, \dots, \tag{10}$$

where $h$ is some measurable bounded function on $[0, 1]$; and we put

$$S_n(t; \xi_\bullet) = S_1(nt; \xi_\bullet)/\sqrt{n} \quad \text{for all} \quad t \in [0, 1], \quad n = 1, 2, \dots. \tag{11}$$

Note that for $k = 0$ we use in (10) the standard agreement that $\sum_\emptyset = 0$.

When $h(t) = h_1(t) := t$, the processes $S_n(\cdot; \xi_\bullet)$ are often called random broken lines. Remind, that in this case it was proved by Donsker (1951) that

$$\frac{1}{\sigma} S_n(\cdot; \varepsilon_\bullet) \Longrightarrow W(\cdot) \quad \text{in} \quad C[0, 1], \tag{12}$$

where $W(\cdot)$ is a standard Wiener process. (Below, in subsection 3.2., we remind, in details, the definition of convergence (12).)

The natural question arise whether we have the following convergence

$$W_n(\cdot) := \frac{1}{\sigma} S_n(\cdot; w_\bullet) \Longrightarrow W(\cdot) \quad \text{in} \quad C[0, 1], \tag{13}$$

for the weighted regression recursive residuals introduced in (7), when $h = h_1$. Underline, that in (12), (13) and everywhere below all limits are taken as $n \to \infty$ in all cases when the inverse is not specified explicitly.

**Example 1.** (See [3]). *Suppose that the random variable $\varepsilon_1$ has a normal distribution. Then convergence* (13) *takes place under the hypothesis $H_0(r)$.*

In [3] this fact was rightly used as obvious corollary of Property A. Remind that the approach from [3] is alternative to one developed by MacNeill (1978) and Bishoff (1998). And the Example 1 shows that the approach of Brown, Durbin and Evans (1975) from [3] has a number of advantages.

However, when the $\{\varepsilon_i\}$ are not normally distributed, the $\{w_r\}$ are not necessarily independent nor identically distributed, and hence, invariance principles for these weighted recursive residuals is not obvious. Moreover, the problem to obtain the invariance principle (13) appeared to be sufficiently difficult.

Below we present the three partial cases of the invariance principle (13), which we have found on page 311 of the more general paper of Sen (1982).

**Example 2.** *Assume that $x_n = x_1 \neq 0$ for all $n \geq 1$. Then convergence* (13) *holds under hypothesis $H_0(r)$.*

In reality, in Sen (1982), p. 311, Example 2 was considered only when $x_1 = 1$. But we always may investigate observations $Y_i/x_1$ instead of $Y_i$.

**Example 3.** *Suppose that*

$$\max_{1 \leq k \leq n} x_k^2 / v_{n-1}^2 = O(n^{-\lambda}), \quad \textit{for some} \quad \lambda > 1/2. \tag{14}$$

*If, in addition, $\mathbb{E}\varepsilon_1^4 < \infty$, then convergence* (13) *holds under hypothesis $H_0(r)$.*

**Example 4.** *Assume that condition* (14) *is true for* $\lambda = 1$, *i.e.*

$$\max_{1\le k\le n} x_k^2/v_{n-1}^2 = O(1/n) \quad as \quad r < n \to \infty. \tag{15}$$

*Suppose, in addition, that for every* $1 \le i \le n-2$

$$(x_i - x_{i+1})x_n/v_{n-1}^2 = O(1/n^2) \quad as \quad r < n \to \infty. \tag{16}$$

*Then convergence* (13) *takes place under hypothesis* $H_0(r)$.

Underline, that in Sen (1982), p. 311, Examples 3 and 4 was considered in a more general case, when the unknown parameter $\beta$ is multidimensional. Here we agree with Bishoff (2016) that the paper of Sen (1982) until now remains to be the only known investigation on the subject.

1.4. **On conditions** (15) **and** (16)**.** The following two simple properties may be interesting.

**Property 1.** *Under assumption* (15)

$$\forall\, n \ge r \quad v_n^2 \le v_r^2 (n/r)^C \quad for \quad 0 \le C := \sup_{n>r} n x_n^2/v_{n-1}^2 < \infty. \tag{17}$$

**Property 2.** *If condition* (15) *holds together with* (16), *then either*

$$3 \le K := \limsup_{n\to\infty} \frac{\log v_n^2}{\log n} \le C < \infty, \tag{18}$$

*or* $x_i = x_1$ *for all* $i \ge 1$.

Thus, condition (16) is restrictive for one-dimensional parameter $\beta$.

## 2. Main results and discussions

2.1. **Main statements.** Below in the paper we suppose that the hypothesis $H_0(r)$ is fulfilled (together with condition (2)).

**Theorem 1.** *Assume that*

$$\mu_n := |x_n|/v_n \to \mu \ge 0. \tag{19}$$

*Then the finite-dimensional distributions of the processes* $\{W_n(t) : t \in [0,1]\}$ *converge to those of the standard Wiener process* $W(\cdot)$.

**Corollary 1.** *The assertion of Theorem 1 takes place in the case when*

$$v_\infty^2 := \sum_{i=1}^{\infty} x_i^2 < \infty. \tag{20}$$

Indeed, in this case $x_n \to 0$ and condition (19) is satisfied for $\mu = 0$.

**Corollary 2.** *The assertion of Theorem 1 takes place in the case when*

$$\sup_{n\ge1} |x_n| < \infty.$$

Indeed, in this case the condition (19) is again satisfied for $\mu = 0$ when

$$v_n^2 \equiv x_1^2 + \cdots + x_n^2 \to \infty \quad \text{as} \quad n \to \infty. \tag{21}$$

On the other hand, if (21) does not hold, then Corollary 2 is a partial case of Corollary 1.

**Theorem 2.** *Suppose that*

(22) $$\mu_n^2 \equiv x_n^2/v_n^2 = O(1/n) \quad \text{as} \quad n \to \infty.$$

*Then the convergence* (13) *takes place.*

**Corollary 3.** *Let the assumption* (15) *from Example 4 is satisfied. Then conditions* (22) *and* (21) *hold, and the convergence* (13) *takes place.*

Thus, Theorem 2 allows us to obtain the assertions from Examples 4 and 3 (when $\lambda \geq 1$) without additional restrictions, but (22). Note, that our Theorems 1 and 2 may be applied, in contrast with Examples 4 and 3, also in the case (20).

2.2. **Examples.**

**Example 5.** *Assume the model of the sinusoidal regressor, that is,*

$$x_i = \sin\left(\frac{\pi k i}{T} + \varphi_0\right), \quad i \geq 1, \quad k, T \in \mathbf{Z}, \quad 0 < k < T, \quad \varphi_0 \in \mathbf{R}.$$

*Then* (2) *holds with* $r \leq 2$, $x_i \not\equiv x_1$, $|x_i| \leq 1$ *and*

$$v_n^2/n = \frac{1}{2n}\sum_{i=1}^{n}\left(1 - \cos\left(\frac{2\pi k i}{T} + \varphi_0\right)\right) \to 1/2 \quad \text{as} \quad n \to \infty$$

*because*

$$\sum_{i=1}^{mT}\cos\frac{2\pi k i}{T} = \sum_{i=1}^{mT}\sin\frac{2\pi k i}{T} = 0 \quad \text{for any} \quad m \in \mathbf{Z}_+.$$

*Thus* (22) *holds whereas* (16) *does not.*

**Example 6.** *Assume that* $x_i$ *are i.i.d. bounded random variables,* $\{x_i\}_{i\geq 1}$ *and* $\{\varepsilon_i\}_{i\geq 1}$ *are independent,* $\mathbb{V}ar\, x_1 > 0$, $\mathbb{P}(x_1 = 0) = 0$. *Then* (2) *holds with* $r = 1$ *a.s.,* (22) *holds a.s. whereas* (16) *does not hold with probability one.*

**Example 7.** *Note that assumptions of Theorem 1 hold for any* $\mu \in [0, 1]$ *threas ones of Theorem 2 do are not true if* $\mu > 0$. *For example, if* $x_i = g^i$ *for each* $i \geq 1$ *and some* $g > 1$ *then* $v_n^2 = g^{2n} - 1$ *and* (19) *takes place with* $\mu = \sqrt{1 - 1/g^2} > 0$; *whereas* (22) *does not hold.*

2.3. **Standard generalization.** For arbitrary function $h$ in (10) we use below notation $\tilde{W}_n(\cdot) := \tilde{S}_n(\cdot; w_\bullet)/\sigma$ instead of the notation $W_n(\cdot)$ which we use when $h(t) = h_1(t) := t$. In subsection 3.1 we prove the next

**Property 3.** *Suppose that* $\sup_{t\in[0,1]} |h(t)| < \infty$ *and that* $\mathbb{E}w_i^2 < \infty$ *for all* $i \geq 1$. *Then*

(23) $$\sup_{t\in[0,1]} |\tilde{W}_n(t) - W_n(t)| \leq \rho_n \sup_{t\in[0,1]} |h(t) - t|,$$

*where*

(24) $$\rho_n := \frac{1}{\sqrt{n}} \max_{1\leq i\leq n} |w_i| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Underline, that if the bounded function $h$ in (10) is not continuous, then $\tilde{W}_n \notin C[0, 1]$ and, as a result, we have no right to use convergence (13) with $\tilde{W}_n$ instead of $W_n$. But, by Property 1, it is possible to find a neighborhood $C_+[0,1]$ of $C[0,1]$ such that

$$\mathbb{P}(\tilde{W}_n \in C_+[0,1]) \to 1 \quad \text{as} \quad n \to \infty. \tag{25}$$

The famous partial case here is when $h(t) = h_0(t) := 0$. In this case processes $\tilde{W}_n(t;\xi_\bullet)$ are called random step-functions.

Consider now a functional $f : C_+[0,1] \to R$ which is a.s. continuous, i.e.

$$\mathbb{P}(f \;\text{ is continuous at random "point" } W) = 1, \tag{26}$$

where, remind, $W$ is a standard Wiener process. Now from (13) and Property 1 we obtain that

$$f(\tilde{W}_n) \Rightarrow f(W) \quad \text{as} \quad n \to \infty. \tag{27}$$

Convergence (27) for all functionals $f : C_+[0,1] \to R$ which satisfy assumptions (25) and (26), is naturally write as the following convergence

$$\tilde{W}_n(\cdot) := \frac{1}{\sigma}\tilde{S}_n(\cdot; w_\bullet) \Longrightarrow W(\cdot) \quad \text{in} \quad C_+[0,1], \tag{28}$$

as $n \to \infty$. It is an obvious generalization of convergence (13). Such generalization is well known, but we do not know a standard notation for it.

2.4. **On applications of Invariance principles.** Denote by

$$s_n^2 := \sum_{i=r}^{n}(Y_i - \widehat{\beta}_n\xi_i)^2/(n-r), \quad n > r,$$

the standard statistical estimator of the unknown variance $\sigma^2$, constructed by a sample of the size $n$. Then we have from (28) that

$$\tilde{W}_n^*(\cdot) := \frac{1}{s_n}\tilde{S}_n(\cdot; w_\bullet) \Longrightarrow W(\cdot) \quad \text{in} \quad C_+[0,1],$$

as $n \to \infty$. It is a form of Invariance principle which is convenient for statistical applications.

Consider now an example. Below in the paper we will suppose that $h(t) = 0$ and, hence, that $\tilde{W}_n(\cdot)$ and $\tilde{W}_n^*(\cdot)$ are step-functions. For any function $x = x(\cdot)$ introduce two functionals

$$f_0(x) := \sup_{t\in[0,1]} |x(t)| \quad \text{and} \quad f_+(x) := \sup_{t\in[0,1]} x(t).$$

Consider two statistics

$$D_n^+ = \max_{r\le k\le n} S_1(k)/\{s_n(n-r)^{1/2}\} \quad \text{and} \quad D_n = \max_{r\le k\le n} |S_1(k)|/\{s_n(n-r)^{1/2}\}.$$

It is easy to see from (27) that, as $n \to \infty$,

$$D_n^+ = f_+(\tilde{W}_n^*) \Rightarrow f(W) = \max_{t\in[0,1]} W(t); \tag{29}$$

$$D_n^+ = f_0(\tilde{W}_n^*) \Rightarrow f_0(W) = \max_{t\in[0,1]} |W(t)| \tag{30}$$

Thus, formulas (29) and (30) allow us to find approximations for the distribution of statistics $D_n^+$ and $D_n$. Underline, that the known CUSUM test is based on these statistics. This test is used in analysis for structural change in linear regression models (see, e.g., Zeileis et al., 2002).

The rest of the paper is devoted to proofs of assertions stated above in Section 2. Remind, that the hypothesis $H_0(r)$ is supposed to hold below.

## 3. Proofs of Properties 1 – 3 and elementary lemmas

3.1. **Proof of Property 1.** For each $k > r$

$$\log v_k^2 - \log v_{k-1}^2 = \log\left(1 + \frac{x_k^2}{v_{k-1}^2}\right) \le \frac{x_k^2}{v_{k-1}^2} \le \frac{C}{k} < \int\limits_{k-1}^{k} \frac{C dx}{x}.$$

Summing up, we obtain for $n > r$ that

$$\log \frac{v_n^2}{v_r^2} = \sum_{k=r+1}^{n} (\log v_k^2 - \log v_{k-1}^2) < \int\limits_{r}^{n} \frac{C dx}{x} = C \log \frac{n}{r}.$$

And (17) follows.

3.2. **Proof of Property 2.** If condition (15) is fulfilled, then we have from Property 2, that $K \le C$, where the value $K$ is defined in (18). But if $K < \infty$, then

$$\forall \varepsilon > 0 \quad v_n^2 = o(n^{K+\varepsilon}) \quad \text{as} \quad n \to \infty. \tag{31}$$

Suppose now that assumption (16) is true for some $i \ge 1$ with $x_i \ne x_{i+1}$. Then

$$x_n / v_{n-1}^2 = O(1/n^2) \quad \text{and} \quad x_n^2 = O(v_n^4/n^4) \quad \text{as} \quad n \to \infty.$$

This fact and (31) imply that

$$\forall \varepsilon > 0 \quad x_n^2 = O(v_n^4/n^4) = o(n^{2K+2\varepsilon-4}).$$

Hence, for each $\varepsilon > 0$ there exists a number $N(\varepsilon) < \infty$ such that

$$x_n^2 \le n^{2K+2\varepsilon-4} \quad \text{for all} \quad n \ge N(\varepsilon).$$

In particular, for all $n > N(\varepsilon)$

$$v_n^2 = \sum_{k=1}^{N(\varepsilon)-1} x_k^2 + \sum_{k=N(\varepsilon)}^{n} x_k^2 \le O(1) + \sum_{k=N(\varepsilon)}^{n} k^{2K+2\varepsilon-4}. \tag{32}$$

Suppose first that $K < 3/2$. Then we have from (32) with $\varepsilon = (3 - 2K)/3 > 0$ that

$$v_n^2 = O(1) + \sum_{k=N(\varepsilon)}^{n} k^{-1-\varepsilon} = O(1).$$

But this fact contradicts to condition $v_n \to \infty$, which follows from (15).

Assume next that $3/2 \le K < 3$. Then we put $\varepsilon = (3 - K)/3 > 0$ in (32) and obtain that

$$v_n^2 = O(1) + \sum_{k=N(\varepsilon)}^{n} k^{K-1-\varepsilon} = O(n^{K-\varepsilon}), \quad \varepsilon < K.$$

Hence

$$\limsup_{n\to\infty} \frac{\log v_n^2}{\log n} \le K - \varepsilon < K.$$

But it contradicts to the definition (18) of the value $K$.

Thus, only the case $K \ge 3$ is possible.

3.3. **Proof of Property 3.** Inequality (23) follows from (10) and (11) with no assumption.

To prove (24), first of all note that

$$\delta_i(N) := \mathbb{E}[w_i^2 : |w_i| > N] \to 0 \quad \text{as} \quad N \to \infty,$$

if only $\mathbb{E}w_i^2 < \infty$. Hence, when $N_n \to \infty$ as $n \to \infty$, we have

$$\overline{\delta}_n(N_n) := \frac{1}{n}\sum_{i=1}^{n} \delta_i(N_n) \to 0 \quad \text{as} \quad n \to \infty.$$

Now, using the idea of Chebyshev, we obtain for each $\varepsilon > 0$ that

$$\begin{aligned}\mathbb{P}(\rho_n > \varepsilon) &\le \sum_{i=1}^{n} \mathbb{P}(|w_i| > \varepsilon\sqrt{n}) \le \frac{1}{(\varepsilon\sqrt{n})^2}\sum_{i=1}^{n} \delta_i(\varepsilon\sqrt{n}) \\ &= \frac{1}{\varepsilon^2}\overline{\delta}_n(\varepsilon\sqrt{n}) \to 0 \quad \text{as} \quad n \to \infty.\end{aligned}$$

Thus, (24) is proved.

3.4. **Two elementary lemmas.** Define

$$x_n^* := \max_{1\le j\le n} |x_j|, \quad \overline{\mu}_n := \sup_{k\ge n} \mu_k = \sup_{k\ge n} \frac{|x_k|}{v_k}, \quad \mu_n^* := \sup_{k\ge n} \frac{x_k^*}{v_k}. \tag{33}$$

Below we frequently use the following evident inequalities:

$$\forall\ j \ge 0 \quad |x_{r+j}| \le x_{r+j}^* \le v_{r+j}, \quad \mu_{r+j} \le \overline{\mu}_{r+j} \le \mu_{r+j}^* \le 1 \le \alpha_{r+j}. \tag{34}$$

**Lemma 1.** *Assume that* (21) *and* (19) *takes place with* $\mu = 0$*. Then*

$$0 \le \mu_n = \frac{|x_n|}{v_n} \le \overline{\mu}_n := \sup_{k\ge n} \mu_k \le \mu_n^* := \sup_{k\ge n} \frac{x_k^*}{v_k} \to 0. \tag{35}$$

*Proof.* For all $n > m \ge r$

$$\begin{aligned}0 \le \mu_n \le \overline{\mu}_n \le \mu_n^* &:= \sup_{k\ge n} \frac{x_k^*}{v_k} = \sup_{k\ge n} \max_{1\le j\le k} \frac{|x_j|}{v_k} \\ &\le \max_{1\le j\le m} \frac{|x_j|}{v_n} + \sup_{j>m} \frac{|x_j|}{v_j} \le \frac{x_m^*}{v_n} + \overline{\mu}_m.\end{aligned}$$

Taking now limit as $n \to \infty$, we get

$$0 \le \limsup_{n\to\infty} \mu_n^* \le \mu_m \quad \text{for all} \quad m \ge r. \tag{36}$$

But the left-hand side in (36) does not depend on $m$. So, taking $m \to \infty$ in (36), we obtain (35) because, by the definition of limits, as $m \to \infty$, the sequence $\overline{\mu}_m$ has the same limit $\mu = 0$ as the sequence $\mu_m$ has. □

For all $j > r$, $m > k \ge 0$ introduce the following notations

$$a_j := \frac{x_j}{v_j v_{j-1}}, \quad \overline{a}_{k,m} := \sum_{i=k+1}^{m} |a_{r+i}|, \quad a_{k,m}^* := \max_{k\le i\le m} v_{r+i}\overline{a}_{i,m}. \tag{37}$$

**Lemma 2.** *Under assumption* (19) *for all* $n \geq l > j \geq m$

$$a^*_{j,l}/\sqrt{n} \leq \sqrt{1 - v^2_{r+m}/v^2_{r+n}}. \tag{38}$$

*In addition, when* $-r < j \leq k$ *and* $0 < k < l \leq m$ *we have*

$$|x_{j+r}|\overline{a}_{k,l} \leq v_{j+r}(l-k)/v_{k+r} \leq m, \tag{39}$$

*and if* $n \geq l > j \geq m$*, then*

$$|x_{j+r}|\overline{a}_{j,l} \leq x^*_j \overline{a}_{j,n} \leq \mu^*_j a^*_{j,n} \leq \mu^*_m a^*_{m,n}. \tag{40}$$

*Proof.* By Schwartz inequality for all $n \geq l > j \geq 0$

$$\overline{a}^2_{j,l} = \left( \sum_{i=j+1}^{l} |a_{r+i}| \right)^2 \leq (l-j) \sum_{i=j+1}^{l} a^2_{r+i} \leq n \sum_{i=j+1}^{n} \frac{x^2_{r+i}}{v^2_{r+i} v^2_{r+i-1}}$$
$$= n \sum_{i=j+1}^{n} \left( \frac{1}{v^2_{r+i-1}} - \frac{1}{v^2_{r+i}} \right) = n \left( \frac{1}{v^2_{r+j}} - \frac{1}{v^2_{r+n}} \right) = \frac{n}{v^2_{r+j}} \left( 1 - \frac{v^2_{r+j}}{v^2_{r+n}} \right).$$

Hence, when $n \geq l > j \geq m$

$$v^2_{r+j} \overline{a}^2_{j,l} \leq 1 - \frac{v^2_{r+j}}{v^2_{r+n}} \leq 1 - \frac{v^2_{r+m}}{v^2_{r+n}}.$$

Thus, (38) is proved.

Inequalities (39) and (40) follows immediately from (34) and definitions (33) and (37). □

## 4. Proof of Theorem 1

4.1. **Representations for recursive residuals.** For all $i > r$ and $k, l \geq 1$ introduce the following notations

$$a(k) := \sum_{i=1}^{k} a_{r+i}, \quad a_k(l) := a(l) - a(k), \quad \text{where} \quad a_i = \frac{x_i}{v_i v_{i-1}} = \frac{x_i}{\alpha_i v^2_{i-1}}. \tag{41}$$

In this case we have from (5) and (7) that

$$w_i = \frac{\varepsilon_{r+i}}{\alpha_{r+i}} - \frac{x_{r+i}}{\alpha_{r+i} v^2_{r+i-1}} \sum_{j=1}^{r+i-1} x_j \varepsilon_j = \frac{\varepsilon_{r+i}}{\alpha_{r+i}} - \sum_{j=1}^{r+i-1} a_{r+i} x_j \varepsilon_j. \tag{42}$$

Hence, for all $k \geq 1$

$$\tilde{S}_1(k) = S_1(k) = \sum_{i=1}^{k} w_i = \sum_{i=1}^{k} \frac{\varepsilon_{r+i}}{\alpha_{r+i}} - \sum_{i=1}^{k} \sum_{j=1}^{r+i-1} a_{r+i} x_j \varepsilon_j \tag{43}$$
$$= \sum_{j=r+1}^{r+k} \frac{\varepsilon_j}{\alpha_j} + \sum_{j=r+1}^{r+k-1} a(j-r) x_j \varepsilon_j - a(k) \sum_{j=1}^{r+k-1} x_j \varepsilon_j.$$

It is convenient also to introduce the following i. i. d. random variables

$$\xi_i := \varepsilon_{r+i}/\sigma, \quad i = 1-r, 2-r, 3-r, \ldots, \quad \text{with} \quad \mathbb{E}\xi_i = 0 \quad \text{and} \quad \mathbb{E}\xi^2_i = 1. \tag{44}$$

4.2. **Key Lemmas.** We use below notations, introduced before Lemmas 1 and 2.

**Lemma 3.** *For all $n \ge k > 0$ random variables $W_n(k/n)$ may be represented in the form*

$$W_n\left(\frac{k}{n}\right) = \frac{S_1(k)}{\sigma\sqrt{n}} = \sum_{j=1-r}^{n} \frac{A_j(k,n)}{\sqrt{n}}\xi_j \tag{45}$$

*for some numbers $A_j(k,n)$. Moreover, for all $n > m > 0$*

$$A(n) := \max\{|A_j(k,n)| : -r < j \le k \le n\} \le 1 + m + \mu_m^* a_{m,n}^*. \tag{46}$$

*Proof.* Representation (45) follows from (43) and (44) with

$$A_j(k,n) = \begin{cases} \frac{1}{\alpha_{r+k}}, & \text{when } j = k; \\ \frac{1}{\alpha_{r+j}} + a_k(j)x_{j+r}, & \text{if } 1 \le j < k; \\ -a(k)x_{j+r}, & \text{when } -r < j \le 0. \end{cases} \tag{47}$$

Comparing definitions (37) and (41), we obtain when $j \le 0 < m \le n$ and $j \le k \le n$ that

$$|a(k)x_{j+r}| \le |x_{j+r}|\overline{a}_{0,k} \le |x_{j+r}|(\overline{a}_{0,m} + \overline{a}_{m,n}) \le m + \mu_m^* a_{m,n}^*. \tag{48}$$

Here we also used (39) and (40).

Similarly, for all $n \ge k > j \ge m$ we have from (40) that

$$|a_k(j)x_{j+r}| \le |x_{j+r}|\overline{a}_{j,k} \le \mu_m^* a_{m,n}^*. \tag{49}$$

Now consider the case, when $j < k \le m$. We have from (39) that

$$|a_k(j)x_{j+r}| \le |x_{j+r}|\overline{a}_{j,k} \le m. \tag{50}$$

At last, in the most difficult case, when $j \le m < k$, we obtain from (39) and (40) that

$$|a_k(j)x_{j+r}| \le |x_{j+r}|\overline{a}_{j,k} = |x_{j+r}|\overline{a}_{j,m} + |x_{j+r}|\overline{a}_{j,k} \le m + \mu_m^* a_{m,n}^*. \tag{51}$$

Since $\alpha_{r+j} \ge 1$, substituting (48) – (51) into (47), we obtain (46). □

**Lemma 4.** *Under assumption* (19)

$$A(n)/\sqrt{n} \to 0 \quad \textit{as} \quad n \to \infty. \tag{52}$$

*Proof.* First, consider the case, when $v_\infty < \infty$. Since $\mu_m^* \le 1$ by (34), we have from (46) and (38) that

$$A(n)/\sqrt{n} \le (1+m)/\sqrt{n} + \sqrt{1 - v_{r+m}^2/v_{r+n}^2}$$

for all $n > m > 0$. Taking limit as $n \to \infty$, we get

$$\limsup_{n\to\infty} A(n)/\sqrt{n} \le \sqrt{1 - v_{r+m}^2/v_\infty^2} \quad \text{for all} \quad m \ge 1. \tag{53}$$

But the left-hand side in (53) does not depend on $m$. So, taking $m \to \infty$ in (53), we obtain (52) because $v_{r+m} \to v_\infty$ as $m \to \infty$.

Second, consider the case, when $v_n \to \infty$ and (19) holds with $\mu = 0$. Then we have from (46) and (38) that

$$A(n)/\sqrt{n} \le (1+m)/\sqrt{n} + \mu_m^* \quad \forall\ n > m > 0.$$

Taking limit, we get

(54) $$\limsup_{n\to\infty} A(n)/\sqrt{n} \le \mu_m^* \quad \text{for all} \quad m \ge 1.$$

But the left-hand side in (54) again does not depend on $m$. So, taking $m \to \infty$ in (54), we obtain (52) because $\mu_m^* \to 0$ as $m \to \infty$ by Lemma 1.

At last, consider the third case, when (19) holds with $\mu > 0$. In this case $\mu \le 1$ and

$$q_n := v_{n-1}/v_n = \sqrt{1-\mu_n^2} \to q := \sqrt{1-\mu^2} < 1 \quad \text{as} \quad n \to \infty.$$

For any $0 < \varepsilon < 1 - q$ there exists $m > 0$ such that the inequality $q_j < q + \varepsilon < 1$ holds for any $j \ge m$. Using (54), we have for the chosen $m$ that

$$\mu_m^* a_{m,n}^* \le a_{m,n}^* = \max_{m\le j\le n} \sum_{i=j+1}^{n} \frac{v_{j+r}|x_{i+r}|}{v_{i+r}v_{i+r-1}} \le \max_{m\le j\le n} \sum_{i=j+1}^{n} \frac{v_{j+r}}{v_{i+r-1}}$$

$$= 1 + \max_{m\le j\le n} \sum_{i=j+2}^{n} \prod_{l=j+r+1}^{i+r-1} q_l \le 1 + \sup_{j\ge m} \sum_{i=j+2}^{\infty} (q+\varepsilon)^{i-j-1} = (1-q-\varepsilon)^{-1}.$$

Hence

$$A(n) \le 1 + m + (1-q-\varepsilon)^{-1} < \infty \quad \forall n > 0.$$

Thus, (52) is true also in the third case. □

4.3. **Properties of random step-functions.** Remind that for the standard Wiener process $W(t)$

(55) $$\forall T \ge t \ge 0 \quad \mathbb{E}[W(t)W(T)] = \mathbb{E}W^2(t) = t \ge 0.$$

Below we use notation $\tilde{W}_n(\cdot) = S_n(\cdot, \varepsilon_\bullet)/\sigma$ only when $h(t) = h_0(t) = 0$ in (10). Thus, everywhere below $\tilde{W}_n(\cdot)$ is a random step-function, $n = 1, 2, \dots$, and $\lfloor a \rfloor$ denotes the integer part of a real number $a$.

**Lemma 5.** *For all $n \ge 1$*

(56) $$\forall T \ge t \ge 0 \quad t \ge \mathbb{E}[\tilde{W}_n(t)\tilde{W}_n(T)] = \mathbb{E}\tilde{W}_n^2(t) = \frac{\lfloor nt \rfloor}{n} \ge t - \frac{1}{n}.$$

*Proof.* In this case for all $t \ge 0$

$$\tilde{S}_1(t) = \tilde{S}_1(\lfloor t \rfloor) = S_1(\lfloor t \rfloor) = \sum_{i=1}^{\lfloor t \rfloor} w_i/\sigma.$$

It is easy to see now that for all $T \ge t \ge 0$

$$\mathbb{E}[\tilde{S}_1(T)\tilde{S}_1(t)] = \mathbb{E}\tilde{S}_1^2(t) = \lfloor t \rfloor \mathbb{E}w_1^2/\sigma^2 = \lfloor t \rfloor.$$

And (56) follows for random step functions. □

4.4. **Proof of Theorem 1.** Note first of all that, by Property 1, it is sufficient for us to prove the convergence of the finite-dimensional distributions for the processes $\{\tilde{W}_n(t) : t \in [0,1]\}$ instead of $\{W_n(t) : t \in [0,1]\}$.

We are going to use the idea of Wold. For arbitrary integer $m \ge 1$ choose arbitrary numbers

$$0 \le t_1 < \ldots < t_m \le 1 \quad \text{and} \quad c_i \ne 0, \quad i = 1, \dots, m,$$

and introduce random variables

$$\eta := \sum_{i=1}^{m} c_i W(t_i) \quad \text{and} \quad \eta_n = \sum_{i=1}^{m} c_i \tilde{W}_n(t_i). \tag{57}$$

Here $\eta$ has a degenerate normal distribution with

$$\begin{aligned} 0 < \sigma^2(c_\bullet) &:= \mathbb{E}\eta^2 = \sum_{i=1}^{m} c_i^2 \mathbb{E}W^2(t_i) + 2\sum_{i=1}^{m}\sum_{j=i+1}^{m} c_i c_j \mathbb{E}[W_i(t_i)W_j(t_j)] \\ &= \sum_{i=1}^{m} c_i t_i \left( c_i + 2\sum_{j=i+1}^{m} c_j \right) = \sum_{i=1}^{m} c_i C_i t_i \quad \text{with} \quad C_i := c_i + 2\sum_{j=i+1}^{m} c_j. \end{aligned} \tag{58}$$

as it follows from (55). Similarly, by (56)

$$0 \le \sigma_n^2(c_\bullet) := \mathbb{E}\eta_n^2 = \sum_{i=1}^{m} c_i \mathbb{E}\tilde{W}_n^2(t_i) \left( c_i + 2\sum_{j=i+1}^{m} c_j \right) = \sum_{i=1}^{m} c_i C_i \mathbb{E}\tilde{W}_n^2(t_i). \tag{59}$$

Hence, by (56) it is easy to obtain from (58) and (59) that

$$|\sigma_n^2(c_\bullet) - \sigma^2(c_\bullet)| \le \frac{1}{n}\sum_{i=1}^{m} |c_i C_i| \to 0 \quad \text{as} \quad n \to \infty. \tag{60}$$

In particular, $\sigma_n(c_\bullet) > 0$ for all $n \ge n_0$ with some $n_0 < \infty$.

Next, from (57) and (45) we have that for $n \ge n_0$

$$\mathring{\eta}_n := \frac{\eta_n}{\sigma_n(c_\bullet)} = \sum_{j=1-r}^{n} \tilde{A}_j(n)\xi_j \quad \text{with} \quad \tilde{A}_j(n) := \sum_{i=1}^{m} \frac{c_i}{\sigma_n(c_\bullet)} A_j(\lfloor nt_i \rfloor, n). \tag{61}$$

And hence, by (52)

$$\tilde{A}(n) := \max_{1\le j\le n} |\tilde{A}_j(n)| \le \frac{A(n)}{\sqrt{n}\sigma_n(c_\bullet)} \sum_{i=1}^{m} |c_i| \to 0 \quad \text{as} \quad n \to \infty. \tag{62}$$

But condition (62) means, that we have right to apply CLT to the sum $\mathring{\eta}_n$ of weighted random variables introduces in (61). As a result we obtain, that

$$\mathring{\eta}_n = \frac{\eta_n}{\sigma_n(c_\bullet)} \Rightarrow \mathring{\eta} := \frac{\eta}{\sigma(c_\bullet)} \quad \text{as} \quad n \to \infty, \tag{63}$$

where $\mathring{\eta}$ has the standard normal distribution.

By the idea of Wold, convergences (63) and (60) imply the assertion of Theorem 1.

Underline, that the fact that condition (62) imply that $\mathring{\eta}_n$ is a sum of independent random variables with Lindeberg condition, was independently proved by many authors. See, for example, Hájek and Šidák (1967), p. 153.

Note that condition (19) plays the key role in Lemma 4, so generalizations of Theorem 1 must deal with new ideas of its proof.

## 5. Proof of Theorem 2

Here our proof differs from the both approaches used by Sen (1982) in the proofs of his Theorems 1 and 2. We use condition (22) in Lemma 8 to match assumptions of Theorem 1 from Móricz (1976) with the function of the special form.

5.1. **Useful equalities.** We may now rewrite (42) in the next form

$$w_i/\sigma = \xi_i - \delta_i \quad \text{with} \quad \delta_i = \sum_{l=1-r}^{\infty} a_{il}\xi_l, \quad i = 1, 2, \dots, \tag{64}$$

where

$$a_{il} = \begin{cases} 1 - \frac{1}{\alpha_{r+i}}, & \text{when } l = i; \\ a_{r+i}x_{r+l}, & \text{if } -r < l < i; \\ 0, & \text{otherwise.} \end{cases} \tag{65}$$

**Lemma 6.** *For all* $i > 0$

$$\mathbb{E}\delta_i^2 = 2a_{ii} = 2\left(1 - \frac{1}{\alpha_{r+i}}\right) \geq 0, \tag{66}$$

$$\mathbb{E}[\delta_i\delta_j] = a_{ij} + a_{ji} = a_{r+i}x_{r+j}, \quad \textit{if} \quad 0 < j < i. \tag{67}$$

*Proof.* It is easy to see from (64) that for all possible $i, j$

$$\mathbb{E}[w_iw_j]/\sigma^2 = \mathbb{E}[(\xi_i - \delta_i)(\xi_j - \delta_j)] = \mathbb{E}[\xi_i\xi_j] - \mathbb{E}[\delta_i\xi_j] - \mathbb{E}[\xi_i\delta_j] + \mathbb{E}[\delta_i\delta_j]. \tag{68}$$

Remind that for all such $i, j$ by (8) and (9)

$$\mathbb{E}[w_iw_j]/\sigma^2 = \mathbb{E}[\xi_i\xi_j].$$

Hence, we have from (68) and (64) that

$$\mathbb{E}[\delta_i\delta_j] = \mathbb{E}[\delta_i\xi_j] + \mathbb{E}[\xi_i\delta_j] = a_{ij} + a_{ji} \tag{69}$$

for all possible $i, j$. Substituting now (65) into (69), we obtain (66) and (67). □

Note now that for all $m > k \geq 0$

$$W_{k,m} := \sum_{i=k+1}^{m} w_i/\sigma = \widehat{S}_{k,m} - S_{k,m}, \tag{70}$$

where

$$\widehat{S}_{k,m} := \sum_{i=k+1}^{m} \xi_i \quad \text{and} \quad S_{k,m} = \sum_{i=k+1}^{m} \delta_i. \tag{71}$$

**Lemma 7.** *For all* $m > k \geq 0$

$$\mathbb{E}S_{k,m}^2 = 2\sum_{i=k+1}^{m}\left(1 - \frac{1}{\alpha_{r+i}}\right) + 2\sum_{i=k+1}^{m} a_{r+i}\sum_{j=k+1}^{i-1} x_{r+j}, \quad a_i = \frac{x_i}{v_iv_{i-1}}. \tag{72}$$

*Proof.* We have from (71) that

$$\mathbb{E}S_{k,m}^2 = \sum_{i=k+1}^{m} \mathbb{E}\delta_i^2 + 2\sum_{i=k+1}^{m}\sum_{j=k+1}^{i-1} \mathbb{E}[\delta_i\delta_j]. \tag{73}$$

Substituting now (66) and (67) into (73), we obtain (72). □

5.2. **Main Lemmas.** Note, that condition (22) may be rewritten in the next form:

$$\forall l > r \quad |x_l| \le K_1 v_{l-1}/\sqrt{l} \quad \text{for some} \quad K_1 < \infty. \tag{74}$$

**Lemma 8.** *For all $m > k \ge 0$*

$$\mathbb{E}S_{k,m}^2 \le 4K_1^2(\sqrt{m} - \sqrt{k})^2. \tag{75}$$

*Proof.* By (74) for all $i > j > 0$

$$a_{r+i}x_{r+j} = \frac{x_{r+i}x_{r+j}}{v_{r+i-1}v_{r+i}} \le \frac{|x_{r+i}|}{v_{r+i-1}} \frac{|x_{r+j}|}{v_{r+j-1}} \le \frac{K_1}{\sqrt{i}} \frac{K_1}{\sqrt{j}}.$$

Hence

$$R_1 := \sum_{i=k+1}^{m} \sum_{j=k+1}^{i-1} a_i x_j \le \sum_{i=k+1}^{m} \sum_{j=k+1}^{i-1} \frac{K_1}{\sqrt{i}} \frac{K_1}{\sqrt{j}}. \tag{76}$$

On the other hand

$$1 - \frac{1}{\alpha_{r+i}} = \frac{v_{r+i} - v_{r+i-1}}{v_{r+i}} = \frac{x_{r+i}^2}{v_{r+i}(v_{r+i} + v_{r+i-1})} \le \frac{x_{r+i}^2}{2v_{r+i-1}^2} \le \frac{K_1^2}{2\sqrt{i}^2}.$$

From this inequality and (76) we have:

$$R_2 := 2\sum_{i=k+1}^{m} \left(1 - \frac{1}{\alpha_{r+i}}\right) + 2R_1 \tag{77}$$

$$\le \sum_{i=k+1}^{m} \frac{K_1^2}{\sqrt{i}^2} + 2\sum_{i=k+1}^{m} \sum_{j=k+1}^{i-1} \frac{K_1}{\sqrt{i}} \frac{K_1}{\sqrt{j}} = \sum_{i=k+1}^{m} \frac{K_1}{\sqrt{i}} \sum_{j=k+1}^{m} \frac{K_1}{\sqrt{j}}.$$

Remind that

$$\sum_{i=k+1}^{m} \frac{1}{\sqrt{i}} < \int_k^m \frac{dx}{\sqrt{x}} = 2(\sqrt{m} - \sqrt{k}).$$

From this fact and (77) we obtain (75). □

Now introduce into consideration random variables

$$M_{k,m} = \max_{k<j\le m} |S_{k,j}|, \quad 0 \le k < m.$$

**Lemma 9.** *For all $m > k \ge 0$*

$$\mathbb{E}M_{k,m}^2 \le 47K_1^2(\sqrt{m} - \sqrt{k})^2. \tag{78}$$

*Proof.* We are going to apply Theorem 1 from the paper of Móricz (1976) with the following parameters:

$$\gamma = \alpha = 2 \quad \text{and} \quad g(F_{k,m}) = 2K_1(\sqrt{m} - \sqrt{k}). \tag{79}$$

(Here we use unusual notation $g(F_{k,m})$ from the mentioned paper.) Note that the values $g(F_{k,m})$ have the next property:

$$g(F_{k,m}) + g(F_{m,n}) = g(F_{k,n}) \quad \text{for all} \quad n > m > k \ge 0. \tag{80}$$

And these notations allows us rewrite (75) in the next form:

$$\mathbb{E}S_{k,m}^2 \le g^2(F_{k,m}) \quad \text{for all} \quad m > k \ge 0. \tag{81}$$

From properties (80) and (81) follows that we may apply Theorem 1 from Móricz (1976) with the parameters given in (79). As a result we obtain that

$$\mathbb{E}M_{k,m}^2 \le C_{\alpha,\gamma} g^2(F_{k,m}) = 4K_1^2 C_{\alpha,\gamma}(\sqrt{m} - \sqrt{k})^2,$$

where

$$C_{\alpha,\gamma} = \frac{1}{(1-2^{(1-\alpha)/\gamma})^\gamma} = \frac{1}{(1-2^{-1/2})^2} = \frac{\sqrt{2}^2}{(\sqrt{2}-1)^2} = \frac{2(\sqrt{2}+1)^2}{(\sqrt{2}^2-1)^2} = 2(3+2\sqrt{2}).$$

So, $4C_{\alpha,\gamma} < 47$, and (78) follows. □

Consider now arbitrary integers $n > m > 0$ and $k_1, \dots, k_m$ which satisfy the following conditions:

$$0 = k_0 < k_1 < \ldots < k_m = n,$$

and introduce random variable

$$\overline{M}_n := \max_{1\le i\le m} M_{k_{i-1},k_i}, \quad n > 0. \tag{82}$$

**Lemma 10.** *For all* $x > 0$

$$\mathbb{P}(\overline{M}_n > x) \le \frac{47K_1^2\sqrt{n}}{x^2} \max_{1\le i\le m}(\sqrt{k_i} - \sqrt{k_{i-1}}). \tag{83}$$

*Proof.* By Chebyshev inequality

$$\begin{aligned} \mathbb{P}(\overline{M}_n > x) &\le \sum_{i=1}^m \mathbb{P}(M_{k_{i-1},k_i} > x) \le \frac{1}{x^2}\sum_{i=1}^m \mathbb{E}M_{k_{i-1},k_i}^2 \\ &\le \frac{1}{x^2}\sum_{i=1}^m 47K_1^2(\sqrt{k_i} - \sqrt{k_{i-1}})^2, \end{aligned} \tag{84}$$

where we used (78).

Since $\sum_{i=1}^m(\sqrt{k_i} - \sqrt{k_{i-1}}) = \sqrt{k_m} - \sqrt{k_0} = \sqrt{n}$, inequality (83) follows from (84). □

5.3. **Proof of Theorem 2.** Using formulas (10) and (11) with $h(t) = h_1(t) = t$ introduce into consideration the following random broken lines:

$$\widehat{S}_n(t) := S_n(t; \xi_\bullet) \quad \text{and} \quad S_n(t) := S_n(t; \delta_\bullet),$$

where variables $\{\xi_i\}$ and $\{\delta_i\}$ were defined in (64). It follows from (70) that

$$W_n(t) := S_n(t; w_\bullet)/\sigma = \widehat{S}_n(t) + S_n(t). \tag{85}$$

For arbitrary integer $n > 1$ and real $\delta \in (0, 1/2)$ introduce integers

$$m = m(\delta) = \lfloor 1/\delta \rfloor \quad \text{and} \quad k_i = k_i(n,\delta) = \lfloor in\delta \rfloor + 1, \tag{86}$$

for $i = 1, \dots, m-1$, with $k_0 = 0$ and $k_m = n$. It is easy to see that in this case for each $i = 1, \dots, m$

$$\sqrt{k_i} - \sqrt{k_{i-1}} = \frac{k_i - k_{i-1}}{\sqrt{k_i} + \sqrt{k_{i-1}}} \le \sqrt{n}\left(\sqrt{\delta} + \frac{2}{n\sqrt{\delta}}\right), \tag{87}$$

because $\sqrt{k_i} \ge \sqrt{k_1} > \sqrt{n\delta}$ for each $i = 1, \dots, m$.

Now for $m$ defined in (86) introduce

$$t_0 = 0, \quad t_m = 1 \quad \text{and} \quad t_i = k_i/n, \quad i = 1, \dots, m.$$

It is not difficult to verify that

$$\Delta_i := \max_{t_{i-1} \le t \le t_i} |S_n(t) - S_n(t_{i-1})| = \frac{M_{k_{i-1},k_i}}{\sqrt{n}}, \quad i = 1, , \dots, m.$$

Hence, for every $\varepsilon > 0$

$$\mathbb{P}(\overline{\Delta}_n := \max_{1 \le i \le n} \Delta_i > \varepsilon) = \mathbb{P}(\overline{M}_n > \varepsilon\sqrt{n}),$$

where we use notation $\overline{M}_n$ from (82). Now, using (83) and (87) we obtain with $x = \varepsilon\sqrt{n} > 0$ that

$$\mathbb{P}(\overline{\Delta}_n > \varepsilon) \le \frac{47K_1^2\sqrt{n}}{\varepsilon^2 n} \max_{1 \le i \le m} (\sqrt{k_i} - \sqrt{k_{i-1}}) \le \frac{47K_1^2}{\varepsilon^2} \left(\sqrt{\delta} + \frac{2}{n\sqrt{\delta}}\right).$$

Hence,

$$\forall \varepsilon > 0 \quad \lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}(\overline{\Delta}_n > \varepsilon) \le \lim_{\delta \to 0} \frac{47K_1^2\sqrt{\delta}}{\varepsilon^2} = 0. \tag{88}$$

Thus, the tightness of the family of distributions of processes $\{S_n(\cdot)\}$ follows from (88).

The tightness of distributions in $C[0,1]$ of the classical processes $\{\widehat{S}_n(\cdot)\}$ is well known. These facts and equality (85) imply, that the distributions of $\{W_n(\cdot)\}$ are also tight. By Prokhorov theorem we have now the convergence (13) of the distributions in functional space $C[0,1]$, because the convergence of the corresponding finite-dimensional distributions was proved in Theorem 1.

So, all results of the paper are proved.

**Acknowledgement**

The work is supported by Mathematical Center in Akademgorodok under agreement No. 075-15-2019-1675 with the Ministry of Science and Higher Education of the Russian Federation.

## References

[1] W. Bischoff, *A functional central limit theorem for regression models*, Ann. Stat., **26**:4 (1998), 1398–1410. Zbl 0936.62072

[2] W. Bischoff, *On designs for recursive least squares residuals to detect alternatives*, in J. Kunert, C.H. Müller, A.C. Atkinson (eds.), *mODa 11 – Advances in Model-Oriented Design and Analysis*, Contributions to Statistics. Springer, 2016, 37–45.

[3] R.L. Brown, J. Durbin, J.M. Evans, *Techniques for testing the constancy of regression relationships over time. Discussion*, J. R. Stat. Soc., Ser. B, **37**:2 (1975), 149–192. Zbl 0321.62063

[4] J. Hájek, Z. Šidák. *Theory of rank tests*, Academic Press, New York–London, 1967. Zbl 0161.38102

[5] I.B. MacNeill, *Limit processes for sequences of partial sums of regression residuals*, Ann. Probab., **6**:4 (1978), 695–698. Zbl 0377.60028

[6] F.Móricz, *Moment inequalities and the strong laws of large numbers*, Z. Wahrscheinlichkeitstheorie Verw. Geb., **35** (1976), 299–314. Zbl 0314.60023

[7] P.K. Sen, *Invariance principles for recursive residuals*, Ann. Stat., **10**:1 (1982), 307–312. Zbl 0484.62098

[8] A. Zeileis, F. Leisch, K. Hornik, Ch. Kleiber, *strucchange: An R package for testing for structural change in linear regression models*, J. Stat. Softw., **7**:2 (2002), 1–38.

Alexander Sakhanenko
Sobolev Institute of Mathematics,
4, Acad. Koptyug ave.,
Novosibirsk, 630090, Russia
Novosibirsk State University,
1, Pirogova str.,
Novosibirsk, 630090, Russia
*Email address*: aisakh@mail.ru

Artyom Kovalevskii
Novosibirsk State Technical University,
20, K. Marksa ave.,
Novosibirsk, 630073, Russia
Novosibirsk State University,
1, Pirogova str.,
630090, Novosibirsk, Russia
*Email address*: artyom.kovalevskii@gmail.com

Anastasiya Shelepova
Novosibirsk State University,
1, Pirogova str.,
Novosibirsk, 630090, Russia
*Email address*: nastya.shelepova.99@mail.ru