ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2022 Управление, вычислительная техника и информатика № 59
Tomsk State University Journal of Control and Computer Science

Original article
doi: 10.17223/19988605/59/9

Non-asymptotic Confidence Estimation of the Autoregressive Parameter in AR(1) Process by Noisy Observations

Sergey E. Vorobeychikov[1], Andrey V. Pupkov[2]

[1, 2] Tomsk State University, Tomsk, Russian Federation
[1] sev@mail.tsu.ru
[2] andrewpupkov@gmail.com

Abstract. For parameter in an AR(1) process corrupted by noise, the paper proposes the construction of confidence interval for unknown parameter with a prescribed coverage probability. The noises both in observable and in unobservable processes are assumed to be Gaussian with unknown variance. The estimation procedure is non-asymptotic and uses a special stopping rule. The results of numerical simulation by Monte-Carlo method are presented.
Keywords: autoregressive process; non-asymptotic estimation; confidence interval

***For citation*:** Vorobeychikov, S.E., Pupkov, A.V. (2022) Non-asymptotic confidence estimation of the autoregressive parameter in AR(1) process by noisy observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 59. pp. 83–90. doi: 10.17223/19988605/59/9

Научная статья
УДК 519.246.2
doi: 10.17223/19988605/59/9

Доверительное неасимптотическое оценивание параметра авторегрессии AR(1) по зашумленным данным

Сергей Эрикович Воробейчиков[1], Андрей Викторович Пупков[2]

[1, 2] Томский государственный университет, Томск, Россия
[1] sev@mail.tsu.ru
[2] andrewpupkov@gmail.com

Аннотация. Рассматривается задача построения доверительного интервала неизвестного параметра процесса авторегрессии первого порядка, зашумленного аддитивным шумом. Предполагается, что управляющий шум процесса и шум в канале наблюдений – гауссовские с неизвестными дисперсиями. Построенная процедура является неасимптотической и опирается на специальное правило остановки наблюдений. В статье приводятся результаты численного моделирования, реализованного методом Монте-Карло.
Ключевые слова: процесс авторегрессии; неасимптотическое оценивание; доверительный интервал

***Для цитирования*:** Воробейчиков С.Э., Пупков А.В. Доверительное неасимптотическое оценивание параметра авторегрессии AR(1) по зашумленным данным // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 59. С. 83–90. doi: 10.17223/19988605/59/9

In many applications related to signal processing the linear models specified by stochastic difference equations are widely used. To identify unknown parameters in such models the methods of least squares

(LSE), maximum likelihood and other have been developed. The properties of the estimators are studied usually in asymptotic when the number of observations tends to infinity. For an autoregressive process the asymptotic properties of the LSE have been studied in [1]. To overcome the problems of investigation the properties of the estimators obtained by a fixed number of observations the sequential methods were developed. Sequential procedures use the sampling schemes with random stopping times. It allows one to study the properties of the estimators. The sequential sampling scheme was proposed in works [2] and [3] to estimate the parameter of a first order autoregressive process

$$x_k = \theta x_{k-1} + \varepsilon_k,\ k = 1, 2, \ldots.$$

The estimator in [2] has the guaranteed mean square deviation.

The problem of estimation of parameters in ARMA processes and in AR processes with noise was studied in [4–7]. A sequential procedure of identification parameters of autoregressive processes by noisy observations was proposed in [8]. This procedure uses the Yule-Walker estimators with guaranteed mean square deviation.

A problem of confidence estimation of the mean in a sequence of independent identically distributed Gaussian variables with unknown variance was studied in [9]. A sequential procedure was proposed because no procedure with fixed sample size can guarantee the prescribed coverage probability.

Recently a procedure for constructing a fixed-size confidence interval with any prescribed coverage probability for parameter in AR(1) process was proposed in [10]. The interval estimation for a first-order autoregressive process was constructed in [11].

The paper is organized as follows. In Section 1 we construct a sequential point estimator of unknown parameter in AR(1) process with noises. In section 2 we construct a confidence interval with any prescribed coverage probability for the parameter in AR(1) model with unknown noise variances. In section 3 the results of numerical simulation are presented.

1. Sequential point estimator

Consider an unobservable AR(1) process described by equations

$$x_k = \theta x_{k-1} + \varepsilon_k,\ k = 1, 2, \ldots, \tag{1}$$

which is observed with noises

$$y_k = x_k + \eta_k,\ \ k = 1, 2, \ldots. \tag{2}$$

Here θ is unknown parameter, $|\theta| \le 1$, $\{\varepsilon_k\}$ and $\{\eta_k\}$ are the sequences of Gaussian independent random variables with zero means $\mathbf{E}\varepsilon_k = \mathbf{E}\eta_k = 0$ and variances $\mathbf{E}\varepsilon_k^2 = \sigma^2$; $\mathbf{E}\eta_k^2 = \Delta^2$ respectively, initial value x_0 and processes $\{\varepsilon_k\}$, $\{\eta_k\}$ are independent. It's assumed that σ^2 and Δ^2 are unknown. The problem is to construct the non-asymptotic confidence interval for the parameter θ using observations $\{y_k\}$.

Note that from (1) and (2), we obtain the equation for the observed process

$$y_k = \theta y_{k-1} + \xi_k, \tag{3}$$

where $\xi_k = \varepsilon_k + \eta_k - \theta\eta_{k-1}$, $\mathbf{E}\xi_k = 0$ and $\mathbf{E}\xi_k^2 = \Delta^2(1+\theta^2) + \sigma^2$.

First we obtain a point estimator of an unknown parameter of AR(1) process. The estimator is used later to construct the confidence interval. The scheme of estimating the parameter θ follows the approach proposed in [10] and includes three stages.

First we obtain the pilot Yule-Walker estimator of θ by a fixed number of observations. On the second stage we construct an estimator of the variance of the random variable ξ_k. On the third stage we construct a sequential modification of the Yule-Walker estimator of the parameter θ.

For an integer $n_1 \ge 3$ define the Yule-Walker estimator of the parameter θ as

$$\tilde{\theta}(n_1) = \left(\sum_{k=3}^{n_1} y_{k-2} y_{k-1}\right)^{-1} \sum_{k=3}^{n_1} y_{k-2} y_k. \tag{4}$$

To compensate an unstable behavior of estimator (4) for small sample size we use the projection of the estimator into the interval $[-1,1]$:

$$\hat{\theta} = \hat{\theta}(n_1) = \begin{cases} \tilde{\theta}(n_1), & \text{if } |\tilde{\theta}(n_1)| < 1, \\ 1, & \text{if } \tilde{\theta}(n_1) \ge 1, \\ -1, & \text{if } \tilde{\theta}(n_1) \le -1. \end{cases}$$

Using the estimator $\hat{\theta}(n_1)$ define the estimator of the variance ξ_k as

$$\Gamma_{n_1,n_2} = C(n_2) S_{n_1,n_2}, \tag{5}$$

where

$$S_{n_1,n_2} = \sum_{k=n_1+1}^{n_1+n_2} (y_k - \hat{\theta} y_{k-1})^2, \qquad C(n_2) = \mathbf{E}\left(\sum_{k=1}^{n_2} \nu_k^2 \right)^{-1}, \qquad \nu_k = \frac{\varepsilon_k + \eta_k}{\sqrt{\Delta^2 + \sigma^2}}. \tag{6}$$

Note that ν_k are standard Gaussian random variables. Hence $\sum_{k=1}^{n_2} \nu_k^2$ is Chi-square random variable with n_2 degrees of freedom and $C(n_2) = (n_2 - 2)^{-1}$. Denote

$$V_{n_1}^l = \sum_{k=n_1+1}^{n_1+l} \left(\varepsilon_k + \eta_k \right)^2, \quad l > 0. \tag{7}$$

The properties of the variables $V_{n_1}^l$ establishes the following theorem

Theorem 1. Let the random processes $\left\{ S_{n_1,l} \right\}$ and $\{V_{n_1}^l\}$ are defined by (5) and (7) respectively. Then the following inequality holds true

$$P_\theta \left(S_{n_1,l} \le z \right) \le P_\theta \left(V_{n_1}^l \le z \right). \tag{8}$$

The proof of Theorem actually proceeds along the lines of the proof of Theorem 1 in [10].

Using Theorem 3 in [10] one has

Theorem 2. The following inequality holds true

$$\mathbf{E} \frac{1}{\Gamma_{n_1,n_2}} \le \frac{1}{\Delta^2 + \sigma^2}.$$

This inequality allows one to obtain the upper bound of the variance ξ_k.

Now we construct a sequential point estimator of the parameter θ. The estimator is based on the estimator proposed in [12] when the variances σ^2 and Δ^2 in (1) and (2) are known.

To construct sequential point estimator of the parameter θ, we divide the set of indexes $T(n) = \{1, 2, \ldots, n\}$ into two subsets

$$T_1(n) = \{2j - 1 : j = 1, 2, \ldots; 2j - 1 \le n\}, \quad T_2(n) = \{2j : j = 1, 2, \ldots; 2j \le n\}.$$

Note that $T(n) = T_1(n) + T_2(n)$. Next we introduce stopping times

$$\tau_i(h) = \inf \left\{ n \ge n_1 + n_2 + 1 : \sum_{k=n_1+n_2+1}^{n} \frac{y_{k-2}^2}{\Gamma_{n_1,n_2}} \chi_{\{k \in T_i(n)\}} \ge h \right\}, \qquad i = 1, 2. \tag{9}$$

and $\chi_{\{A\}}$ is indicator of event A. The value h is a parameter of the identification procedure and defines the quality of the estimator.

Define sequential estimators of the parameter θ from odd and even observations by formulas

$$\hat{\theta}_i(h) = \left(\sum_{k=n_1+n_2+1}^{\tau_i(h)} \sqrt{\beta_{i,k}} \frac{y_{k-2} y_{k-1}}{\sqrt{\Gamma_{n_1,n_2}}} \right)^{-1} \sum_{k=n_1+n_2+1}^{\tau_i(h)} \sqrt{\beta_{i,k}} \frac{y_{k-2} y_k}{\sqrt{\Gamma_{n_1,n_2}}}, \qquad i = 1, 2, \tag{10}$$

where the coefficients $\beta_{i,k}$, $i=1,2$, are defined as

$$\beta_{i,k}=\begin{cases}1, & \text{if } k\in T_i(\tau_i(h)) \text{ and } k<\tau_i(h);\\ \alpha_i(h), & \text{if } k=\tau_i(h);\\ 0, & \text{if } k\notin T_i(\tau_i(h)).\end{cases} \tag{11}$$

Correction factors $0<\alpha_i(h)\le 1, i=1,2$ are determined by the equalities

$$\sum_{k=n_1+n_2+1}^{\tau_i(h)-1}\frac{y_{k-2}^2}{\Gamma_{n_1,n_2}}\chi_{\{k\in T_i(n)\}}+\alpha_i(h)\frac{y_{\tau_i(h)-2}^2}{\Gamma_{n_1,n_2}}=h.$$

Consider the deviation of the estimators (10). Substituting y_k from (3) into (10) we have

$$\hat{\theta}_i(h)-\theta=\frac{1}{s_i(h)}\sum_{k=n_1+n_2+1}^{\tau_i(h)}\sqrt{\beta_{i,k}}\,\frac{y_{k-2}(\theta y_{k-1}+\xi_k)}{\sqrt{\Gamma_{n_1,n_2}}}-\theta=\frac{\zeta_i(h)}{s_i(h)},\quad i=1,2,$$

where

$$\zeta_i(h)=\sum_{k=n_1+n_2+1}^{\tau_i(h)}\sqrt{\beta_{i,k}}\,\frac{y_{k-2}\xi_k}{\sqrt{\Gamma_{n_1,n_2}}},\qquad s_i(h)=\sum_{k=n_1+n_2+1}^{\tau_i(h)}\sqrt{\beta_{i,k}}\,\frac{y_{k-2}y_{k-1}}{\sqrt{\Gamma_{n_1,n_2}}}. \tag{12}$$

Denote

$$\tilde{\zeta}_i(h)=\sum_{k=n_1+n_2+1}^{\tau_i(h)}\sqrt{\beta_{i,k}}\,\frac{y_{k-2}\tilde{\xi}_k}{\sqrt{\Gamma_{n_1,n_2}}},\quad \tilde{\xi}_k=\kappa_\theta^{-1}\xi_k,\quad \kappa_\theta=\sqrt{\Delta^2(1+\theta^2)+\sigma^2}. \tag{13}$$

Hence, the deviation $\hat{\theta}_i(h)-\theta$ takes the form

$$\hat{\theta}_i(h)-\theta=\kappa_\theta\frac{\tilde{\zeta}_i(h)}{s_i(h)},\qquad i=1,2.$$

According to Lemma 2 in [12] the following result holds.

Lemma 1. Let $\{\varepsilon_k\}$ and $\{\eta_k\}$ in (3) be the sequences of Gaussian random variables with zero means and variances $\mathbf{E}\varepsilon_k^2=\sigma^2$, $\mathbf{E}\eta_k^2=\Delta^2$, and random variables $\tilde{\zeta}_i(h)$, $i=1,\ 2$ are described by the equations (13). Then for each $h>0$ the random variables $h^{-1/2}\tilde{\zeta}_1(h)$ and $h^{-1/2}\tilde{\zeta}_2(h)$ are standard Gaussian.

In the next section we construct a non-asymptotic confidence interval for parameter θ.

2. Non-asymptotic confidence interval

The main result of the paper gives the following Theorem.

Theorem 3. Let $\{\varepsilon_k\}$ and $\{\eta_k\}$ in (3) be the sequences of Gaussian random variables with zero means and variances $\mathbf{E}\varepsilon_k^2=\sigma^2$, $\mathbf{E}\eta_k^2=\Delta^2$, and the sequential point estimators $\hat{\theta}_i(h), i=1,2$ are defined by formulas (9), (10), (11). Then for all $h>0$ and $z>0$

$$P_\theta\left(\frac{|s_*(h)|}{h\sqrt{\Gamma_{n_1,n_2}}}\left|\frac{\hat{\theta}_1(h)+\hat{\theta}_2(h)}{2}-\theta\right|<z\right)\ge 1-\alpha(h,z), \tag{14}$$

where Γ_{n_1,n_2} is defined by (5) and $s_*(h)=\min\left(s_1(h),s_2(h)\right)$ and

$$\alpha(h,z)=4\int_0^\infty\left[1-\Phi\left(z\sqrt{yhC(n_2)}\right)\right]\frac{y^{n_2/2-1}}{\Gamma(n_2/2)}\exp(-y)dy.$$

Here $C\left(n_2\right)$ is defined by (6), $\Gamma(n_2/2)$ is the Gamma-function and

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left(-\frac{y^2}{2}\right)dy. \tag{15}$$

Proof: Using estimators (5) and (10) we obtain the inequality

$$\frac{|s_*(h)|}{h\sqrt{\Gamma_{n_1,n_2}}}\left|\frac{\hat\theta_1(h)+\hat\theta_2(h)}{2}-\theta\right| = \frac{\kappa_\theta\,|s_*(h)|}{2h\sqrt{\Gamma_{n_1,n_2}}}\left|\frac{\tilde\zeta_1(h)}{s_1(h)}+\frac{\zeta_2(h)}{s_2(h)}\right| \le$$
$$\le \frac{\kappa_\theta\,|s_*(h)|}{2h\sqrt{\Gamma_{n_1,n_2}}}\left[\left|\frac{\tilde\zeta_1(h)}{s_1(h)}\right|+\left|\frac{\zeta_2(h)}{s_2(h)}\right|\right] \le \frac{\kappa_\theta}{2h\sqrt{\Gamma_{n_1,n_2}}}\left[|\tilde\zeta_1(h)|+|\zeta_2(h)|\right]. \tag{16}$$

Introduce the filtration

$$\mathcal{F}_0=\sigma\{y_0\},\quad \mathcal{F}_n=\sigma\{y_0,\varepsilon_1,\ldots,\varepsilon_n,\eta_1,\ldots,\eta_n\}.$$

To construct the confidence interval, we obtain the upper bound for probability

$$P_\theta\left(\frac{|s_*(h)|}{h\sqrt{\Gamma_{n_1,n_2}}}\,\left|\frac{\hat\theta_1(h)+\hat\theta_2(h)}{2}-\theta\right|\ge z\right) \le P_\theta\left(\frac{\kappa_\theta}{2h\sqrt{\Gamma_{n_1,n_2}}}\left[|\tilde\zeta_1(h)|+|\zeta_2(h)|\right]\ge z\right)=$$
$$=\mathbf{E}P_\theta\left(\frac{1}{\sqrt{h}}\left[|\tilde\zeta_1(h)|+|\tilde\zeta_2(h)|\right]\ge\frac{2z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right). \tag{17}$$

Then

$$P_\theta\left(\frac{1}{\sqrt{h}}\left[|\tilde\zeta_1(h)|+|\tilde\zeta_2(h)|\right]\ge\frac{2z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)\le P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}\ge\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)+$$
$$+P_\theta\left(\frac{|\tilde\zeta_2(h)|}{\sqrt{h}}\ge\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)=2P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}\ge\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right). \tag{18}$$

As a result from (17) and (18) we obtain

$$\mathbf{E}P_\theta\left(\frac{1}{\sqrt{h}}\left(|\tilde\zeta_1(h)|+|\tilde\zeta_2(h)|\right)<\frac{2z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)\ge 1-2\mathbf{E}P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}\ge\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)=$$
$$=2\left[1-\mathbf{E}P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}\ge\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\right)\right]-1=2\mathbf{E}P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}<\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)-1. \tag{19}$$

Let $G(y)$ be the distribution function of random variable $\Gamma_{n_1,n_2}/\kappa_\theta^2$. Taking into account that $|\theta|\le 1$ we come to inequality $\kappa_\theta^2=\Delta^2(1+\theta^2)+\sigma^2\le 2(\Delta^2+\sigma^2)$. Using expression (5) for Γ_{n_1,n_2} and inequality (8) we obtain the upper bound for the distribution function $G(y)$

$$P_\theta\left(\frac{\Gamma_{n_1,n_2}}{\kappa_\theta^2}<z\right)=P_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}(y_k-\hat\theta y_{k-1})^2<\frac{z\kappa_\theta^2}{C(n_2)}\right)\le P_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}(\varepsilon_k+\eta_k)^2<\frac{z\kappa_\theta^2}{C(n_2)}\right)\le$$
$$\le P_\theta\left(\left(\sigma^2+\Delta^2\right)\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2<\frac{2z(\sigma^2+\Delta^2)}{C(n_2)}\right)=P_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2<\frac{2z}{C(n_2)}\right), \tag{20}$$

where $\nu_k=\left(\sigma^2+\Delta^2\right)^{-1/2}(\varepsilon_k+\eta_k)$.

Analogously to [10] using integration by parts and (20) we have

$$\mathbf{E}P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}<\frac{z\sqrt{h\Gamma_{n_1,n_2}}}{\kappa_\theta}\;|\mathcal{F}_{n_1+n_2}\right)=\int_0^\infty P_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}<z\sqrt{hy}\;|\mathcal{F}_{n_1+n_2}\right)dG(y)=$$
$$=1-\int_0^\infty G(y)\,dP_\theta\left(\frac{|\tilde\zeta_1(h)|}{\sqrt{h}}<z\sqrt{hy}\;|\mathcal{F}_{n_1+n_2}\right)\ge \tag{21}$$

$$\geq 1-\int_0^\infty P_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2<\frac{2y}{C(n_2)}\right)dP_\theta\left(\frac{|\tilde{\zeta}_1(h)|}{\sqrt{h}}<z\sqrt{hy}\;|\mathcal{F}_{n_1+n_2}\right)=$$

$$=\int_0^\infty P_\theta\left(\frac{|\tilde{\zeta}_1(h)|}{\sqrt{h}}<z\sqrt{hy}\;|\mathcal{F}_{n_1+n_2}\right)dP_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2<\frac{2y}{C(n_2)}\right).$$

Taking into account that $\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2$ has Chi-square distribution with n_2 degrees of freedom, inequalities (17), (19), (21) and Lemma 1 we can construct a non-asymptotic confidence interval for the parameter θ in the form

$$P_\theta\left(\frac{|s_*(h)|}{h\sqrt{\Gamma_{n_1,n_2}}}\,\left|\frac{\hat{\theta}_1(h)+\hat{\theta}_2(h)}{2}-\theta\right|<z\right)\geq$$

$$\geq 2\int_0^\infty P_\theta\left(\frac{|\tilde{\zeta}_1(h)|}{\sqrt{h}}<z\sqrt{hy}\;|\mathcal{F}_{n_1+n_2}\right)dP_\theta\left(\sum_{k=n_1+1}^{n_1+n_2}\nu_k^2<\frac{2y}{C(n_2)}\right)-1=$$

$$=2\int_0^\infty\left[2\Phi\left(z\sqrt{hy}\right)-1\right]\frac{(1/2)^{n_2/2}}{\Gamma(n_2/2)}\left(\frac{2y}{C(n_2)}\right)^{n_2/2-1}\exp\left\{-\frac{y}{C(n_2)}\right\}\frac{2dy}{C(n_2)}-1= \qquad (22)$$

$$=2\int_0^\infty\left[2\Phi\left(z\sqrt{yhC(n_2)}\right)-1\right]\frac{y^{n_2/2-1}}{\Gamma(n_2/2)}\exp(-y)dy-1,$$

where $\Phi(x)$ is determined by the formula (15).

From (22) one has the equality to defining the value *h* of the form

$$\alpha(h,z)=4\int_0^\infty\left[1-\Phi\left(z\sqrt{yhC(n_2)}\right)\right]\frac{y^{n_2/2-1}}{\Gamma(n_2/2)}\exp(-y)dy.$$

We come to the result of Theorem 3.

Remark 1. Note that the random coefficient $s_*(h)/h$ in the confidence interval converges to constant almost surely due to Proposition 1 in [12].

3. Simulation results

In this section, we report and discuss the results of Monte Carlo experiments. Table 1 presents selected data obtained by the simulations.

Represent the interval for parameter θ

$$\frac{|s_*(h)|}{h\sqrt{\Gamma_{n_1,n_2}}}\left|\frac{\hat{\theta}_1(h)+\hat{\theta}_2(h)}{2}-\theta\right|<z \qquad (23)$$

in the form

$$\bar{\theta}(h)-\delta<\theta<\bar{\theta}(h)+\delta, \qquad \bar{\theta}(h)=\frac{\hat{\theta}_1(h)+\hat{\theta}_2(h)}{2}, \qquad (24)$$

where $\delta=zh\sqrt{\Gamma_{n_1,n_2}}/|s_*(h)|$ is the confidence semi-interval.

The results of Monte Carlo simulation are reported in Table. All the results were obtained by 1000 replications by making use of the programming language *R*. The values n_1 and n_2 were chosen equal to 30. The averaged values of sequential estimators $\bar{\theta}(h)$ as well as mean length of semi-interval δ have been calculated for every value of the parameter θ. The confidence probability $1-\alpha(h,z)$ was equal to 0,9 for different values of *h* and *z*. The quantity $\bar{\tau}$ denotes the observed averages of the stopping time $(\tau_1(h)+\tau_2(h))/2$. The quantity $\hat{p}$ denotes the frequency count of the number of times when the confidence interval contains the true values θ.

Averaged confidence estimates of parameter AR(1)

θ	σ^2	Δ^2	h = 779, z = 0,1				h = 3 115, z = 0,05			
			$\overline{\theta}(h)$	δ	$\overline{\tau}$	$\hat{p}$	$\overline{\theta}(h)$	δ	$\overline{\tau}$	$\hat{p}$
–0,99	1	0,25	–0,990	0,098	156,5	1	–0,991	0,05	359,6	1
–0,8	1	0,25	–0,802	0,133	954,3	1	–0,800	0,067	3504,4	0,999
–0,6	1	0,25	–0,598	0,187	1545,9	1	–0,600	0,095	6079,5	1
–0,4	1	0,25	–0,401	0,290	2171,2	0,999	–0,402	0,147	8315,1	1
–0,3	1	0,25	–0,300	0,387	2447,0	0,999	–0,300	0,198	9445,0	1
0,3	1	0,25	0,302	0,448	2420,0	1	0,303	0,212	9627,7	1
0,4	1	0,25	0,400	0,323	2202,5	1	0,402	0,156	8446,0	1
0,6	1	0,25	0,599	0,202	1611,1	1	0,600	0,098	6113,8	1
0,8	1	0,25	0,799	0,140	913,3	1	0,800	0,069	3572,4	1
0,99	1	0,25	0,991	0,103	158,4	1	0,990	0,051	340,0	1
0,99	4	0,25	0,990	0,102	144,9	1	0,990	0,051	293,0	1
0,99	4	1	0,990	0,103	159,4	1	0,990	0,051	347,0	1
0,99	1	4	0,991	0,114	426,0	1	0,990	0,056	1286,3	1

Note that as the variance σ^2 of the process increases, the sample size $\overline{\tau}$ decreases. Increasing the variance Δ^2 of additive noise leads to increasing the sample size $\overline{\tau}$. This fact follows from Proposition 1 in [12].

Conclusions

The proposed sequential procedure allows one to construct the confidence non-asymptotic interval for autoregressive parameter θ in the presence of additive noise in observations. The procedure is independent of variances of noises in unobservable and observable processes. It is based on a special rule of determining the needed sample size. The results can be used in identification and control problems.

References

1. Lai, T.L. & Wei, C.Z. (1983) Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. *Journal of Multivariate Analysis.* 13. pp. 1–23. DOI: 10.1016/0047-259X(83)90002-7
2. Borisov, V.Z. & Konev, V.V. (1977) On sequential parameter estimation in discrete time processes. *Automation and Remote Control.* 38. pp. 1475–1480.
3. Lai, T.L. & Siegmund, D. (1983) Fixed accuracy estimation of an autoregressive parameter. *Annals of Statistics.* 11. pp. 478–485. DOI: 10.1214/aos/1176346154
4. Anderson, T.W. (1971) *The Statistical Analysis of Time Series.* New York: Wiley.
5. Hasan, Md.K., Chowdhury, A.K.M.Z.R. & Khan, M.R. (2003) Identification of autoregressive signals in colored noise using damped sinusoidal model. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications.* 50. pp. 966–969.
6. Mahmoudi, A., Karimi, M. & Amindavar, H. (2012) Parameter estimation of autoregressive signals in presence of colored AR(1) noise as a quadratic eigenvalue problem. *Signal Processing*. 92. pp. 1151–1156. DOI: 10.1016/j.sigpro.2011.11.015
7. Xia, Y. & Zheng, W.X. (2015) Novel parameter estimation of autoregressive signals in the presence of noise. *Automatica.* 62. pp. 98–105. DOI: 10.1016/j.automatica.2015.09.008
8. Vasiliev, V.A. & Konev, V.V. (1985) Sequential identification of parameters in dynamical systems in presence of multiplicative and additive noises. *Automation and Remote Control.* 46. pp. 706–716.
9. Stein, C. (1945) A Two-Sample test for a linear hypothesis' whose power is independent of the variance. *Annals of Mathematical Statistics.* 16. pp. 253–258. DOI: 10.1214/aoms/1177731088
10. Vorobeychikov, S. & Burkatovskaya, Yu. (2020) Non-asymptotic confidence estimation of the autoregressive parameter in AR(1) Process with an Unknown Noise Variance. *Austrian Journal of Statistics.* 49. pp. 19–26. DOI: 10.17713/ajs.v49i4.1121
11. Wei, C., Hao, Y. & Ching, K. (2018) Interval estimation for a first-order positive autoregressive process. *Journal of Time Series Analysis.* 39. pp. 1475–1480. DOI: 10.1111/jtsa.12297
12. Konev, V. & Pupkov, A. (2021) Confidence estimation of autoregressive parameters based on noisy data. *Automation and Remote Control*. 82. pp. 1030–1048. DOI: 10.1134/S0005117921060059

Information about the authors:
Vorobeychikov Sergey E. (Doctor of Physics and Mathematics, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: sev@mail.tsu.ru
Pupkov Andrey V. (Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation).
E-mail: andrewpupkov@gmail.com

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Информация об авторах:
Воробейчиков Сергей Эрикович – доктор физико-математических наук, профессор кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: sev@mail.tsu.ru

Пупков Андрей Викторович – аспирант кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: andrewpupkov@gmail.com

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Received 24.03.2022; accepted for publication 30.05.2022

Поступила в редакцию 24.03.2022; принята к публикации 30.05.2022