МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Международная лаборатория статистики случайных
процессов и количественного финансового анализа

Международная научная конференция «Робастная статистика и финансовая математика – 2017»

(03–05 июля 2017 г.)

Сборник статей

Под редакцией
д-ра физ.-мат. наук, профессора С.М. Пергаменщикова,
канд. физ.-мат. наук, доцента Е.А. Пчелинцева

Томск
Издательский Дом Томского государственного университета
2017

Non-asymptotic distribution of the sequential estimates of parameters in a first-order unstable autoregression with unknown mean [*]

Konev V. V., Nazarenko B. N.
Tomsk State University, Tomsk
e-mail: vvkonev@mail.tsu.ru

Abstract

The paper proposes new sequential estimates for the parameters in an autoregressive process $AR(1)$ with unknown mean. The property of uniform asymptotic normality is proved in the case of non-stable process with unspecified noise distributions. In case of Gaussian noises non-asymptotic distribution of estimates has been derived.

Keywords: sequential estimate, autoregression, fixed-accuracy estimation, uniform asymptotic normality.

Introduction. It is well-known that autoregressive models are widely used in the problems of automatic control, time series analysis, filtering theory, image processing, spectral analysis and others because they provide adequate description of different processes in applications. For estimating parameters in these models a wide variety of methods have been developed. Much efforts have been made to investigate the asymptotic properties of the least squares estimates (LSE) of an autoregressive parameter in the AR(1) model obeying the equation

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

where $\{\varepsilon_k\}$ is a sequence of i.i.d. random variables with $\mathsf{E}\varepsilon_k = 0$, $0 < \mathsf{E}\varepsilon_k^2 = \sigma^2 < \infty$. It is well-known (see [5] for details and other references) that the LSE based on $(x_0, \ldots, x_n)$

$$\widehat{\theta}_n = \frac{\sum\limits_{k=1}^{n} x_{k-1}x_k}{\sum\limits_{k=1}^{n} x_{k-1}^2} \tag{2}$$

has three different limiting distributions, depending on the value of unknown parameter θ each demanding its own normalizing factor.

Lai and Siegmund [4] proposed for estimating parameter θ in the model (1) to use a sequential sampling scheme based on the stopping time
$$\tau = \tau(h) = \{n \geqslant 1 : \ I_n \geqslant h\}, \ h > 0,$$
where $I_n = \sum\limits_{k=1}^{n} x_{k-1}^2$ is the observed Fisher information. They have proved that the sequential estimate $\widehat{\theta}_{\tau(h)}$ obtained from (2) by replacing n with $\tau(h)$, has the important property of uniform asymptotic normality
$$\lim_{h \to \infty} \sup_{|\theta| \leqslant 1} \left[P_\theta \left(I_{\tau(h)}^{1/2} (\theta_{\tau(h)} - \theta) \leqslant z \right) - \Phi \left(\frac{z}{\sigma} \right) \right] = 0.$$

In this paper we consider the problem of estimating two unknown parameters θ_1 and θ_2 in the model of type
$$x_k = \theta_1 x_{k-1} + \theta_2 + \varepsilon_k, \ k \geqslant 1, \tag{3}$$
where $(\varepsilon_k)_{k\geqslant 1}$ is a sequence of i.i.d. random variables with $\mathsf{E}\varepsilon_k = 0$, $0 < \mathsf{E}\varepsilon_k^2 = \sigma^2 < \infty$.

As is pointed out in [5], the multiparametric case is more complicated and should be treated differently to obtain the estimator with the property of uniform asymptotic normality. Galtchouk and Konev [2] treated problem of estimating parameters θ_1 and θ_2 in the model (3) from the standpoint of sequential analysis. To construct the sequential estimate of the vector $\theta = (\theta_1, \theta_2)'$ they use the LS-estimate based on the sample $(x_0, \ldots, x_n)$
$$\widehat{\theta}(n) = M_n^{-1} \sum_{k=1}^{n} Y_{k-1} x_k,$$
where $Y_k = (x_k, 1)'$, the prime denotes the transposition; M_n is the sample Fisher information matrix, that is
$$M_n = \sum_{k=1}^{n} Y_{k-1} Y_{k-1}'. \tag{4}$$

The sequential estimate is defined as
$$\widehat{\theta}_{\tau(h)} = M_{\tau(h)}^{-1} \sum_{k=1}^{\tau(h)} Y_{k-1} x_k \tag{5}$$
where $h > 0$, $\tau(h)$ is the stopping time of the form
$$\tau(h) = \inf \left\{ n \geqslant 1 : \ \sum_{k=1}^{n} \|Y_{k-1}\| \geqslant h \right\}, \ \inf\{\emptyset\} = \infty,$$
$$\|Y_{k-1}\|^2 = 1 + x_{k-1}^2.$$

The paper [2] has established the uniform asymptotic normality property of estimate (5) which is given in the following theorem.

Theorem 1. *Let the process (3) be stable, $|\theta_1| < 1$, and (ε_n) be i.i.d. with mean 0 and unit variance and be independent of x_0. Then for any $0 \leqslant l < \infty$, $0 < s < 1$,*

$$\lim_{h \to \infty} \sup_{\theta \in \Theta_{l,s}} \sup_{t \in R^2} \left[P_\theta \left(M_{\tau(h)}^{1/2} (\widehat{\theta}_{\tau(h)} - \theta) \leqslant t \right) - \Phi_2(t) \right] = 0,$$

where $\Phi_2(t) = \Phi(t_1)\Phi(t_2)$, Φ is the standard normal distribution function,

$$\Theta_{l,s} = \{ (\theta_1, \theta_2)' : -1 + s \leqslant \theta_1 \leqslant 1 - s, |\theta_2| \leqslant l \}.$$

Remark 1. *In order to estimate parameters θ_1 and θ_2 in (3) with prescribed mean square precision one can apply the theory of guaranteed estimation developed in the paper of Konev and Pergamenshchikov [3].*

We aim here at constructing sequential estimates which possess the following properties.

First, they are asymptotically uniformly normal for an unstable process (3).

Second, in the case of Gaussian disturbances $\{\varepsilon_n\}_{n \geqslant 1}$ and appropriate choice of normalizing factor, the estimates have standard two-dimensional normal distribution.

Construction of sequential estimates We will need the following sample Fisher information matrices

$$M_{1,n} = \sum_{k=1}^{n} Y_{k-1}Y_{k-1}', \quad M_{2,m} = \sum_{k=n+1}^{m} Y_{k-1}Y_{k-1}', \ m > n, \quad (6)$$

calculated by the sets of observations $(x_0, \ldots, x_{n-1})$ and $(x_n, x_{n+1}, \ldots, x_{m-1})$ respectively.

Further we introduce two stopping rules $\tau_1 = \tau_1(h)$ and $\tau_2 = \tau_2(h)$, for each $h > 0$ as

$$\tau_1(h) = \inf \left\{ n \geqslant 1 : \sum_{k=1}^{n} x_{k-1}^2 \geqslant h \right\},$$

$$\tau_2(h) = \inf \left\{ n > \tau_1(h) : \sum_{j=\tau_1(h)+1}^{n} 1 \geqslant h \right\}, \ \inf\{\varnothing\} = \infty,$$

and modify the matrices (6) as follows

$$\widehat{M}_{1,\tau_1(h)} = \sum_{k=1}^{\tau_1(h)} \sqrt{\beta_{1,k}(h)} Y_{k-1}Y_{k-1}', \quad \widehat{M}_{2,\tau_2(h)} = \sum_{j=\tau_1(h)+1}^{\tau_2(h)} \sqrt{\beta_{2,j}(h)} Y_{j-1}Y_{j-1}',$$

$$(7)$$

where
$$\beta_{1,k}(h) = \begin{cases} 1 & if \ \ k < \tau_1(h), \\ \alpha_1(h) & if \ \ k = \tau_1(h); \end{cases} \qquad \beta_{2,j}(h) = \begin{cases} 1 & if \ \ j < \tau_2(h), \\ \alpha_2(h) & if \ \ j = \tau_2(h). \end{cases}$$

Here $\alpha_1(h)$ and $\alpha_2(h)$, $0 < \alpha_i(h) \leqslant 1$, $i = 1, 2$, are the correction factors defined by the equations

$$\sum_{k=1}^{\tau_1(h)-1} x_{k-1}^2 + \alpha_1(h)x_{\tau_1(h)-1}^2 = h, \qquad \sum_{j=\tau_1(h)+1}^{\tau_2(h)-1} 1 + \alpha_2(h) = h. \qquad (8)$$

Let $v(h) = (v_1(h), v_2(h))'$ be the vector with the components

$$v_1(h) = \sum_{k=1}^{\tau_1(h)} \sqrt{\beta_{1,k}(h)}x_{k-1}x_k; \quad v_2(h) = \sum_{j=\tau_1(h)+1}^{\tau_2(h)} \sqrt{\beta_{2,j}(h)}x_j. \qquad (9)$$

By making use of matrices (7) we define a sequential version of the sample Fisher information matrix (4) as

$$m(h) = ||m_{i,j}(h)|| = \begin{bmatrix} \langle \widehat{M}_{1,\tau_1(h)} \rangle_{11} & \langle \widehat{M}_{1,\tau_1(h)} \rangle_{12} \\ \langle \widehat{M}_{2,\tau_2(h)} \rangle_{21} & \langle \widehat{M}_{2,\tau_2(h)} \rangle_{22} \end{bmatrix}, \qquad (10)$$

where $\langle A \rangle_{ij}$ denotes (i,j)-th element of a matrix A.

Finally, we construct the sequential estimate for vector $\theta = (\theta_1, \theta_2)'$ in (3) as

$$\theta^*(h) = \begin{pmatrix} \theta_1^*(h) \\ \theta_2^*(h) \end{pmatrix} = m^{-1}(h)v(h). \qquad (11)$$

Uniform asymptotic normality of $\theta^*(h)$. First we will study the asymptotic distribution of the estimate (11). We assume that the process (3) is unstable, that is $|\theta_1| \leqslant 1$ and that the distribution of disturbances ε_k is not specified. In this case we arrive at the following result.

Theorem 2. *Let sequential estimates $\theta^*(h) = (\theta_1^*(h), \theta_2^*(h))'$ for $\theta = (\theta_1, \theta_2)'$ in (3) be defined, for each $h > 0$, by (11). If the process (3) is unstable, that is $|\theta| \leqslant 1$, and $\{\varepsilon_n\}$ are i.i.d. random variables with $\mathsf{E}\varepsilon_n = 0$, $0 < \mathsf{E}\varepsilon_n^2 = \sigma^2 < \infty$ and are independent of x_0, then for any $0 < L < \infty$*

$$\lim_{h\to\infty} \sup_{\theta\in\Theta_L} \sup_{t\in R^2} \left| P_\theta \left(\frac{m(h)}{\sqrt{h}}(\theta^*(h) - \theta) \leqslant t \right) - \Phi_2 \left(\frac{t}{\sigma} \right) \right| = 0$$

where $t = (t_1, t_2)'$, $\Phi_2(t) = \Phi(t_1)\Phi(t_2)$, Φ is the standard normal distribution function,

$$\Theta_L = \{\theta = (\theta_1, \theta_2)' : \ |\theta_1| \leqslant 1, |\theta_2| \leqslant L\}.$$

Proof. Substituting x_k from (3) in (11) yields
$$m(h)\theta^*(h) = v(h) = m(h)\theta + \eta(h), \ \eta(h) = (\eta_1(h), \eta_2(h))'$$
where
$$\eta_1(h) = \sum_{k=1}^{\tau_1(h)} \sqrt{\beta_{1,k}(h)} x_{k-1}\varepsilon_k, \ \ \eta_2(h) = \sum_{j=\tau_1(h)+1}^{\tau_2(h)} \sqrt{\beta_{2,j}(h)}\varepsilon_j.$$
This implies that
$$\frac{m(h)}{\sqrt{h}}(\theta^*(h) - \theta) = \frac{1}{\sqrt{h}}\eta(h). \tag{12}$$
Therefore we have to establish that $\eta(h)/\sqrt{h}$ converges uniformly in $\theta \in \Theta_L$ to two-dimensional normal distribution, that is
$$\frac{1}{\sqrt{h}}\eta(h) \Rightarrow \mathcal{N}(0, \sigma^2 I), \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
as $h \to \infty$. To this end it suffices to show that for each constant vector $\lambda = (\lambda_1, \lambda_2)' \in R^2$ with $\|\lambda\| = 1$, $\lambda_1 \neq 0$, $\lambda_2 \neq 0$,
$$\lim_{h\to\infty} \sup_{\theta\in\Theta_L} \sup_{-\infty<z<\infty} \left| P_\theta\left(\lambda'\frac{\eta(h)}{\sqrt{h}} \leqslant z\right) - \Phi\left(\frac{z}{\sigma}\right) \right| = 0. \tag{13}$$
The linear combination
$$\zeta(h) = \lambda_1\eta_1(h) + \lambda_2\eta_2(h)$$
can be rewritten as
$$\zeta(h) = \sum_{j=1}^{\tau_2(h)} y_{j-1}\varepsilon_j + \zeta_1(h)$$
where
$$y_j = \lambda_1 x_j \chi_{(\tau_1(h)>j+1)} + \lambda_1\sqrt{\alpha_1(h)}x_{\tau_1(h)-1}\chi_{(\tau_1(h)=j+1)} + \lambda_2\chi_{(\tau_1(h)<j+1)},$$
$$\zeta_1(h) = (\sqrt{\alpha_2(h)} - 1)\varepsilon_{\tau_2(h)}.$$
Further we prove that
$$\lim_{h\to\infty} \sup_{\theta\in\Theta_L} \sup_{-\infty<z<\infty} \left| P_\theta\left(\lambda'\frac{1}{\sqrt{h}}\sum_{j=1}^{\tau_2(h)} y_{j-1}\varepsilon_j \leqslant z\right) - \Phi\left(\frac{z}{\sigma}\right) \right| = 0. \tag{14}$$
and that for any $\Delta > 0$
$$\lim_{h\to\infty} \sup_{\theta\in\Theta_L} P_\theta\left(\frac{1}{\sqrt{h}}|\zeta_1(h) > \Delta|\right) = 0. \tag{15}$$
The argument in the proof of (14) is similar to that of Proposition 2.1 of the paper [4]. Combining (14) and (15) one comes to (13). Hence Theorem 2. $\qquad\square$

Non-asymptotic normality of $\theta^*(h)$. Now we assume that the noises $\{\varepsilon_k\}$ in (3) is a sequence of i.i.d. random variables with Gaussian distributions. In this case we can derive non-asymptotic distribution of

the sequential estimate (11) under appropriate choice of the normalizing factor.

Theorem 3. *Let sequential estimates* $\theta^*(h) = (\theta_1^*(h), \theta_2^*(h))'$ *for* $\theta = (\theta_1, \theta_2)'$ *in (3) be defined by (11). If* $\{\varepsilon_n\}$ *are i.i.d. standard normal random variables, that is* $\varepsilon_n \sim \mathcal{N}(0,1)$, *then for any* $\theta \in R^2$ *and any* $h > 0$

$$P_\theta \left(\frac{m(h)}{\sqrt{h}} (\theta^*(h) - \theta) \leqslant t \right) = \Phi_2(t)$$

where $m(h)$ *is defined by (10).*

Proof. Taking into account (12) we establish that the characteristic function of the vector $\eta(h)/\sqrt{h}$ has the form

$$\varphi_{\frac{\eta(h)}{\sqrt{h}}}(u) = \mathsf{E}exp(i(\eta, u)) = exp\left(-\frac{u_1^2}{2} - \frac{u_2^2}{2} \right).$$

This completes the proof of Theorem 3. $\square$

Monte-Carlo simulation results.   In this section we report some numerical results to verify the asymptotic uniform normality property of sequential estimate proved in Theorem 2. The basic experiment consisted of 20000 replications of the sequential procedure (11) with $h = 20$ for each value of the parameter vector $\theta = (\theta_1, \theta_2)'$ indicated in Table 1.

After each simulation of the procedure the normalized deviation vector

$$z(h) = (z_1(h), z_2(h))' = \frac{m(h)}{\sqrt{h}} (\theta^*(h) - \theta)$$

was calculated. Table 1 gives frequency estimates for the probability
$$P(z_1(h) \leqslant a, \ z_2(h) \leqslant b)$$
for two points (a, b) and different values of unknown parameters θ_1, θ_2.

The results of simulations show good performance of the procedure (11) and confirm the results of Theorem 2.

Concluding remarks.   In this paper we propose new sequential estimates for estimating parameters in autoregressive process $AR(1)$ with unknown mean.

We prove the property of uniform asymptotic normality when the process is unstable and the noise distributions are not specified.

Table 1 — Test on the uniform normality

$a = -0.3,\ b = 1.7,\ h = 20,\ p = 0.365$						
$\theta_1\|\theta_2$	-2.1	-1.4	-0.7	0	0.7	1.4	2.1
-1	0.367	0.364	0.365	0.365	0.367	0.369	0.365
-0.5	0.370	0.355	0.365	0.363	0.369	0.365	0.357
0	0.366	0.364	0.368	0.361	0.360	0.363	0.365
0.5	0.365	0.368	0.368	0.367	0.362	0.361	0.366
1	0.369	0.367	0.364	0.369	0.363	0.364	0.368
$a = 0.2,\ b = -0.5,\ h = 20,\ p = 0.179$						
$\theta_1\|\theta_2$	-2.1	-1.4	-0.7	0	0.7	1.4	2.1
-1	0.175	0.179	0.182	0.180	0.181	0.180	0.182
-0.5	0.181	0.182	0.181	0.181	0.180	0.183	0.183
0	0.181	0.177	0.178	0.174	0.176	0.177	0.182
0.5	0.179	0.183	0.182	0.182	0.178	0.175	0.178
1	0.175	0.176	0.179	0.182	0.181	0.177	0.180

References

1. Borisov V. Z. On sequential parameters estimation in discrete time processes /V. Z. Borisov, V. V. Konev // Automat. Remote Control. – 1977. – No 6. – pp. 58–64.

2. Galtchouk L. On Uniform Asymptotic Normality of Sequential Estimators for the Parameters in a Stable AR(1) / L. Galtchouk, V. V. Konev // Sequential Analysis. – 2003. – vol. 22. – Nos 1 & 2. – pp.31-54.

3. Konev V. V. Sequential Parameter Estimation with Guaranteed Mean-Square Accuracy for Unstable Linear Stochastic Systems / V. V. Konev, S. M. Pergamenshchikov // Problemy Peredachi Informatsii. – 1992. – Vol. 28. – No. 4. – pp. 35–48.

4. Lai T. L. Fixed-accuracy estimation of an autoregressive parameter / T. L. Lai and D. Siegmund // Ann. Statist. – 1983. – Vol. 11. – pp. 478–485.

5. Shiryaev A. N. Statistical experiments and decisions: Asymptotic theory /Shiryaev A. N. and Spokoiny V. G. – World Scientific. – 2000