One investigation method of ratio type estimators*

Vyacheslav A. Vasiliev

Department of Applied Mathematics
and Cybernetics
Tomsk State University
vas@mail.tsu.ru

Abstract

This paper presents a truncated modification of basic ratio type estimators constructed by dependent sample of finite size.

This method gives a possibility to obtain estimators with guaranteed accuracy in the sense of L_m-norm, $m \geq 2$. As an illustration, parametric and non-parametric estimation problems on a time interval of a fixed length are considered. In particular, parameters of linear (autoregressive) and non-linear (ARARCH) discrete-time processes are estimated. Moreover, the parameter estimation problem of non-Gaussian Ornstein-Uhlenbeck process by discrete-time observations and the estimation problem of a logarithmic derivative of a noise density of an autoregressive process with guaranteed accuracy are solved.

In addition to non-asymptotic properties, the limiting behavior of presented estimators is investigated. It is shown, in particular, that all parametric truncated estimators have rates of convergence of basic estimators. Non-parametric estimator has optimal (as compared to the case of independent inputs) rate of convergence.

AMS classification: 62G05; 62G07; 62F99; 93B30; 93E10

Ratio estimation; truncated estimation method; dependent observations; guaranteed accuracy; finite sample size; autoregression; ARARCH model; non-Gaussian Ornstein-Uhlenbeck process; non-parametric logarithmic density derivative estimation

* The research on this paper was supported by RFBR 11-01-12053 Grant

1 Introduction

Modern evolution of mathematical statistics is turned to development of data processing methods by dependent sample of finite size.

One of such possibilities gives a well-known sequential estimation method, which was successfully applied in parametric and non-parametric problems.

This approach for various statistical problems for a scheme of independent observations has been primarily proposed by [Wald(1947)]. Then this idea has been applied to parameter estimation problem of continuous and discrete-time dynamic systems in many papers and books (see [Dobrovidov et al.(2012), Konev(1985), Küchler and Vasiliev(2010), Liptser and Shiryaev(1977), Vasiliev et al.(2004)] among others).

Sequential approach has been also applied to non-parametric regression, autoregression and density function estimation problems as well (see, e.g., [Arkoun(2011), Arkoun and Pergamenchtchikov(2008), Dobrovidov et al.(2012), Efroimovich(2007), Vasiliev et al.(2004)]).

To obtain sequential estimators with an arbitrary accuracy one needs to have a sample of unbounded size. However in practice the observation time of a system is usually not only finite but fixed. One of the possibilities for finding estimators with the guaranteed quality of inference using a sample of fixed size is provided by the approach of truncated sequential estimation. The truncated sequential estimation method was developed by [Konev and Pergamenshchikov(1990a), Konev and Pergamenshchikov(1990b), Fourdrinier et al.(2009)] (and others) for parameter estimation problems in discrete–time dynamic models. Using a sequential approach, estimators of dynamic systems parameters with known variance by sample of fixed size were constructed in these papers.

Non-parametric truncated sequential estimators of a regression function were presented by [Politis and Vasiliev(2012a), Politis and Vasiliev(2012b)] on the basis of Nadaraya–Watson estimators calculated at a special stopping time. These estimators have known mean square errors as well. The duration of observations is also random but bounded from above by a non-random fixed number.

The main purpose of this paper is to obtain a modification of ratio type estimators from a wide class, having guaranteed accuracy by dependent sample of finite size.

When estimating, for example, the ratio type functionals one uses as a rule the substitution statistics of [Borovkov(1997)], that is ratio of some estimators. Studying the properties of such estimators, we face certain difficulties that are connected with finding dominating sequences [Cramér(1948)]. In some cases, for instance, in reconstruction of the logarithmic derivative of a distribution density one can use estimators for which an exact asymptotic expression of the mean square error (MSE) is available, [Dobrovidov et al.(2012), Vasiliev et al.(2004)].

For this problem the theory of smoothing can be also used.It gives a possibility to find the principal term of the MSE of the ratio estimators with an improved rate of convergence, similar to the case of independent observations. Moreover, the rate of convergence of the estimators of their ratio in metric $L_m,\ m \geq 2$, can be obtained, see [Dobrovidov et al.(2012), Vasiliev et al.(2004)].

In this paper the truncated estimation method for ratio type functionals constructed by dependent sample of finite size is presented. This method gives a possibility to obtain estimators with guaranteed accuracy in the sense of L_m-norm, $m \geq 2$. Examples of parametric and non-parametric estimation problems on a time interval of a fixed length are considered.

2 Statement of the problem. Main result

Let $(\Omega,\ \mathcal{F},\ \mathsf{P})$ be a probability space with a filtration $\{\mathcal{F}_n\}_{n\geq 0}$ and let $(f_n)_{n\geq 1}$ and $(g_n)_{n\geq 1}$ be $\{\mathcal{F}_n\}$-adapted sequences of random numbers.

Let

$$\Psi_N = \frac{f_N}{g_N}, \quad N \geq 1 \tag{1}$$

be an estimator of a number Ψ. For example,

$$\Psi = \frac{f}{g}$$

if f_N and g_N are estimators of some numbers f and $g \neq 0$.

Consider the following modification of the estimator Ψ_N :

$$\tilde{\Psi}_N = \Psi_N \cdot \chi(|g_N| \geq H), \quad N \geq 1, \tag{2}$$

where H is a positive number, defined below and $\chi(a \geq b) = 1$ for $a \geq b$ and 0 when $a < b$.

Our main aim is to formulate general conditions on the sequences (f_N) and (g_N) and on the number H giving a possibility to estimate Ψ with a guaranteed accuracy in the sense of the L_m-norm, $m \geq 2$.

These conditions contained to the following

Theorem 2.1 *Let $\Psi_N = f_N/g_N,\ N \geq 1$ be an estimator of a number Ψ. Assume for some integer $m \geq 1$ and $\mu \geq 1$ there exist sequences of positive numbers $(\varphi_N(m))_{N\geq 1}$ and $(w_N(\mu))_{N\geq 1}$, decreasing to zero, as well as a number $g \neq 0$ such that*

(i) $E(f_N - \Psi g_N)^{2m} \leq \varphi_N(m)$;

(ii) $E(g_N - g)^{2\mu} \leq w_N(\mu)$.

Then for every $H \in (0, |g|)$ and $N \geq 1$ the estimator $\tilde{\Psi}_N$ has the property

$$E(\tilde{\Psi}_N - \Psi)^{2m} \leq V_N(m, \mu), \tag{3}$$

where

$$V_N(m,\mu) = \frac{1}{H^{2m}}\varphi_N(m) + \frac{\Psi^{2m}}{(g-H)^{2\mu}} w_N(\mu).$$

P r o o f. From the definition of the estimator $\tilde{\Psi}_N$ we find its deviation

$$\tilde{\Psi}_N - \Psi = \frac{f_N - \Psi g_N}{g_N} \cdot \chi(|g_N| \geq H) - \Psi \cdot \chi(|g_N| < H).$$

Then, using the Chebyshev inequality and the definition of V_N we can estimate the desired moment

$$E(\tilde{\Psi}_N - \Psi)^{2m} = E\frac{(f_N - \Psi g_N)^{2m}}{g_N^{2m}} \cdot \chi(|g_N| \geq H) +$$

$$+\Psi^{2m} \cdot P(|g_N| < H) \leq \frac{1}{H^{2m}} E(f_N - \Psi g_N)^{2m} +$$

$$+\Psi^{2m} \cdot P(|g_N - g| > |g| - H) \leq$$

$$\leq \frac{1}{H^{2m}} \varphi_N(m) + \Psi^{2m} \cdot \frac{E(g_N - g)^{2\mu}}{(|g| - H)^{2\mu}} \leq V_N(m, \mu).$$

Theorem 2.1 is proved.

Corollary 2.1 *Assume that $\Psi = f/g$ for some number f, where g is defined in Theorem 2.1 and, instead of the assumption (i), for some $\nu \geq 1$ there exists sequence $(\upsilon_N(\nu))_{N\geq 1}$ of non-negative numbers, decreasing to zero, such that*

$$E(f_N - f)^{2\nu} \leq \upsilon_N(\nu), \quad N \geq 1.$$

Then the condition (i) of Theorem 2.1 is fulfilled, where the function $\varphi_N(m)$ should be substituted by the following one:

$$\varphi_N(m) = \frac{2^{2m-1}}{g^{2m}} [g^{2m} \upsilon_N^{m/\nu}(m) + f^{2\nu} w_N^{m/\mu}(\nu)], \quad m = \nu \wedge \mu.$$

Corollary 2.2 *If it is known, that $\Psi \in [A, B]$, then the estimator (2) can be taken in the form*

$$\tilde{\Psi}_N^* = \Psi_N \cdot \chi(|g_N| \geq H) + L\chi(|g_N| < H), \quad N \geq 1, \tag{4}$$

where $L = (A + B)/2$. In this case the number Ψ in the upper bound $V(m, \mu)$ in (3) should be substituted by the number $\Psi^* = (B - A)/2$.

Remark 1 *The function $V_N(m, \mu)$ could be unknown. At the same time the knowledge of the rate of L_m-convergence of proposed estimators can be useful in various adaptive procedures (control, prediction etc) for the construction of pilot estimators (see, e.g., [Dobrovidov et al.(2012), Vasiliev et al.(2004)]–[Vasiliev(1997)]).*

3 Examples

3.1 Estimation of parameters of a stable first order autoregression

Consider the process satisfying the following equation

$$x_n = \lambda x_{n-1} + \xi_n, \quad n \geq 1, \tag{5}$$

where noises $\xi_n,\ n \geq 1$ are i.i.d. zero mean random variables with finite (for some even number $\gamma \geq 2$) moments $\sigma^{2\gamma} = E\xi_n^{2\gamma}$, as well as $Ex_0^{2\gamma} < \infty$ and $|\lambda| < 1$.

It should be noted, that under these conditions there exist functions $\sigma_x^{2\gamma}(\theta)$, $\theta = (\lambda, \sigma^2, \sigma^{2\gamma})$, such that

$$\sup_n E_\theta x_n^{2\gamma} \le \sigma_x^{2\gamma}(\theta) < \infty. \tag{6}$$

Consider the estimation problem of λ and σ^2 with a guaranteed accuracy.

a) Non-asymptotic estimation of λ

We define the estimator of the type (2) on the basis of the least squares estimator (LSE) of the type (1)

$$\hat\lambda_N = \frac{\frac{1}{N}\sum\limits_{n=1}^{N} x_n x_{n-1}}{\frac{1}{N}\sum\limits_{n=1}^{N} x_{n-1}^2}, \quad N \ge 1.$$

According to general notation, in this case we have

$$\Psi = \lambda, \quad \Psi_N = \hat\lambda_N,$$

$$f_N = \frac{1}{N}\sum_{n=1}^{N} x_n x_{n-1}, \quad g_N = \frac{1}{N}\sum_{n=1}^{N} x_{n-1}^2$$

and $\tilde\Psi_N = \tilde\lambda_N$,

$$\tilde\lambda_N = \hat\lambda_N \cdot \chi(g_N \ge H). \tag{7}$$

Using formula (5) it is easy to verify that

$$g = \frac{\sigma^2}{1-\lambda^2},$$

as well as for $m = \gamma/2$, $\mu = m$, constants $C_1(m,\theta)$ and $C_2(m,\theta)$ are exist such that

$$w_N(m,\theta) = \frac{C_1(m,\theta)}{N^m} + \frac{C_2(m,\theta)}{N^{2m}}, \tag{8}$$

where, e.g., for $m = 1$,

$$\begin{array}{l} C_1(1,\theta) = \frac{1}{(1-\lambda^2)^2}\left[12\lambda^2\sigma^2\left(\lambda^2 E x_0^2 + \frac{\sigma^2}{1-\lambda^2}\right) + 3E\xi_1^4\right], \\ C_2(1,\theta) = \frac{1}{(1-\lambda^2)^2}\left[24\left(\lambda^4 E x_0^4 + \frac{4(\sigma^2)^2}{(1-\lambda^2)^2}\right) + E x_0^4\right]. \end{array}$$

Moreover, using the Burkholder and Hölder inequalities, we have

$$E_\theta(f_N - \lambda g_N)^{2m} = \frac{1}{N^{2m}} E_\theta\Big(\sum_{n=1}^{N} x_{n-1}\xi_n\Big)^{2m} \le$$

$$\le \frac{B_{2m}^{2m}\sigma^{2m}}{N^{2m}} E_\theta\Big(\sum_{n=1}^{N} x_{n-1}^2\Big)^m \le \frac{B_{2m}^{2m}\sigma^{2m}}{N^{m+1}} \sum_{n=1}^{N} E_\theta x_{n-1}^{2m} \le$$

$$\le B_{2m}^{2m}\sigma^{2m}\sigma_x^{2m}(\theta)\frac{1}{N^m} =: \varphi_N(m,\theta),$$

where B_{2m} is the coefficient from the Burkholder inequality. In particular,

$$\varphi_N(1,\theta) = \frac{(\sigma^2)^2}{1-\lambda^2}\frac{1}{N}.$$

Thus all conditions of Theorem 2.1 hold, hence for obviously defined numbers $\tilde{C}_1(m,\theta)$, $\tilde{C}_2(m,\theta)$ and known $0 < H < g$ (e.g., for $0 < H < \sigma^2$),

$$E_\theta(\tilde{\lambda}_N - \lambda)^{2m} \leq \frac{\tilde{C}_1(m,\theta)}{N^m} + \frac{\tilde{C}_2(m,\theta)}{N^{2m}}, \quad N \geq 1. \tag{9}$$

For the parameter estimation with a guaranteed accuracy we have to know that, e.g., $\theta \in \Theta$, where $\Theta = \{\theta = (\lambda, \sigma^2, \sigma^{2\gamma}) : \ |\lambda| \leq r < 1, \ 0 < \underline{\sigma}^2 \leq \sigma^2, \ \sigma^{2\gamma} \leq \overline{\sigma}^{2\gamma}\}$.

In this case we can find known functions

$$\overline{\varphi}_N(m) = \sup_{\theta\in\Theta} \varphi_N(m,\theta) \quad \text{and} \quad \overline{w}_N(m) = \sup_{\theta\in\Theta} w_N(m,\theta)$$

such that

$$\sup_\Theta E_\theta(f_N - \lambda g_N)^{2m} \leq \overline{\varphi}_N(m),$$
$$\sup_\Theta E_\theta(g_N - g)^{2m} \leq \overline{w}_N(m).$$

In particular, for $m = 1$,

$$\overline{\varphi}_N(1) = \frac{(\overline{\sigma}^2)^2}{1-r^2}\frac{1}{N},$$

as well as in (8) we can replace $C_1(1,\theta)$ and $C_1(2,\theta)$ with

$$\overline{C}_1(1) = \frac{1}{(1-r^2)^2}\left[12r^2\overline{\sigma^2}\left(r^2 Ex_0^2 + \frac{\overline{\sigma^2}}{1-r^2}\right) + 3\overline{\sigma}^4\right],$$

$$\overline{C}_2(1) = \frac{1}{(1-r^2)^2}\left[24\left(r^4 Ex_0^4 + \frac{4(\overline{\sigma^2})^2}{(1-r^2)^2}\right) + Ex_0^4\right]$$

and obtain

$$\overline{w}_N(1) = \frac{\overline{C}_1(1)}{N} + \frac{\overline{C}_2(1)}{N^2}.$$

In general, for $0 < H < \underline{\sigma}^2$ we can find the numbers

$$\tilde{C}_1(m) = \sup_\Theta \tilde{C}_1(m,\theta) < \infty, \quad \tilde{C}_2(m) = \sup_\Theta \tilde{C}_2(m,\theta) < \infty$$

and then, according to (9) we have

$$\sup_\Theta E_\theta(\tilde{\lambda}_N - \lambda)^{2m} \leq \frac{\tilde{C}_1(m)}{N^m} + \frac{\tilde{C}_2(m)}{N^{2m}}, \quad N \geq 1. \tag{10}$$

In particular, for $\gamma = 2$ and $m = 1$,

$$\sup_\Theta E_\theta(\tilde{\lambda}_N - \lambda)^2 \leq$$

$$\leq \left[\frac{(\overline{\sigma}^2)^2}{(1-r^2)H^2} + \frac{r^2\overline{C}_1(1)}{(\underline{\sigma}^2 - H)^2}\right]\frac{1}{N} + \frac{r^2\overline{C}_2(1)}{(\underline{\sigma}^2 - H)^2}\frac{1}{N^2}.$$

Remark 2 *It is well known, that for the stable process (5) the function $g_N \to g$ as $N \to \infty$ almost surely. Then*

$$P_\theta(\tilde{\lambda}_N = \hat{\lambda}_N) = 1, \quad \theta \in \Theta,$$

for N large enough. Moreover, it is easy to establish the uniform on Θ asymptotic normality of estimator $\tilde{\lambda}_N$ with the smallest asymptotic variance. It gives a possibility to prove the minimax optimality of $\tilde{\lambda}_N$.

b) Non-asymptotic estimation of σ^2

Consider the estimation problem of the noise variance σ^2 in the model (5) under the assumption $\gamma = 4$ ($\sigma^8 < \infty, \ Ex_0^8 < \infty$).

In the definition of the LSE type estimator $\hat{\sigma}_N^2$ defined as

$$\hat{\sigma}_N^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \tilde{\lambda}_N x_{n-1})^2, \quad N \geq 1,$$

we use the estimator $\tilde{\lambda}_N$ of λ, defined in (7), having known non-asymptotic properties (10) for $m = 1$ and $m = 2$.

Thus, using (6) and (10), we have

$$E_\theta(\hat{\sigma}_N^2 - \sigma^2)^2 = E_\theta(\tilde{\lambda}_N - \lambda)^2 \frac{1}{N}\sum_{n=1}^{N} x_{n-1}^2 +$$

$$+2E_\theta(\tilde{\lambda}_N - \lambda)\frac{1}{N}\sum_{n=1}^{N} x_{n-1}\xi_n \leq \sqrt{\sigma_x^4(\theta)E_\theta(\tilde{\lambda}_N - \lambda)^4} +$$

$$+2\sqrt{\sigma^2\sigma_x^2(\theta)E_\theta(\tilde{\lambda}_N - \lambda)^2\frac{1}{N}} \leq \frac{C_N(\theta)}{N},$$

where

$$C_N(\theta) = \left\{\sqrt{\sigma_x^4(\theta)[\tilde{C}_1(2) + \tilde{C}_2(2)\frac{1}{N^2}]} + \right.$$

$$\left. +2\sqrt{\sigma^2\sigma_x^2(\theta)[\tilde{C}_1(1) + \tilde{C}_2(1)\frac{1}{N}]}\right\}.$$

Since the numbers

$$\overline{C}_N = \sup_\Theta C_N(\theta) < \infty, \quad N \geq 1$$

are known, we have obtained estimator of σ^2 with a guaranteed accuracy:

$$\sup_\Theta E_\theta(\hat{\sigma}_N^2 - \sigma^2)^2 \leq \frac{\overline{C}_N}{N}, \quad N \geq 1. \tag{11}$$

It should be noted, that this estimator is asymptotically equivalent to the corresponding LSE. In particular, it has optimal rate of convergence as $N \to \infty$.

3.2 Estimation of parameters of a stable ARARCH(1,1)

Consider the process satisfying the following equation

$$x_n = \lambda x_{n-1} + \sqrt{\sigma_0^2 + \sigma_1^2 x_{n-1}^2} \cdot \xi_n, \quad n \geq 1, \tag{12}$$

where noises $\xi_n,\ n \geq 1$ are i.i.d. zero mean random variables with variance equal to one and finite fourth moment $\sigma^4 = E\xi_1^4$, as well as $Ex_0^4 < \infty$ and $|\lambda| < 1$.

Define the LSE $\hat{\lambda}_N$ of λ of the form:

$$\hat{\lambda}_N = \frac{\frac{1}{N}\sum\limits_{n=1}^{N} x_n x_{n-1}}{\frac{1}{N}\sum\limits_{n=1}^{N} x_{n-1}^2}, \quad N \geq 1,$$

which is strongly consistent (see, e.g., [Malyarenko(2010)]) under the following stability condition

$$\lambda^4 + 6\lambda^2\sigma_1^2 + (\sigma_1^2)^2\sigma^4 < 1. \tag{13}$$

a) Non-asymptotic estimation of λ

According to general notation, in this case we have

$$\Psi = \lambda, \quad \Psi_N = \hat{\lambda}_N,$$

$$f_N = \frac{1}{N}\sum_{n=1}^{N} x_n x_{n-1}, \quad g_N = \frac{1}{N}\sum_{n=1}^{N} x_{n-1}^2$$

and $\tilde{\Psi}_N = \tilde{\lambda}_N$,

$$\tilde{\lambda}_N = \hat{\lambda}_N \cdot \chi(g_N \geq H). \tag{14}$$

Define for some known numbers $r \in (0,1)$, $\underline{\sigma}_0^2$, $\overline{\sigma}_0^2$, $\underline{\sigma}_1^2$, and $\overline{\sigma}_1^2$ the set

$$\Theta = \{\theta = (\lambda, \sigma_0^2, \sigma_1^2) : \ \lambda^4 + 6\lambda^2\sigma_1^2 + (\sigma_1^2)^2\sigma^4 \leq r,$$

$$\underline{\sigma}_0^2 \leq \sigma_0^2 \leq \overline{\sigma}_0^2, \quad \underline{\sigma}_1^2 \leq \sigma_1^2 \leq \overline{\sigma}_1^2\}.$$

Using the following representation for the process (x_n^2) :

$$x_n^2 = (\lambda^2 + \sigma_1^2)x_{n-1}^2 + \sigma_0^2 + \varsigma_n, \quad n \geq 1,$$

where

$$\varsigma_n = 2\lambda x_{n-1}\sqrt{\sigma_0^2 + \sigma_1^2 x_{n-1}^2} \cdot \xi_n + \sigma_0^2(\xi_n^2 - 1) + \sigma_1^2 x_{n-1}^2(\xi_n^2 - 1),$$

it is easy to find the numbers

$$\sigma_x^2 = \sup_{\Theta,n} E_\theta x_n^2, \quad \sigma_x^4 = \sup_{\Theta,n} E_\theta x_n^4$$

and $\lim\limits_{N\to\infty} g_N = g$ P_θ-a.s. for $\theta \in \Theta$, where

$$g = \frac{\sigma_0^2}{1 - \lambda^2 - \sigma_1^2}.$$

For this model we can calculate

$$\overline{w}_N = \{12(\overline{\sigma}_0^2\sigma_x^2 + \overline{\sigma}_1^2\sigma_x^4) + \sigma_\xi^4 \cdot ((\overline{\sigma}_0^2)^2 + 2\overline{\sigma}_0^2\overline{\sigma}_1^2\sigma_x^2 +$$

$$+(\overline{\sigma}_1^2)^2\sigma_x^4)\}\frac{1}{N} + \{Ex_0^4 + 3\sigma_x^4\}\frac{1}{N^2}, \tag{15}$$

where σ_ξ^4 is an upper bound for $E(\xi_1^2 - 1)^2$.

Now we can find the function $\overline{\varphi}_N$. By the definition of f_N and g_N we have

$$\sup_\Theta E_\theta(f_N - \lambda g_N)^2 =$$

$$= \frac{1}{N^2}\sup_\Theta E_\theta(\sum_{n=1}^N x_{n-1}\sqrt{\sigma_0^2 + \sigma_1^2 x_{n-1}^2} \cdot \xi_n)^2 =$$

$$= \frac{1}{N^2}\sum_{n=1}^N \sup_\Theta E_\theta x_{n-1}^2(\sigma_0^2 + \sigma_1^2 x_{n-1}^2) \leq (\overline{\sigma}_0^2\sigma_x^2 + \overline{\sigma}_1^2\sigma_x^4)\frac{1}{N}$$

and we can put

$$\overline{\varphi}_N = (\overline{\sigma}_0^2\sigma_x^2 + \overline{\sigma}_1^2\sigma_x^4)\frac{1}{N}.$$

Then for $0 < H < \dfrac{\underline{\sigma}_0^2}{1 - \underline{\sigma}_1^2}$ and every $N \geq 1$,

$$\sup_\Theta E_\theta(\tilde{\lambda}_N - \lambda)^2 \leq \frac{1}{H^2}\overline{\varphi}_N + \frac{(1 - \underline{\sigma}_1^2)^2}{(\underline{\sigma}_0^2 - (1 - \underline{\sigma}_1^2)H)^2}\overline{w}_N. \tag{16}$$

It should be noted, that the rate of convergence of the obtained upper bound is the same that the rate of the LSE and is optimal.

b) Non-asymptotic estimation of σ_0^2 and σ_1^2

We will construct estimators with guaranteed accuracy on the bases of correlation estimators for the following cases:

(i) of σ_0^2 with known σ_1^2 :

$$\hat{\sigma}_0^2(N) = \frac{1}{N}\sum_{n=1}^N [x_n^2 - (\hat{\lambda}_N^2 + \sigma_1^2)x_{n-1}^2];$$

(ii) of σ_1^2 with known σ_0^2 :

$$\hat{\sigma}_1^2(N) = \frac{\sum\limits_{n=1}^N (x_n^2 - \sigma_0^2)}{\sum\limits_{n=1}^N x_{n-1}^2} - \hat{\lambda}^2(N),$$

which are strongly consistent under the condition (13), see [Malyarenko(2010)].

Define estimators for considered cases

$$(i) \qquad \tilde{\sigma}_0^2(N) = \frac{1}{N}\sum_{n=1}^N [x_n^2 - ((\lambda_N^*)^2 + \sigma_1^2)x_{n-1}^2];$$

$$(ii) \quad \tilde{\sigma}_1^2(N) = \frac{\frac{1}{N}\sum\limits_{n=1}^{N}(x_n^2 - \sigma_0^2)}{\frac{1}{N}\sum\limits_{n=1}^{N}x_{n-1}^2}\chi(g_N \geq H) - (\lambda_N^*)^2,$$

where

$$\lambda_N^* = \mathrm{proj}_{[-1,1]}\tilde{\lambda}_N,$$

$\tilde{\lambda}_N$ and g_N are defined in (14).

Similar to previous sections, the upper bounds for the MSE's of these estimators with known constants C_0 and C_1 can be found

$$(i) \quad \sup_{\Theta_0} E_\theta(\tilde{\sigma}_0^2(N) - \sigma_0^2)^2 \leq \frac{C_0}{N}, \tag{17}$$

where $\Theta_0 = \{\theta = (\lambda, \sigma_0^2) : \ \lambda^4 + 6\lambda^2\sigma_1^2 + (\sigma_1^2)^2\sigma^4 \leq r, \ \underline{\sigma}_0^2 \leq \sigma_0^2 \leq \overline{\sigma}_0^2\}$ and

$$(ii) \quad \sup_{\Theta_1} E_\theta(\tilde{\sigma}_1^2(N) - \sigma_1^2)^2 \leq \frac{C_1}{N}, \tag{18}$$

where $\Theta_1 = \{\theta = (\lambda, \sigma_1^2) : \ \lambda^4 + 6\lambda^2\sigma_1^2 + (\sigma_1^2)^2\sigma^4 \leq r, \ \underline{\sigma}_1^2 \leq \sigma_1^2 \leq \overline{\sigma}_1^2\}, \quad r \in (0,1).$

3.3 Parameter estimation of a stable non-Gaussian Ornstein–Uhlenbeck process by discrete–time observations

Presented below results give a possibility to make statistical inferences for continuous-time stochastic systems by finite size of observations. Moreover, one of the main assumption is a discrete scheme of observations. It coincides to numerous real situations, in particular, in problems of financial mathematics.

Consider the following regression model

$$dx(t) = ax(t)dt + d\xi(t), \ 0 \leq t \leq T \tag{19}$$

with an initial condition $x(0) = x_0$, having all moments. Here $\xi(t) = \rho_1 W(t) + \rho_2 Z(t)$, $\rho_1 \neq 0$ and ρ_2 – some constants, $(W(t), t \geq 0)$ is a standard Wiener process, given on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$, adaptive to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$, $Z(t) = \sum\limits_{k=1}^{N_t} Y_k$, where $Y_k, \ k \geq 1$ are i.i.d.r.v's having all moments and (N_t) is a Poisson process with the intensity $\lambda > 0$.

It should be noted, that for $\rho_2 = 0$ the process (19) is an Ornstein-Uhlenbeck process.

We suppose, that the unknown parameter $a \in [-\Delta, -\delta]$, where δ and Δ are known positive numbers.

The problem is to estimate the parameter a by observations of the discrete–time process $y = (y_k)$,

$$y_k = x(t_k), \quad t_k = \frac{k}{n}T, \quad k = \overline{o,n}.$$

Using the following representation for the solution of the equation (19)

$$x(t) = \mathrm{e}^{at}x_0 + \int\limits_0^t \mathrm{e}^{a(t-s)}d\xi(s), \quad 0 \leq t \leq T\,,$$

we get the recurrent equation for the observations (y_k) :

$$y_k = by_{k-1} + \eta_k, \quad k = \overline{1,n}, \tag{20}$$

where $b = e^{aT/n}$, $\eta_k = \int\limits_{t_{k-1}}^{t_k} e^{a(t_k-s)} d\xi(s)$ – i.i.d.r.v's with

$$E_a\eta_k = 0, \quad \sigma^2 := D_a\eta_k = \frac{1}{2a}(\rho_1^2 + \lambda\rho_2^2)[b^2-1].$$

Moreover, for this model all moments $\sigma^{2m} = E_a\eta_k^{2m}$ and their upper bounds $\overline{\sigma}^{2m} = \sup\limits_{a\le-\delta} \sigma^{2m}$, $m \ge 1$ are exist.

Define the estimator $\tilde{a}_n$ of a with a guaranteed accuracy using an estimator $\tilde{b}_n$ of b as follows

$$\tilde{a}_n = \frac{n}{T}\ln\tilde{b}_n, \quad n \ge 1, \tag{21}$$

where the estimator $\tilde{b}_n$ we construct on the basis of the LSE $\hat{b}_n$, obtained using the equation (20) and Corollary 2.2:

$$\tilde{b}_n = \hat{b}_n \cdot \chi(g_n \ge H) + L\chi(g_n < H), \quad \hat{b}_n = \frac{f_n}{g_n}.$$

Here $L = [e^{-\delta T/n} + e^{-\Delta T/n}]/2$,

$$f_n = \frac{1}{n}\sum_{k=1}^n y_k y_{k-1}, \quad g_n = \frac{1}{n}\sum_{k=1}^n y_{k-1}^2$$

and the number g is defined as

$$g = \frac{\sigma^2}{1-b^2}.$$

Then the estimator $\tilde{b}_n$ has all properties of the estimator $\tilde{\lambda}_N$, defined in (7). In particular, according to Theorem 1, which holds for this model for all $m \ge 1$ and $\mu \ge 1$, the following inequalities

$$\sup_{a\le-\delta} E_a(\tilde{b}_n - b)^{2m} \le \frac{C_1^*(m)}{n^m} + \frac{C_2^*(\mu)}{n^\mu}, \quad n \ge 1 \tag{22}$$

for an arbitrary $\mu > m$ and $0 < H \le \underline{\sigma}^2$ hold, where

$$\underline{\sigma}^2 = \frac{1}{2\delta}(\varrho_1^2 + \lambda\varrho_2^2)[1-r^2], \quad r = e^{-\delta}$$

and numbers $C_1^*(m)$, $C_2^*(\mu)$ are known.

From (21) and (22) it is easy to verify the following property of estimators $\tilde{a}_n$ for every $\mu > m \ge 1$:

$$\sup_{a\in[-\Delta,-\delta]} E_a(\tilde{a}_n - a)^{2m} \le$$

$$\le (nT^{-1}e^{\Delta T/n})^{2m}\left\{\frac{C_1^*(m)}{n^m} + \frac{C_2^*(\mu)}{n^\mu}\right\}, \quad n \ge 1. \tag{23}$$

3.4 Non-parametric estimation of a logarithmic density derivative

Consider the problem of estimating the logarithmic derivative

$$\Psi(t) = f'(t)/f(t)$$

of a distribution density $f(t)$ of the i.i.d. noises ξ_n in the model

$$x_n = \lambda x_{n-1} + \xi_n, \quad n \geq 1. \tag{24}$$

It is assumed, that $(\xi_n)_{n\geq 1}$ is a sequence of zero mean random numbers with finite (for some even number $\kappa \geq 1$) moments $\sigma^{2(2\kappa+1)} = E\xi_n^{2(2\kappa+1)}$, as well as $Ex_0^{2(2\kappa+1)} < \infty$ and $|\lambda| < 1$.

We will suppose that the function $f(t) \in \Phi$, i.e. satisfy the following conditions

$$0 < c_f \leq f(t), \quad \sup_{s\in\mathcal{R}^1} f(s) \leq C_f$$

and for some $\mathcal{L} > 0$ and $\gamma \in (0,1)$

$$|f^{(1+2\kappa)}(x) - f^{(1+2\kappa)}(y)| \leq \mathcal{L}|x-y|^\gamma.$$

The knowledge of $\Psi(t)$ is important in various statistical problems. In particular, it is needed when: constructing the algorithm of optimal control of an autoregressive process; estimating of a regression curve; testing close hypotheses. These problems are of a peculiar interest in the case of dependent observations: for example, the logarithmic derivative of a density is used when forming the optimal algorithms of nonlinear filtering and adaptive control of random processes (see, e.g., [Dobrovidov et al.(2012), Vasiliev et al.(2004)] and references therein).

We will construct estimators of $f(t)$ and $f'(t)$ using the following estimators $\hat{\xi}_n$ of noises ξ_n in (24):

$$\tilde{\xi}_n = x_n - \tilde{\lambda}_{n-1}x_{n-1}, \quad n = \overline{1, N}, \tag{25}$$

where $\tilde{\lambda}_n$ is the estimator defined in (7).

For estimation of the parameter λ with a guaranteed quality we have to know that, e.g., $\theta \in \Theta$, where $\Theta = \{\theta = (\lambda, \sigma^2, \sigma^{2(2\kappa+1)}) : \ |\lambda| \leq r, \ 0 < \underline{\sigma}^2 \leq \sigma^2, \ \sigma^{2(2\kappa+1)} \leq \overline{\sigma}^{2(2\kappa+1)}\}$, $r \in (0,1)$.

As an estimator of the ratio $\Psi(t)$ from the observations $(x_n)_{n\geq 1}$, one can take the ratio of statistics $\hat{f}_N^{(1)}(t)$ and $\hat{f}_N(t)$ of the form

$$\hat{f}_N^{(s)}(t) = \frac{1}{Nh_{s,N}} \sum_{n=1}^{N} K_a^{(s)}\left(\frac{t-\tilde{\xi}_n}{h_{s,N}}\right), \quad s = 0; 1, \tag{26}$$

where $K^{(s)}(z)$ are kernel functions, $h = (h_{s,N})_{N\geq 1}$ are sequences of positive numbers, $s = 0; 1$.

Estimators like (26) of the density and its derivatives from observations (25) were considered in [Dobrovidov et al.(2012), Vasiliev et al.(2004)], Section 4.1, where it was shown that we can establish asymptotic normality and convergence with probability one for estimators

$$\hat{\Psi}_N(t) = \hat{f}_N^{(1)}(t)/\hat{f}_N(t). \tag{27}$$

The results on asymptotic ratio estimation of the partial derivatives of the noise distribution density in multivariate dynamic systems are given in [Vasiliev et al.(2004)], Section 5.1. (see [Dobrovidov et al.(2012)] as well).

To obtain estimators of $\Psi(t)$ with a known MSE we apply Theorem 2.1.

Define the estimator

$$\tilde{\Psi}_N = \hat{\Psi}_N \chi(\hat{f}_N(t) \geq H), \quad N \geq 1.$$

By the definition (25), the estimators $\tilde{\xi}_n$ can be represented in the form

$$\tilde{\xi}_n = \xi_n + (\lambda - \tilde{\lambda}_{n-1}) x_{n-1}, \quad n = \overline{1, N}.$$

Using properties of estimators $\tilde{\lambda}_n$, defined in (7), we can find the known numbers C_1 and C_m, such that

$$\sup_{\Theta} \sum_{n=1}^{N} E_\theta (\lambda - \tilde{\lambda}_{n-1})^{2m} x_{n-1}^{2m} \leq \begin{cases} C_1 \log N, & m = 1, \\ C_m, & m > 1. \end{cases} \tag{28}$$

Similar relations where obtained in [Dobrovidov et al.(2012), Vasiliev et al.(2004)]. Then, using technique of Theorems 4.3.1 and 5.1.3 from [Dobrovidov et al.(2012), Vasiliev et al.(2004)] and (28), by appropriate chosen kernels in (26), we can find known numbers $C_{i,1}$ and $C_{i,2}$, $i = 0; 1$, such that

$$\sup_{\Theta, \Phi} E_f (\hat{f}_N(t) - f(t))^2 \leq \frac{C_{0,1}}{N h_{0,N}} + C_{0,2} h_{0,N}^{2\kappa},$$

$$\sup_{\Theta, \Phi} E_f (\hat{f'}_N(t) - f'(t))^2 \leq \frac{C_{1,1}}{N h_{1,N}^3} + C_{1,2} h_{1,N}^{2\kappa}.$$

Then it is natural to put

$$h_{0,N} = \left(\frac{C_{0,1}}{2\kappa C_{0,2}} \right)^{\frac{1}{2\kappa+1}} N^{-\frac{1}{2\kappa+1}},$$

$$h_{1,N} = \left(\frac{3 C_{1,1}}{2\kappa C_{1,2}} \right)^{\frac{1}{2\kappa+3}} N^{-\frac{1}{2\kappa+3}}$$

and for obviously defined numbers $\tilde{C}_0$ and $\tilde{C}_1$, we have

$$\sup_{\Theta, \Phi} E_f (\hat{f}_N(t) - f(t))^2 \leq \tilde{C}_0 N^{-\frac{2\kappa}{2\kappa+1}},$$

$$\sup_{\Theta, \Phi} E_f (\hat{f'}_N(t) - f'(t))^2 \leq \tilde{C}_1 N^{-\frac{2\kappa}{2\kappa+3}}.$$

It gives a possibility to apply Theorem 2.1 (taking into account Corollary 2.1) to estimation of $\Psi(t)$ for $H \in (0, c_f)$ with a known upper bound:

$$\sup_{\Theta, \Phi} E_f (\tilde{\Psi}(t) - \Psi(t))^2 \leq \tilde{C}_1 N^{-\frac{2\kappa}{2\kappa+3}} + \tilde{C}_0 N^{-\frac{2\kappa}{2\kappa+1}}. \tag{29}$$

4 Conclusion

We have presented a truncated modification of basic ratio type estimators constructed by dependent samples of finite size. This method gives a possibility to obtain estimators with a guaranteed accuracy in the sense of L_m-norm, $m \geq 2$ (3) on a time interval of a fixed length.

As an illustration, parametric and non-parametric estimation problems are considered. The presented method was applied to estimation of parameters of a linear autoregressive and a non-linear ARARCH-type process, as well as to estimation of a parameter of a non-Gaussian Ornstein-Uhlenbeck process by discrete-time observations (see properties (10), (11), (23), (16), (17), (18)).

Moreover, the estimators with a guaranteed accuracy in the mean square sense of a logarithmic derivative of noises density of an autoregressive process with an unknown dynamic parameter was investigated, see (29).

The presented method will be applied in the future to parametric and non-parametric estimation problems for multivariate dependent samples.

Truncated estimation method will be developed in adaptive problem statement for an unknown number g (see Theorem 2.1).

Acknowledgements

The author is grateful to Uwe Küchler for support of the paper.

References

[Arkoun(2011)] O. Arkoun. Sequential adaptive estimators in nonparametric autoregressive models. *Sequential Analysis: Design Methods and Applications*, 30 (2), 228-246, 2011.

[Arkoun and Pergamenchtchikov(2008)] O. Arkoun and S. Pergamenchtchikov. Nonparametric estimation for an autoregressive models. *Vestnik of Tomsk State University, Ser. Mathematics and Mechanics*, 2 (3), 20-30, 2008.

[Borovkov(1997)] Borovkov A.A. *Mathematical statistics.* Novosibirsk: Nauka and Sobolev Institute Press, 1997 (in Russian).

[Dobrovidov et al.(2012)] A.V.Dobrovidov, G.M.Koshkin and V.A.Vasiliev *Nonparametric models and statistical inference from dependent observations.* USA.: Kendrick Press, 2012. (to be published).
(Russian original: V.A.Vasiliev, A.V.Dobrovidov, G.M.Koshkin. Nonparametric estimation of functionals of stationary sequences distributions. Moscow.: Nauka, 2004. - 508 p.)

[Efroimovich(2007)] Efroimovich Sam. Sequential Design and Estimation in Heteroscedastic Nonparametric Regression. *Sequential Analysis: Design Methods and Applications*, 26 (1), 3-25, 2007. DOI: 10.1080/07474940601109670

[Konev(1985)] Konev V.V. *Sequential parameter estimation of stochastic dynamical systems.* Tomsk: Tomsk Univ. Press, 1985 (in Russian).

[Konev and Pergamenshchikov(1990a)] Konev V.V. and Pergamenshchikov S.M. Truncated sequential estimation of the parameters in random regression. *Sequential analysis*, 9:1 19-41, 1990.

[Konev and Pergamenshchikov(1990b)] Konev V.V. and Pergamenshchikov S.M. On truncated sequential estimation of the drifting parametermeant in the first order autoregressive models. *Sequential analysis*, 9:2 193-216, 1990.

[Cramér(1948)] Cramér G. *Mathematical methods of statistics.* Princeton Univ. Press, 1948.

[Küchler and Vasiliev(2010)] U. Küchler and Vasiliev V. On guaranteed parameter estimation of a multiparameter linear regression process. *Automatica*, Journal of IFAC, Elsevier, 46, 637-646, 2010.

[Liptser and Shiryaev(1977)] Liptser R.Sh. and Shiryaev A.N. *Statistics of random processes. I: General theory.* N. Y.: Springer-Verlag, 1977. *II: Applications.* N. Y.: Springer-Verlag, 1978.

[Malyarenko(2010)] Malyarenko A.A. Estimating the generalized autoregression model parameters for unknown noise distribution. *Automation and Remote Control*, 71 (2), 291-302, 2010.

[Politis and Vasiliev(2012a)] D. N. Politis and V.A.Vasiliev. Sequential kernel estimation of a multivariate regression function. *Proceedings of the IX International Conference 'System Identification and Control Problems', SICPRO'12*, Moscow, 30 January – 2 February, 2012, V. A. Tapeznikov Institute of Control Sciences, pages 996-1009.

[Politis and Vasiliev(2012b)] D. N. Politis and V.A.Vasiliev. Non-parametric sequential estimation of a regression function based on dependent observations. *Sequential Analysis*, pages 1-26, 2011 (submitted).

[Fourdrinier et al.(2009)] Fourdrinier D., Konev V. and Pergamenshchikov S. Truncated sequential estimation of the parameter of a first order autoregressive process with dependent noises. *Mathematical Methods of Statistics*, 18, 1, 43-58, 2009.

[Vasiliev et al.(2004)] V.A.Vasiliev, A.V.Dobrovidov and G.M.Koshkin, *Nonparametric estimation of functionals of stationary sequences distributions.* Moscow.: Nauka, 2004 (in Russian).

[Vasiliev(1997)] Vasiliev V.A. On Identification of Dynamic Systems of Autoregressive Type. *Automat. and Remote Control*, 12, 106-118, 1997.

[Vasiliev and Koshkin(1998)] Vasiliev V.A. and Koshkin G.M. Nonparametric Identification of Autoregression. *Probability Theory and Its Applications*, 43 (3), 577-588, 1998.

[Wald(1947)] Wald A. *Sequential analysis.* N. Y.: Wiley, 1947.