# On the Asymptotics of Distributions of Two-Step Statistical Estimates

Yu. Yu. Linke
Sobolev Institute of Mathematics

Consider a sequence $X_{n1}, X_{n2}, \dots, X_{nn}$, $n = 1, 2, \dots$, of series of observations independent in each series, whose distributions (in general, different within each series) depend on an unknown parameter $\theta_n$. At the first step, we find some estimate $\theta_n^* = \theta_n^*(X_{n1}, \dots, X_{nn})$ satisfying, say, a consistency requirement. At the second step we use $\theta_n^*$ to construct an estimate $\theta_n^{**}$ which more exactly approximates $\theta_n$, and in a series of cases is in a certain sense optimal.

The idea of two-step estimation procedures goes back to the articles of Fisher, who in addition used an iteration of Newton's method to calculate maximal likelihood estimates approximately. Following Fisher, the two-step estimates are usually expressed as $\theta_n^{**} = \theta_n^* + \sum\limits_{i=1}^{n} W_{ni}(\theta_n^*, X_{ni}) \Big/ \sum\limits_{i=1}^{n} V_{ni}(\theta_n^*, X_{ni})$, where $\{W_{ni}(\theta_n^*, X_{ni})\}$ and $\{V_{ni}(\theta_n^*, X_{ni})\}$ are chosen so that $\theta_n^{**}$ approximate the unknown parameter $\theta_n$ more exactly than the first-step estimate $\theta_n^*$.

The study of the asymptotic behavior of these estimates depends substantially on the choice of a first-step estimate $\theta_n^*$. Constructing consistent first-step estimates in various models constitutes a distinguished and difficult problem in general. But in a whole series of regression models, including various statements of linear and fractional linear regression problems, we succeed in constructing a first-step estimate $\theta_n^*$ for a parameter $\theta_n \in (-\infty, \infty)$ admitting the expression $\theta_n^* - \theta_n = \sum\limits_{i=1}^{n} u_{ni} \Big/ \Big(1 + \sum\limits_{i=1}^{n} v_{ni}\Big)$ for some transformations $\{u_{ni} = u_{ni}(\theta_n, X_{ni})\}$ and $\{v_{ni} = v_{ni}(\theta_n, X_{ni})\}$.

One of the our goals of this investigation is to obtain necessary and sufficient conditions for the convergence $(\theta_n^{**} - \theta_n)/d_n \Longrightarrow \eta$ as $n \to \infty$, where $\eta$ is a random variable whose distribution can be arbitrary, although the most widespread version in applications involves a normally distributed $\eta$.

We obtain the main results under sufficiently weak restrictions. In particulary, we essentially require only that the functions determining the estimate $\theta_n^{**}$ satisfy Hölder's condition, instead classical restriction, starting with the work of Cramér (for instance, the continuity of the first and second derivatives of the functions determining these estimates and the uniform boundedness of the third derivatives of these functions).

In addition, there is a series of regression problems in which we succeed in finding two-step estimates. Some of these problems were considered in the articles A.I. Sakhanenko and Yu.Yu. Linke devoted to estimations of a one-dimensional parameter in linear and fractional linear regression problems, allowing the possibility that a series of classical assumptions may not be satisfied (the variances of observations can depend on the unknown parameter $\theta_n$ and the index $n$ of observations; the observations are not assumed to be normally distributed; the coefficients may be measured with random errors). This investigation will enable us to substantially shorten proofs and make them possible under minimal restrictions on the functions determining two-step estimates.

[1] Linke Yu. Yu. *On the Asymptotics of Distributions of Two-Step Statistical Estimates* Siberian Math. J. 2011 V.52, N 4.