Abstracts of the Ninth Workshop on Simulation (2018) 1-2

Combined nonparametric estimators of probability characteristics

Yury G. Dmitriev and Gennady M. Koshkin [1]

Consider probability characteristics presented by the following functionals:

$$\theta = \theta(H) = \int_{R^s} \varphi(\vec{x}) dH(\vec{x}), \tag{1}$$

$$b_j = b_j(H) = \int_{R^s} \psi_j(\vec{x}) dH(\vec{x}), \quad j = \overline{1, m}, \tag{2}$$

where $\vec{x} = (x_1, \ldots, x_s)$, $H(\vec{x}) = \prod_{l=1}^{s} F(x_l)$, $dH(\vec{x}) = \prod_{l=1}^{s} dF(x_l)$, $F(x)$ is an unknown distribution function on R^1, $\varphi, \psi_j : R^s \to R^1$ are given functions.

Let us estimate θ under condition that b_j takes one of known values $\{\beta_{j,1}, \ldots, \beta_{j,k_j}\}$, $k_j \geq 1$, using independent sample $X_1, \ldots, X_N$ from $F(x)$. Thus, the problem can be formulated as follows: it is necessary to estimate (1) under the condition

$$\Delta_j(H) = \prod_{t=1}^{k_j} \Delta_{j,t}(H) = \prod_{t=1}^{k_j} (b_j(H) - \beta_{j,t}) = 0, \ j = \overline{1, m}. \tag{3}$$

The problem is to estimate (1) taking into account condition (3). In the special case $k_j = 1$, the problem was studied in [4] and [2]. With aim of using auxiliary information, several authors [1, 5, 6] have employed the empirical likelihood method. In all these works, it is shown that the estimators which take into account auxiliary information have smaller variances than the estimators without using auxiliary information.

Instead of unknown $H(\vec{x})$ in (1) and (3) substitute the estimates

$$\hat{H}_\tau(x_1, \ldots, x_s) = \frac{1}{|\omega_\tau|} \sum_{\{i_j\} \in \omega_\tau} \prod_{j=1}^{s} c(x_j - X_{i_j}), \tag{4}$$

where $c(t) = \{1 : t > 0; \ 0 : t \leq 0\}$, ω_τ is the set of index compositions $(i_1, \ldots, i_s)$, selected according to some rule with the index τ $(i_j = 1, \ldots, N; \ j = 1, \ldots, s)$, $|\omega_\tau|$ is the number of elements in the set ω_τ. Then, obtain $\hat{\theta}_\tau = \theta(\hat{H}_\tau)$, $\hat{\Delta}_{\tau,j} = \Delta_j(\hat{H}_\tau)$.

Introduce the following class of estimators:

$$\hat{\theta}_{\lambda_\tau} = \hat{\theta}_\tau - \lambda_\tau^{\mathrm{T}} \hat{\boldsymbol{\Delta}}_{\boldsymbol{\tau}}, \tag{5}$$

[1]Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russia, E-mail: dmit@mail.tsu.ru, kgm@mail.tsu.ru

where $\hat{\mathbf{\Delta}}_{\boldsymbol{\tau}} = (\hat{\Delta}_{\tau,1}, \ldots, \hat{\Delta}_{\tau,m})^{\mathrm{T}}$, $\lambda_\tau = V_\tau^{-1} C_\tau$, the matrix $V_\tau = E_F \hat{\mathbf{\Delta}}_{\boldsymbol{\tau}} \hat{\mathbf{\Delta}}_{\boldsymbol{\tau}}^{\mathbf{T}}$ is a non-degenerate matrix, V_τ^{-1} is the inverse matrix of V_τ, $C_\tau = E_F(\hat{\theta}_\tau - \theta)\hat{\mathbf{\Delta}}_{\boldsymbol{\tau}}$ is the matrix-column. Here, λ_τ minimizes MSE $E_F[\hat{\theta}_\lambda - \theta]^2$ for the given τ, and this minimum is equel to

$$E_F[\hat{\theta}_{\lambda_\tau} - \theta]^2 = E_F[\hat{\theta}_\tau - \theta]^2 - C_\tau^T V_\tau^{-1} C_\tau.$$

The non-negative quantity $C_\tau^T V_\tau^{-1} C_\tau$ determines the decrease of the MSE by attracting auxiliary information (3).

The estimator (5) combines (joins) the available empirical and prior information about the functionals (1) and (2). Statistical information about θ is contained by $\theta(\hat{H}_\tau)$, a priori information about b_j is given by $\{\beta_{j,1}, \ldots, \beta_{j,k_j}\}$, and statistical information is in $b_j(\hat{H}_\tau)$. Index τ corresponds to various methods of construction of an estimate, differing in the number of computational operations and the accuracy of estimation on the base of MSE.

It is important to note that the amount of this MSE's decrease depends on functionals (2), $\Delta_{j,t}$ in (3), τ, m, and k_j. For example, the increase of $|\Delta_{j,t}|$ and m reduces MSE, the increase of k_j increases MSE.

The statistic (5) can be used as an estimator for θ if you know λ_τ; otherwise you need to construct its estimate $\hat{\lambda}_\tau$. In the paper, adaptive estimators

$$\hat{\theta}_{\hat{\lambda}_\tau} = \hat{\theta}_\tau - \hat{\lambda}_\tau^{\mathrm{T}} \hat{\mathbf{\Delta}}_{\boldsymbol{\tau}} \tag{6}$$

are proposed. The asymptotic normality of all the proposed estimators is proved. Also, the mean square convergence for piecewise-smooth approximations [3] of estimators (6) is stated.

References

[1] Chen J., Qin J. *Empirical likelihood estimation for finite populations and the effective usage of auxiliary information* // Biometrika, 1993, v. 80, p. 107–116.

[2] Dmitriev Yu. G., Tarasenko F. P. *On the use of a priori information in estimated linear functionals of distribution* // Problems Control and Inform. Theory, 1978, v. 7, p. 459–469.

[3] Koshkin G. M. *Deviation moments of the substitution estimator and of its piecewise-smooth approximations* // Siberian Mathematical Journal, 1999, v. 40, p. 515–527.

[4] Levit B. Ya. *Conditional estimation of linear functionals* // Problems Inform. Trans., 1975, v. 11, p. 291–302.

[5] Owen A. B. *Empirical likelihood for linear models* // Ann. Statist., 1991, v. 19, p. 1725–1747.

[6] Qin J., Lawless J. *Empirical likelihood and general estimating equations* // Ann. Statist., 1994, v. 22, p. 300–325.