ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 23 МАТЕМАТИКИ Выпуск 2
2016

I. A. K a r e e v (Kazan, Kazan Federal University (KFU)). **Lower Bounds for the Expected Sample Size of Sequential Procedures for the Multinomial Selection Problem.**

Let $\xi = (\xi_1, \xi_2, \ldots, \xi_k)$ be a k-variate multinomial vector with outcome probabilities $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$. We consider the problem of selecting the most probable *population* ξ_i, i. e. the component of ξ associated with the largest-valued parameter $\theta_{[k]}$. We study the *indifference zone* approach where a selection procedure has to comply with the lower constraints on the probability of correct selection $P^* > 1/k$ only when $\theta \in \Theta_\Delta = \{\vartheta \in \Theta\colon \Delta\vartheta_{[k-1]} < \vartheta_{[k]}\}$, where $\Delta > 1$ defines size of the indifference zone.

Our main aim is to construct a lower estimate for the optimal expected sample size needed for solving the multinomial selection problem. Similar problems were considered early for general statistical selection [1] and ranking [2] problems.

Let $Sp(P^*, \Delta)$ be the set of all selection procedures for given P^*, Δ; ν_φ be the sample size induced by φ.

Theorem. *The expected sample size of any procedure* $\varphi \in Sp(P^*, \Delta)$ *and* $\theta \in \Theta_\Delta$ *is lower bounded by the value*:

$$\mathbf{E}_\theta \nu_\varphi \geqslant \omega(1 - P^*, 1 - P^*)\Big(\theta_{[k]} \ln \theta_{[k]} + (\theta_{[k-1]} + \theta_{[k]}) \ln \frac{1 + \Delta}{\theta_{[k-1]} + \theta_{[k]}} + \theta_{[k-1]} \ln \frac{\theta_{[k-1]}}{\Delta}\Big)^{-1},$$

where $\omega(x, y) = x \ln \frac{x}{1-y} - (1 - x) \ln \frac{1-x}{y}$.

One of the natural applications of the obtained lower bounds consists in its use for estimating *efficiency* of a selection procedure φ. Under the efficiency we consider the ratio $\inf_{\psi \in Sp(P^*,\Delta)} \mathbf{E}_\theta \nu_\psi / \mathbf{E}_\theta \nu_\varphi$ for $\theta \in \Theta_\Delta$. By substituting the lower bound for the numerator we gain the lower estimate for the efficiency of the φ.

Using this approach the efficiency of several selection procedures were investigated including single-sample procedure by Bechhofer–Elmaghraby–Morse [3] (depending on the parameters its efficiency varies from 0.1 to 0.5) and sequential procedure by Bechhofer–Goldsman [4] (the efficiency varies from 0.25 to 0.95).

REFERENCES

1. *Kareev I. A* Lower bounds for average sample size and efficiency of sequential selection procedures. — Theory Probab. Appl., 2013, v. 57, is. 2, p. 227–242.
2. *KareevĨ. A.* Lower bound for the average sample size and the efficiency of ranking sequential procedures. — Theory Probab. Appl., 2014, v. 58, is. 3, p. 503–509.
3. *Bechhofer R. E.*, *Elmaghraby S.*, *Morse N.* A single-sample multiple-decision procedure for selecting the multinomial event which has the highest probability. — Ann. Math. Statist., 1959, v. 30, №1, p. 102–119.

4. *Bechhofer R. E.*, *Goldsman D. M.* Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the multinomial event which has the largest probability (II): extended tables and an improved procedure. — Comm. Statist. Simulation Comput., 1986, v. 15, № 3, p. 829–851.