OPTIMAL CHOICE OF ORDER STATISTICS UNDER CONFIDENCE REGION ESTIMATION IN CASE OF LARGE SAMPLES

A. Zaigrajew[1], M. Alama-Bućko[2]
[1] *Nicolaus Copernicus University of Toruń*
[2] *University of Technology and Life Sciences of Bydgoszcz*
[1] *Toruń and* [2] *Bydgoszcz, POLAND*
e-mail: [1]alzaig@mat.umk.pl

Abstract

Let $x = (x_1, \ldots, x_n)$ be a sample from a distribution $P_\theta, \theta = (\theta_1, \theta_2)$, where $\theta_1 \in R$ is a location parameter and $\theta_2 > 0$ is a scale parameter. To estimate θ strong two-dimensional confidence regions of given confidence level $\alpha \in (0, 1)$ are considered. The quality of a Borel confidence set $B(x)$ is characterized by the risk function defined as $R(\theta, B) = \mathrm{E}_\theta \lambda_2(B(x))$, where $\lambda_2(B(x))$ is the Lebesgue measure of $B(x)$. Among confidence regions we distinguish those having the minimal risk and call them optimal. The method for construction of an optimal confidence region is well-known (see, e.g., [1]) and is based on using a pivot. Let $x_{i:n}$ represents the ith order statistic of the sample x for $i = 1, \ldots, n$. To construct a pivot two statistics t_1 and t_2 are taken; both statistics depend on given $k \leq n$ order statistics, say $t_1(x) = \sum_{i=1}^{k} a_i x_{m_i:n},\ t_2(x) = \sum_{i=1}^{k} b_i x_{m_i:n}$, where $1 \leq m_1 < m_2 < \ldots < m_k \leq n$. The case $k = 2$ was considered in [4]. If $k > 2$, then the problem of choosing $\{a_i, b_i\}$ is appeared. Here given $\{m_i\}$ the coefficients $\{a_i, b_i\}$ are taken in such a way that t_1 and t_2 are the asymptotically best linear estimators of θ_1 and θ_2, respectively (see, e.g., [3]). The main goal of the paper is to make the best choice of order statistics, that is the best choice of $\{m_i\}$, to minimize the risk function, as $n \to \infty$, under the assumptions that $m_i/n \to p_i,\ i = 1, \ldots, k,\ 0 \leq p_1 < p_2 < \ldots < p_k \leq 1$. It turns out that such a problem is quite close to that considered in e.g. [2], Section 10.4. In the paper the problem of choice the value of k is also discussed. Several examples of location-scale families of distributions are presented.

References

[1] Alama-Bućko M., Nagaev A.V., Zaigraev A. (2006). Asymptotic analysis of minimum volume confidence regions for location-scale families. *Applicationes Mathematicae (Warszawa)*. Vol. **33**, pp. 1-20.

[2] David H.A., Nagaraja H.N. (2003). *Order Statistics*. Wiley, New York.

[3] Masoom Ali M., Umbach D. (1998). Optimal linear inference using selected order statistics in location-scale models. In: *Handbook of Statistics*. Vol. **17**, pp. 183-213. North-Holland, Amsterdam.

[4] Zaigraev A., Alama-Bućko M. (2013). On optimal choice of order statistics in large samples for the construction of confidence regions for the location and scale. *Metrika*. Vol. **76**, pp. 577-593.