SIGNIFICANCE LEVEL ANALYSIS FOR ADAPTIVE ALGORITHM OF STATIONARY POISSON STREAM PROCESSING

V. I. Nikitsionak, A. M. Bachar
Belarusian State University
Minsk, BELARUS

Depending on the analyzed sample values the stationary Poisson stream (SPS) of events has the Poisson or the exponential law of intervals between the adjacent events. Here we consider the exponential law under the fixed number m of incoming events. Decision-making time T is assumed to be random. Decision-making means testing simple hypothesis H_0: the distribution parameter (or SPS intensity) $\lambda = \lambda_0$, against alternative H_1: $\lambda = \lambda_1 > 0$. In adaptive algorithm for the aim of finding the optimum decision threshold the intensity λ_0 is estimated using the classified training SPS of events (or the training set, TS), corresponding to SPS processing with $\lambda = \lambda_0$.

The significance level of adaptive algorithm is obtained from the one for optimal algorithm by averaging over all values of unknown parameter λ_0. Using the known approximation of probability integral, we get:

$$F_a(m, F; m_0) \cong a\sqrt{J}(F/a)^J, \ J = \left(1 + 2bm_0^{-1}\left(\sqrt{1/b\ln(a/F)} - d - \sqrt{m}\right)^2\right)^{-1}, \quad (1)$$

where F is a significance level, m_0 is the volume of TS, $(a, b, d) = (0.65, 0.443, 0.75)$. Note that the values $\sqrt{m}$ to be used depend on the required quality parameters of adaptive algorithm. Namely, they depend on power and significance level of decision rule, as well as on the ratio $\Lambda = \lambda_1/\lambda_0$, characterizing the “distance” between hypotheses. From the calculations based on (1) it follows: 1) at the small “distance” $\Lambda = 1.1$ the significance level $F = 10^{-4}$ is reached at rather large TS of volume $m_0 = 21000$, about 40 times greater than the one $m_0 = 500$, providing power parity of adaptive and optimum algorithms. The significance levels $10^{-5} \dots 10^{-6}$ are reached at even larger values of m_0; 2) at the larger “distance” $\Lambda = 2$ the significance levels $10^{-4} \dots 10^{-6}$ are reached at $m_0 = 500$ with adaptive and optimum algorithms both of power equal 0.9.

Thus, the adaptive algorithm has satisfactory quality for $\Lambda = 2$, unlike the case $\Lambda = 1.1$, when the required volume of TS grows dramatically.

Further analysis shows the following.

At a close hypothesis and alternative, when $\Lambda = 1.1$ is small enough, the distributions of the observations under hypothesis and alternative are rather close. So as the significance level depends on the left quite a gentle “tail” of the hypothetical distribution, there is a need for highly accurate estimate of a decision threshold. This high accuracy in its turn is achievable at a very large volumes of TS.

On the other hand, for the well separated hypothesis and alternative the accuracy of a threshold estimate at $m_0 = 500$ appears suitable for good quality of decision rule.