Goodness-of-fit tests based on the empirical characteristic function

Aleksej Bakshaev and Rimantas Rudzkis

Vilnius University, Lithuania, aleksej.bakshaev@gmail.com, rimantas.rudzkis@mii.vu.lt

Keywords: goodness-of-fit, empirical characteristic function

Let $X^n = (X_1, ..., X_n)$ be a sample of observations of a random vector X with unknown cumulative distribution function $F(x)$ and probability density function $f(x)$, $x \in \mathbb{R}^d$. The talk is devoted to the supremum-type multivariate goodness-of-fit tests based on the empirical characteristic function (ecf). Particular attention is devoted to the composite hypothesis of normality and Gaussian distribution mixtures model, which are widely applicable in classification problems. The null hypothesis assumes, that f_0 is a Gaussian density or a mixture of a known number of Gaussian densities. The alternative assumes the existence of an additional small distribution cluster g, that is

$$H_1 : f = (1 - \epsilon)f_0 + \epsilon g, \quad \epsilon \leq 1/2.$$

The problem of analytical approximation of the null distribution of the proposed test statistics, and therefore establishment of the critical region of the test, is briefly discussed. The results are obtained using the theory of high excursions of Gaussian (and, in some sense, close to Gaussian) random fields developed in [1]. Simulation study shows that the precision of the derived approximations is good enough even for small sample sizes and moderate test significance levels.

The comparative Monte Carlo power study shows that the considered tests are powerful competitors to the existing classical criteria, clearly dominating in verification of the goodness-of-fit hypotheses against the specific types of alternatives.

References

[1] Rudzkis, R. and Bakshaev A. (2012). Probabilities of high excursions of Gaussian fields. *Lithuanian Mathematical Journal* **52**, 196–213.