Tests based on characterizations, and their efficiencies

Ya. Yu. Nikitin

Saint-Petersburg University, Russia, y.nikitin@spbu.ru

Keywords: goodness-of-fit, asymptotic efficiency, characterizations

Suppose we have a sample $X_1, \dots, X_n$ of i.i.d. observations with distribution function (df) F, and we are testing the composite hypothesis $H_0 : F \in \mathcal{F}$, where $\mathcal{F}$ is some family of distributions, against the alternative $H_1 : F \notin \mathcal{F}$. Often the class $\mathcal{F}$ is characterized by the same distribution of two statistics $g_1(X_1, \dots, X_r)$ and $g_2(X_1, \dots, X_s)$.

Let us introduce two U-empirical df's

$$H_n^1(t) = \binom{n}{r}^{-1} \sum_{1 \le i_1 < \dots < i_r \le n} \mathbf{1}\{g_1(X_{i_1}, \dots, X_{i_r}) < t\}, \quad t \in R^1,$$

$$H_n^2(t) = \binom{n}{s}^{-1} \sum_{1 \le i_1 < \dots < i_s \le n} \mathbf{1}\{g_2(X_{i_1}, \dots, X_{i_s}) < t\}, \quad t \in R^1.$$

According to Glivenko-Cantelli theorem for U-empirical df's, $H_n^1(t)$ and $H_n^1(t)$ become very close under H_0 as $n \to \infty$. Consequently we can build the integral and Kolmogorov type goodness-of-fit tests based on the difference of $H_n^1(t)$ and $H_n^1(t)$.

Consider as an example of this idea the famous Polya's characterization [1]: *If X and Y are i.i.d. centered random variables (rv's), then the relation $X \stackrel{d}{=} (X+Y)/\sqrt{2}$ is valid iff X and Y are normal.* Another example is the Shepp's characterization [2] of normality with zero mean by the equal distribution of rv's X and *nonlinear* statistic $2XY/\sqrt{X^2+Y^2}$. We can construct the $U-$empirical goodness-of-fit statistics, then study their limiting distributions, and calculate their Pitman and Bahadur efficiencies. There are plenty of characterizations for the exponential, Pareto, Cauchy, arcsine, logistic, and other laws, and one can build and study corresponding tests as well. Some of them turn out to be rather efficient.

Goodness-of-fit tests based on characterizations of distributions were proposed by Yu. V. Linnik [3]. Due to technical difficulties, Linnik's idea was implemented only in last 10-15 years when the theory of $U-$empirical measures was sufficiently elaborated. We present a survey of recent results on goodness-of-fit and symmetry testing obtained by two small groups of researchers in Saint-Petersburg and Belgrade.

The author was supported by the grant of RFBR No. 16-01-00258.

References

[1] Polya, G. (1923). Herleitung des Gauss'schen Fehlergesetzes aus einer Funktionalsgleichung. *Math. Zeitschr.* **18**, 96–108.

[2] Shepp, L. (1964). Normal functions of normal random variables. *SIAM Rev.* **6**, 459–460.

[3] Linnik, Yu.V. (1953). Linear forms and statistical criteria I, II. *Ukrainian Math. J.*, **5**:2, 207–243; **5**:3, 247–290.