Metrika manuscript No.
(will be inserted by the editor)

Local exact Bahadur efficiencies of two scale-free tests of normality based on a recent characterization

Ya. Yu. Nikitin

Received: date / Accepted: date

Abstract We consider two scale-free tests of normality based on the characterization of the symmetric normal law by Ahsanullah, Kibria and Shakil (2014). Both tests have an U-empirical structure, but the first one is of integral type, while the second one is of Kolmogorov type. We discuss the limiting behavior of the test statistics and calculate their local exact Bahadur efficiency for location, skew and contamination alternatives.

Keywords Testing of normality · U-statistics · Bahadur efficiency · Kolmogorov test

Mathematics Subject Classification (2000) 62G10 · 62G20

1 Introduction

Testing normality is one of the oldest and most studied goodness of fit problems. Statistical tests for this problem are often based on *characterizations* of the normal law. Among various tests of this kind one may mention the tests developed in the papers of Csörgő et al. (1975), Sakata (1977), Lin and Mudholkar (1980), Muliere and Nikitin (2002), Ahmad and Mugdadi (2003), Volkova and Nikitin (2009),Villaseñor-Alva and Gonzalez-Estrada (2015), Litvinova and Nikitin (2016), and Bera et al.(2016).

In this paper we build and study two scale-free tests for the normal law based on a recent characterization of normality, which appeared in the book by Ahsanullah et al. (2014), see Theorem 8.2.8 there, and it looks as follows.

Theorem 1 *Let X,Y be symmetric independent random variables with a density p. Then the equality in distribution of* $\max^2(X,Y)$ *and* X^2 *is valid iff* $p(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{x^2}{2\sigma^2}), x \in R^1$, *with some variance* $\sigma^2 > 0$. *The same statement is true when replacing* $\max^2(X,Y)$ *by* $\min^2(X,Y)$.

Supported by Grant RFBR 16-01-00258 and Grant SPbGU - DFG 6.65.37.2017

Ya. Yu. Nikitin
Saint-Petersburg State University, Universitetskaya nab. 7/9, 199034, Russia

National Research University - Higher School of Economics, Souza Pechatnikov, 16,
St.Petersburg 190008, Russia,
E-mail: y.nikitin@spbu.ru

We consider testing the composite hypothesis H_0 according to which the sample $X_1,\dots,X_n$ of i.i.d. observations having a density belongs to a normal law $N(0,\sigma^2)$ with unspecified variance, against the general alternatives.

To this end, we introduce the so-called U-empirical distribution function (df),

$$G_n(t) = \binom{n}{2}^{-1} \sum_{1\le i\le j\le n} \mathbf{1}\{\max^2(X_i,X_j) < t\},\, t\ge 0,$$

and the so-called V-empirical df

$$\overline{G}_n(t) = n^{-2} \sum_{i,j=1}^{n} \mathbf{1}\{\max^2(X_i,X_j) < t\},\, t\ge 0.$$

Let $F_n(t)$ be the empirical df of $X_1^2,\dots,X_n^2$. Here and in what follows $\mathbf{1}\{A\}$ denotes the indicator function of the event A.

We consider the integral statistic

$$I_n = \int_0^\infty \left(\overline{G}_n(t) - F_n(t)\right) dF_n(t)$$

and the Kolmogorov type statistic

$$K_n = \sup_{t\ge 0} |\, G_n(t) - F_n(t)|.$$

The first of these statistics is motivated by the well-known statistic for testing exponentiality proposed by Hollander and Proshan (1972). One can construct also the ω^2-type statistic but the latter seems to be too complex for analytical evaluations; see, however, the paper of Allison and Pretorius (2017) dedicated to a Monte Carlo study of similar statistics. The Kolmogorov type statistic is well-known and needs no justification.

It is clear that under H_0 both statistics are scale-invariant so we may assume that the initial sample has unit variance. According to the Glivenko-Cantelli theorem for U-empirical df's (e.g., Helmers et al. (1988)), both statistics tend to zero a.s. under H_0, and this gives us the possibility to detect the alternative. We are primarily interested in the efficiency of the new tests in comparison with other tests of normality.

2 Bahadur efficiency

Bahadur efficiency is one of several possible approaches to evaluate the asymptotic relative efficiency (ARE) of two statistical tests. The Bahadur approach, proposed in Bahadur (1967, 1970), consists in fixing the power of concurrent tests, and then comparing the exponential rates of decrease of their sizes for an increasing number of observations under some fixed alternative.

For a sequence of statistics $\{T_n\}$ this exponential rate is usually proportional to some non-random function $c_T(\theta)$ (depending on the alternative parameter θ) which is called the *exact slope* of the sequence $\{T_n\}$. The Bahadur ARE $e^B_{V,T}(\theta)$ of two sequences of statistics $\{V_n\}$ and $\{T_n\}$ is defined by the formula

$$e^B_{V,T}(\theta) = c_V(\theta) \big/ c_T(\theta)\,.$$

The Bahadur exact slope of the sequence of test statistics $\{T_n\}$ can be evaluated as $c_T(\theta) = 2f(b_T(\theta))$, where $b_T(\theta)$ is the limit in probability of T_n under the alternative, while the continuous function $f(t)$ describes the logarithmic large deviation asymptotics of this sequence under the null-hypothesis, see details in Bahadur (1971) or Nikitin (1995).

It is important to note that there exists an upper bound for the exact slopes, see Bahadur (1967, 1970). We have

$$c_T(\theta) \le 2K(\theta)$$

where the Kullback–Leibler information number $K(\theta)$ measures the "statistical distance" between the alternative and the null-hypothesis. In the literature on estimation theory it is sometimes compared with the Cramér–Rao inequality. Therefore the absolute (non-relative) Bahadur efficiency of the sequence $\{T_n\}$ can be defined as $e_T(\theta) := c_T(\theta)/2K(\theta)$.

Computing the exact Bahadur ARE for arbitrary alternatives depending on θ is often infeasible; but it is possible to calculate the local Bahadur ARE as the alternative approaches the null-hypothesis. Then one speaks of *local* Bahadur slopes and *local* Bahadur efficiency

$$e_T^* := \lim_{\theta\to 0} c_T(\theta)/2K(\theta), \tag{1}$$

see Nikitin (1995).

The indisputable merit of Bahadur efficiency is its ability to handle statistics with non-normal asymptotic distributions. This is the main reason for using it in the present paper, as the Kolmogorov type statistics have a non-normal limiting distribution.

3 Integral statistic

We start with the integral statistic I_n. Note that

$$\int_0^\infty F_n(x)dF_n(x) = n^{-1}\sum_{j=1}^n F_n(X_j^2) = n^{-2}\sum_{j,k=1}^n \mathbf{1}\{X_k^2 \le X_j^2\}$$

$$= n^{-2}\left(n + \sum_{j\neq k}\mathbf{1}\{X_k^2 < X_j^2\}\right) = n^{-2}\left(n + \frac{n(n-1)}{2}\right) = \frac{1}{2} + \frac{1}{2n}\ .$$

Hence the statistic I_n differs only by $\frac{1}{2n}$ from the V-statistic (or the von Mises functional)

$$\widehat{I_n} = n^{-3}\sum_{i,j,k=1}^n H\{X_i, X_j, X_k\},$$

where H is the centered symmetric kernel

$$H(x,y,z) = \frac{1}{3}\left(\mathbf{1}\{\max{}^2(x,y) < z^2\} + \mathbf{1}\{\max{}^2(x,z) < y^2\} + \mathbf{1}\{\max{}^2(y,z) < x^2\}\right) - \frac{1}{2}.$$

It is well-known, see the seminal paper of Hoeffding (1948), that under weak conditions U- and V-statistics with the same kernel H are asymptotically equivalent. Consider the U-statistic

$$J_n = \binom{n}{3}^{-1}\sum_{1\le i<j<k\le n} H(X_i, X_j, X_k).$$

The following relationship is almost evident, see, however, section 1.3 of Koroluk and Borovskikh(1984):

$$\widehat{I_n} = \frac{(n-1)(n-2)}{n^2} J_n + 3n^{-3} \sum_{1 \le i < j \le n} (H(X_i, X_i, X_j) + H(X_i, X_j, X_j)) + n^{-3} \sum_{i=1}^{n} H(X_i, X_i, X_i).$$

We see that for bounded kernels H one has asymptotic equivalence $I_n \sim J_n$, $n \to \infty$. Hence in large samples we can consider the simpler U-statistic J_n instead of I_n. In the sequel, we will only deal with J_n.

In order to describe its limiting distribution and large deviation asymptotics, we must calculate the projection of this kernel and the variance of this projection. To simplify the calculations we first prove the following simple Lemma:

Lemma 1 *If X, Y are independent standard normal rv's , then we have*

$$P(X < |Y|) = P(|Y| > -X) = \frac{3}{4},$$

which follows from geometric considerations and is valid also for any absolutely continuous spherically symmetric distribution in the plane.

The projection $h(z), z \in R^1$, which we need has the form

$$h(z) = \mathbf{E}H(X, Y, z) = \frac{1}{3} P(\max^2(X, Y) < z^2) + \frac{2}{3} P(\max^2(X, z) < Y^2) - \frac{1}{2}. \qquad (2)$$

It is easy to calculate the first probability in (2):

$$P(\max^2(X,Y) < z^2) = P(-|z| < \max(X,Y) < |z|) = \Phi^2(|z|) - \Phi^2(-|z|) = 2\Phi(|z|) - 1. \quad (3)$$

The calculation of the second probability is more involved. We have

$$\begin{aligned} P(\max^2(X,z) < Y^2) &= P(\max(X,z) < |Y|) - P(\max(X,z) < -|Y|) \\ &= P(X < |Y|, z < |Y|) - P(X < -|Y|, z < -|Y|) =: P_1(z) - P_2(z). \quad (4) \end{aligned}$$

Obviously $P_1(z) = \frac{3}{4}$ if $z < 0$ and $P_2(z) = 0$ if $z > 0$. Now using Lemma 1, we have for $z \ge 0$

$$P_1(z) = \frac{3}{4} - P(X < |Y| < z) = \frac{3}{4} - \int_0^z 2\varphi(x)\Phi(x)dx = \frac{3}{4} - \int_0^z d\Phi^2(x) = 1 - \Phi^2(z).$$

Similarly, for $z < 0$ one has, again by Lemma 1,

$$\begin{aligned} P_2(z) &= P(|Y| < X, z < Y < -z) = P(|Y| < X) - P(-z < |Y| < X) \\ &= \frac{1}{4} - \int_{-z}^{\infty} 2\varphi(y)dy \int_y^{\infty} \varphi(x)dx = \frac{1}{4} + (1 - \Phi(y))^2|_{-z}^{\infty} = \frac{1}{4} - \Phi^2(z). \end{aligned}$$

Hence the second probability in (2) is $1 - \Phi^2(z)$ for $z > 0$ and $\frac{1}{2} + \Phi^2(z)$ for $z \le 0$. Now from (2) – (4) we obtain the following formula for the required projection $h(z)$ indicating that it is an odd function:

$$h(z) = \begin{cases} \frac{2}{3}(\Phi(z) - \Phi^2(z)) - \frac{1}{6}, \text{ if } z \ge 0; \\ \\ \frac{1}{6} - \frac{2}{3}(\Phi(z) - \Phi^2(z)), \text{ if } z < 0. \end{cases}$$

It follows that the variance of this projection is equal to

$$\int_0^1 \left(\frac{2}{3}(u - u^2) - \frac{1}{6} \right)^2 du = \frac{1}{180}.$$

Hence our U-statistic J_n is a non-degenerate one, and according to Hoeffding's theorem, see Hoeffding(1948) or Korolyuk and Borovskikh (1994), we have convergence in distribution

$$\sqrt{20n}\, J_n \xrightarrow{d} N(0,1) \text{ as } n \to \infty.$$

Using this theorem, we can establish the asymptotic critical domain for the rejection of the null-hypothesis under any prescribed significance level.

We can use now the large deviation statement under H_0 for non-degenerate U-statistics with bounded kernels which follows from Nikitin and Ponikarov (1999):

$$\lim_{n\to\infty} n^{-1} \ln P(J_n > v) = -f(v),$$

where $f(v)$ is continuous in a neighborhood of zero, and moreover

$$f(v) \sim 10v^2, \text{ as } v \to 0.$$

This result allows us to calculate the local Bahadur exact slope under any alternative to H_0. Consider, for example, the location alternative under which the sample has a df $\Phi(x+\theta)$ for some location parameter θ. The function $b_J(\theta)$ satisfies the relation,

$$b_J(\theta) \sim 3 \int_{R^1} h(t)t\varphi(t)dt \cdot \theta, \quad \text{as } \theta \to 0,$$

see Nikitin and Peaucelle (2004), and consequently the local exact slope has the form

$$c_J(\theta) \sim 180 \left(\int_{R^1} h(t)t\varphi(t)dt \right)^2 \cdot \theta^2.$$

Now we calculate numerically

$$720 \left(\int_0^\infty \left(\frac{2}{3}(\Phi(t) - \Phi^2(t)) - \frac{1}{6} \right) t\varphi(t)dt \right)^2 \approx 0.977.$$

As in this case $2K(\theta) \sim \theta^2$ as $\theta \to 0$, the local Bahadur efficiency (1) of the integral test is close to 0.977. This is a high value. The skew normal alternative gives the same value. We omit the calculations which can be done along the lines of Durio and Nikitin (2003, 2016).

To try other options, consider the contamination alternative

$$F(x,\theta) = (1-\theta)\Phi(x) + \theta\Phi^2(x) = \Phi(x) - \theta(\Phi(x) - \Phi^2(x)).$$

For this alternative it is easy to show that $2K(\theta) \sim \frac{4}{5}\theta^2$ as $\theta \to 0$. We need to calculate the integral

$$720 \left(\int_0^\infty \left(\frac{2}{3}(\Phi(t) - \Phi^2(t)) - \frac{1}{6} \right) \varphi(x)(1-2\Phi(x))dx \right)^2 = \frac{5}{16}.$$

Hence the local efficiency is not so high as compared to other tests and equals $\frac{25}{64} \approx 0.391$.

4 Kolmogorov type statistic

We return to the statistic $K_n = \sup_{t\geq 0}|G_n(t) - F_n(t)|$. It is based on the supremum of the *family* of U-statistics with the centered symmetric kernels depending on t :

$$\Psi(X,Y;t) = \mathbf{1}\{\max^2(X,Y) < t\} - \frac{1}{2}\left(\mathbf{1}\{X^2 < t\} + \mathbf{1}\{Y^2 < t\}\right), t \geq 0.$$

The limiting distribution of this statistic is unknown but one can use simulation to obtain its critical values.

Let us calculate the projection of this kernel for any fixed $t \geq 0$, namely

$$\psi(s;t) = \mathbf{E}\left(\Psi(X,Y;t)|Y = s\right) = P(\max^2(X,s) < t) - \frac{1}{2}(2\Phi(\sqrt{t}) - 1) - \frac{1}{2}\cdot\mathbf{1}\{s^2 < t\}.$$

It is evident that

$$\psi(s;t) = P(-\sqrt{t} < \max(X,s) < \sqrt{t}) - \frac{1}{2}(2\Phi(\sqrt{t}) - 1) - \frac{1}{2}\cdot\mathbf{1}\{|s| < \sqrt{t}\}.$$

But

$$P(\max(X,s) < \sqrt{t}) = P(X < \sqrt{t})\mathbf{1}\{s < \sqrt{t}\} = \Phi(\sqrt{t})\mathbf{1}\{s < \sqrt{t}\},$$
$$P(\max(X,s) < -\sqrt{t}) = P(X < -\sqrt{t})\mathbf{1}\{s < -\sqrt{t}\} = \Phi(-\sqrt{t})\mathbf{1}\{s < -\sqrt{t}\}.$$

Hence we have

$$P(-\sqrt{t} < \max(X,s) < \sqrt{t}) = \Phi(\sqrt{t})\mathbf{1}\{s < \sqrt{t}\} - \Phi(-\sqrt{t})\mathbf{1}\{s < -\sqrt{t}\},$$

Consequently,

$$\psi(X;t) = -\Phi(\sqrt{t})\mathbf{1}\{X > \sqrt{t}\} - \Phi(-\sqrt{t})\mathbf{1}\{X < -\sqrt{t}\} + \frac{1}{2}\cdot\mathbf{1}\{|X| > \sqrt{t}\}.$$

Thus, the variance function of our family of U-statistics (see Nikitin (2010)) is given by

$$\begin{aligned}\sigma^2(t) := \mathbf{E}\psi^2(X;t) = \mathbf{E}\left(\Phi(\sqrt{t})\mathbf{1}\{X > \sqrt{t}\} + \Phi(-\sqrt{t})\mathbf{1}\{X < -\sqrt{t}\} - \frac{1}{2}\cdot\mathbf{1}\{|X| > \sqrt{t}\}\right)^2 \\ = \Phi^2(\sqrt{t}) - \Phi^3(\sqrt{t}) + 1 - 3\Phi(\sqrt{t}) + 3\Phi^2(\sqrt{t}) - \Phi^3(\sqrt{t}) + \frac{1}{2} - \frac{1}{2}\Phi(\sqrt{t}) \\ - \Phi(\sqrt{t}) + \Phi^2(\sqrt{t}) - 1 + 2\Phi(\sqrt{t}) - \Phi^2(\sqrt{t}) \\ = -2\Phi^3(\sqrt{t}) + 4\Phi^2(\sqrt{t}) - \frac{5}{2}\Phi(\sqrt{t}) + \frac{1}{2}.\end{aligned}$$

Consider now the function

$$q(u) = -2u^3 + 4u^2 - \frac{5}{2}u + \frac{1}{2} > 0$$

in the interval $(1/2, 1)$. It is clear that the roots of the derivative are $\frac{1}{2}$ and $\frac{5}{6}$, so that the variance function attains its maximum $\frac{1}{27}$ at $u = 5/6$. Hence the large deviation asymptotics, see Nikitin (2010), has the following form:

$$\lim_{n\to\infty} n^{-1}\ln P(K_n > a) = -k(a),$$

where k is a continuous function in the neighborhood of zero, such that

$$k(a) = \frac{27a^2}{8}(1+o(1)),\, a \to 0.$$

Furthermore, the function $b_K(\theta)$ allows for the following asymptotics as $\theta \to 0$:

$$\begin{aligned}
b_K(\theta) &= \sup_{t\geq 0}|P_\theta(\max{}^2(X,Y)<t) - P_\theta(X^2<t)| \\
&= \sup_{t\geq 0}|\Phi^2(\sqrt{t}+\theta) - \Phi^2(-\sqrt{t}+\theta) - (\Phi(\sqrt{t}+\theta) - \Phi(-\sqrt{t}+\theta))| \\
&\sim 2\sup_{t\geq 0}|(2\Phi(\sqrt{t})-1)\varphi(\sqrt{t})|\theta = \sqrt{\frac{2}{\pi}}\sup_{t\geq 0}\big(2\Phi(\sqrt{t})-1)\exp(-t/2)\big)\,\theta \approx 0.337\theta.
\end{aligned}$$

It follows that finally $c_K(\theta) \sim \frac{27}{4}(0.337\theta)^2 \sim 0.764\theta^2$ as $\theta \to 0$.

The local Bahadur efficiency of the Kolmogorov test for the location alternative is then 0.764 which is distinctly high for a supremum type statistic. The same value appears for the skew alternative. In the case of the contamination alternative

$$F(x,\theta) = (1-\theta)\Phi(x) + \theta\Phi^2(x) = \Phi(x) - \theta(\Phi(x) - \Phi^2(x))$$

we have

$$F^2(x,\theta) = \Phi^2(x) - 2\theta\Phi(x)(\Phi(x)-\Phi^2(x)) + O(\theta^2), \theta \to 0.$$

Here the function $b_K(\theta)$ looks differently as $\theta \to 0$. After simple algebra we get

$$\begin{aligned}
b_K(\theta) &= \sup_{t\geq 0}|P_\theta(\max{}^2(X,Y)<t) - P_\theta(X^2<t)| \\
&= \sup_{t\geq 0}|F^2(\sqrt{t},\theta) - F^2(-\sqrt{t},\theta) - F(\sqrt{t},\theta) + F(-\sqrt{t},\theta)| \\
&\sim \sup_{x\geq 0}(2\Phi(\sqrt{x})(1-\Phi(\sqrt{x}))(2\Phi(\sqrt{x})-1))\,\theta \\
&= \sup_{1/2\leq z\leq 1}(2z(1-z)(2z-1))\,\theta \approx 0.192\theta,\, \theta \to 0.
\end{aligned}$$

Consequently, as $\theta \to 0$, we have $c_K(\theta) \sim \frac{27}{4}\cdot b_K^2(\theta) \approx 0.250\theta^2$, and the local Bahadur efficiency is 0.312.

5 Conclusion

Let us summarize the values of local efficiency in Table 1 where the first row corresponds to the integral statistic, while the second row corresponds to the Kolmogorov statistic.

Table 1 Local efficiencies of new tests

Test/Alternative	location	skew	contamination
I,J	0.977	0.977	0.391
K	0.764	0.764	0.312

We see that the integral test is rather efficient for the location and the skew alternatives. Remarkably many other integral tests of normality have *lower* local efficiency. For instance, the corresponding value for the classical ω^2-test is 0.907 (see Nikitin (1995)), for the test based on the Polya characterization (see Muliere and Nikitin (2002)) it is 0.966, and for the test based on the Shepp property (see Volkova and Nikitin (2009)) it is 0.955.

The Kolmogorov type test shows also high local efficiency. One should keep in mind that this test is *always consistent* while the integral test has mostly one-sided character, and its consistency depends on the alternative. The contamination alternative was taken on the grounds of facility, and our tests are not so efficient in this case.

Also, one should keep in mind that any hypothesis has to be tested with several possible criteria. The point of the matter is that with absolute confidence we can only reject it, while each new test which fails to reject the null-hypothesis gradually brings the statistician closer to the perception that this hypothesis is true. Hence, we are interested in developing new statistical tests based on novel ideas, specifically using characterizations. In our opinion, both tests proposed above can be added to the existing set of normality tests due to their high efficiency properties and relative simplicity.

6 Conflict of Interests

The author declares that there is no conflict of interests.

7 References

Ahmad I, Mugdadi AR (2003) Testing Normality Using Kernel Methods. Journ Nonpar Stat 15: 273 – 288.

Ahsanullah M, Kibria BG, Shakil M (2014) Normal and Student's t-distributions and their applications. Springer Science & Business Media.

Allison JS, Pretorius C (2017) A Monte Carlo evaluation of the performance of two new tests for symmetry. Comput Stat 32: 1323 – 1338

Bahadur RR (1967) Rates of convergence of estimates and test statistics. Ann Math Stat 38: 303–324.

Bahadur RR (1971) Some limit theorems in statistics. SIAM, Philadelphia.

Bera AK, Galvao AF, Wang L, Xiao Z (2016). A new characterization of the normal distribution and test for normality. Econom Theor, 32: 1216 – 1252.

Csörgő M, Seshadri V, Yalovsky M (1975). Applications of characterizations in the area of goodness of fit. Statist. Distrib. in Sci. Work, 2, 79 – 90.

Durio A., Nikitin YY (2003) Local Bahadur efficiency of some goodness-of-fit tests under skew alternatives. Journ Stat Plann Infer 115: 171 – 179.

Durio A., Nikitin YY (2016) Local efficiency of integrated goodness-of-fit tests under skew alternatives. Statist & Probab Lett 117: 136 – 143.

Helmers R, Janssen P, Serfling R (1988) Glivenko-Cantelli properties of some generalized empirical df's and strong convergence of generalized L-statistics, Probab Theor Relat Fields 79:75 – 93.

Hoeffding W (1948) A class of statistics with asymptotically normal distribution, Ann Math Stat 19: 293 – 395.

Hollander M, Proschan F (1972) Testing whether new is better than used. Ann Math Stat 43: 1136 – 1146.

Korolyuk VS, Borovskikh YuV (1994) Theory of U-statistics. Kluwer, Dordrecht.

Lin CC, Mudholkar GS (1980) A simple test for normality against asymmetric alternatives. Biometrika 67: 455 – 461.

Litvinova VV, Nikitin YY (2016) Kolmogorov tests of normality based on some variants of Polya's characterization. Journ Math Sci 5(219): 782 – 788.

Muliere P, Nikitin YY (2002) Scale-invariant test of normality based on Polya's characterization. Metron 60: 21 - 33.

Nikitin YY (2010) Large deviations of U-empirical Kolmogorov-Smirnov tests, and their efficiency. J Nonparam Stat 22: 649 – 668.

Nikitin YY, Peaucelle I (2004) Efficiency and local optimality of distribution-free tests based on U- and V-statistics. Metron 62: 185 – 200.

Nikitin YY, Ponikarov EV (1999) Rough large deviation asymptotics of Chernoff type for von Mises functionals and U-statistics. Proc. Saint-Petersburg Math Soc 7: 124 – 167; Engl. transl. in AMS Transl, ser 2, 2001, 203: 107 – 146.

Sakata T (1977) A test of normality based on some characterization theorem. Memoirs Fac of Sci Kyushu Univ, Ser A Math 31: 221 – 225.

Villaseñor-Alva JA, Gonzalez-Estrada Y (2015) A correlation test for normality based on the Lévy characterization. Comm Stat Simul Comput 44: 1225 - 1238.

Volkova KY, Nikitin YY (2009) On the asymptotic efficiency of normality tests based on the Shepp property. Vestnik St. Petersburg Univ: Mathematics 42: 256 – 261.