PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 34, Fasc. 2 (2014), pp. 317–336

DATA DRIVEN TESTS FOR UNIVARIATE SYMMETRY ABOUT AN UNKNOWN MEDIAN

BY

TADEUSZ INGLOT (WROCŁAW) AND ALICJA JANIC (WROCŁAW)

Dedicated to Teresa Ledwina on the occasion of her 65th birthday

Abstract. We propose new data driven score rank tests for univariate symmetry about an unknown center. We construct test statistics, state assumptions and define estimators of nuisance parameters. We prove that the test statistics are asymptotically distribution-free under the null hypothesis. Using simulations, we verify these asymptotic results for finite samples and show that, under the assumptions and when they are somewhat violated, the size of the test is stable when changing the null distribution. We also compare the empirical behaviour of the new tests with those proposed in the literature. We show that for families of distributions commonly applied to model asymmetry the new tests overcome their competitors on average and for most individual alternatives.

2000 AMS Mathematics Subject Classification: Primary: 62G10; Secondary: 62G30, 65C05, 65C60.

Key words and phrases: Testing symmetry, unknown center of symmetry, data driven score test, effective score, Schwarz type rule, model selection, rank test, kernel estimation, robust estimation, Monte Carlo study.

1. INTRODUCTION

The symmetry or asymmetry of the distribution of data about some center is important for appropriate interpretation of observed phenomena, for correct identification of models and for the validity of some statistical procedures (see, e.g., Ekström and Jammalamadaka [6] or Fernandes et al. [8] for more discussion). If a center of symmetry is known, many good tests have been constructed and investigated. For an up-to-date overview of them the reader is referred to Inglot et al. [17]. However, the more realistic case of when the center of symmetry is unknown leads to a far more difficult problem (cf. Lehmann and Romano [22]). Although it has been investigated by many authors, it still has no satisfactory solution. The recent constructions of Cabilio and Masaro [3], Mira [23], Holgerson [14], Zheng and

Gastwirth [29], Ekström and Jammalamadaka [6] and [7], or Ghosh [12] do not provide significant progress in comparison with the classical ones of Gupta [13], Gastwirth [11], Doksum et al. [5], Randles et al. [25] or Bhattacharya et al. [1]. Many of them exploit some coefficients of skewness, the most simple one being the difference between the sample mean and sample median. Even constructions which are asymptotically distribution-free are not able to keep a stable significance level for finite sample sizes. Also, robustness of an estimator of the median in a diverse set of null distributions is hard to obtain.

In this paper we propose new asymptotically distribution-free tests combining the general method of effective score tests and the method of model selection. Some attempt in this direction was made in Inglot et al. [18], where the good empirical behaviour of such a construction was shown. In Section 2 we state the problem, define a model, make assumptions and construct effective score statistics, as well as establish their asymptotic distribution under the null hypothesis. In Section 3 we specify estimators of all the nuisance parameters and quantities appearing in the test statistics for which the assumptions stated in Section 2 are fulfilled. In Section 4 we report the results of a simulation study. We compare the empirical behaviour of the new tests mainly with the test of Cabilio and Masaro [3], which seems to be one of the best solutions in the literature. To see how much loss in power is caused by the estimation of the median, we also compare them with the test of Modarres and Gastwirth [24] and with the data driven test of Inglot et al. [17], both designed for the case of the known center of symmetry. Some conclusions and an example of real data analysis are given in Section 5. All proofs are postponed until Section 6.

2. DATA DRIVEN SCORE STATISTICS

Let $X_1, \ldots, X_n$ be i.i.d. real random variables with an unknown median μ, a continuous distribution function $F(x-\mu)$ and density $f(x-\mu)$, where $F(0) = 1/2$. We are going to test

$$H_0 : F(x) = 1 - F(-x), \quad x \in \mathbb{R},$$

i.e., to test the symmetry of F about zero (or, equivalently, symmetry of the distribution of X_i about μ).

Furthermore, denote by $F_s(x) = \frac{1}{2}\big(F(x) + 1 - F(-x)\big)$ the distribution function of the symmetric part of F and by $F_a = F - F_s$ the distribution function of the antisymmetric part (signed measure) of F. Then H_0 is equivalent to testing whether $F = F_s$. Transform the shifted data $X_1 - \mu, \ldots, X_n - \mu$ into the unit interval using F_s to obtain $U_1, \ldots, U_n$ with $U_i = F_s(X_i - \mu),\ i = 1, \ldots, n$. Since F is absolutely continuous with respect to its symmetric part F_s, the transformed data U_i have the distribution function

$$F \circ F_s^{-1}(t) = t + F_a \circ F_s^{-1}(t) = t + A(t), \quad t \in [0, 1],$$

where $A = F_a \circ F_s^{-1}$ is an absolutely continuous function, symmetric with respect to $t = 1/2$. Equivalently, they have a density function on $[0, 1]$ of the form

$$p(t) = 1 + a(t),$$

where $a(t)$ is the antisymmetric, with respect to $t = 1/2$, derivative of $A(t)$. So, H_0 is equivalent to testing whether $a = 0$. Observe that $|a(t)| \leqslant 1$ a.s. and contains all the information about the asymmetry of F.

Let $d(n)$ be a nondecreasing sequence of natural numbers. Consider the nested sequence $\{\mathcal{G}_k;\ 1 \leqslant k \leqslant d(n)\}$ of exponential families given by the densities

$$g_k(x, \vartheta, \mu, f_s) = c_k(\vartheta) \exp\Big\{ \sum_{j=1}^{k} \vartheta_j \psi_j\big(F_s(x - \mu)\big)\Big\} f_s(x - \mu), \tag{2.1}$$

where $\psi_1, \psi_2, \ldots$ denotes an orthonormal system of bounded functions on $[0, 1]$, antisymmetric with respect to 1/2, $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_k)^T \in \mathbb{R}^k$ is a vector of parameters, y^T stands for the transposition of the vector y, and $c_k(\vartheta)$ is the normalizing constant.

Fix k between 1 and $d(n)$ and assume that the density f belongs to $\mathcal{G}_k$. Then H_0 reduces to the parametric hypothesis $H_0' : \vartheta = 0$ in the presence of nuisance parameters μ and f_s. In the sequel, we shall denote a null distribution by $P_{\mu f_s}$.

Our basic assumptions are as follows:

(A1) f_s is absolutely continuous with finite Fisher information

$$J = J_{f_s} = \int_{\mathbb{R}} \frac{\big(f_s'(x)\big)^2}{f_s(x)} dx;$$

(A2) f_s'/f_s is linearly independent of $\psi_1\big(F_s(\cdot)\big), \psi_2\big(F_s(\cdot)\big), \ldots$

Under the above assumptions we have the following proposition.

PROPOSITION 2.1. *For every $1 \leqslant k \leqslant d(n)$, under* (A1) *and* (A2), *the densities g_k are differentiable in quadratic mean at $\vartheta = 0$ for any μ and any f_s with respect to three parameters of the model $\mathcal{G}_k$, and the effective score vector has the form*

$$\ell_*(x) = \Psi\big(F_s(x - \mu)\big) + \frac{v}{J} \frac{f_s'}{f_s}(x - \mu) \ a.e., \tag{2.2}$$

where $v = -\int_{\mathbb{R}} \Psi\big(F_s(x)\big) f_s'(x) dx$ is a column matrix depending only on f_s and $\Psi(t) = \big(\psi_1(t), \ldots, \psi_k(t)\big)^T$.

Note that v consists of the Fourier coefficients of $-f_s'/f_s$ with respect to $\Psi \circ F_s$ under $P_{0 f_s}$. Proposition 2.1 can be proved by a standard argument (see Inglot and Janic [16]), so here we omit its proof. Some remarks are also given in

Section 6 below. A typical example of the system $\psi_1, \psi_2, \ldots$, which we shall apply below, is the sequence $b_1, b_3, \ldots$ of the odd Legendre polynomials on the unit interval (cf. Sansone [26]).

By the orthonormality of Ψ, the covariance matrix of ℓ_* under $P_{\mu f_s}$ takes the form $I_* = I_k - J^{-1}vv^T$, where I_k denotes the unit $(k \times k)$ matrix.

In consequence, the effective score statistic for testing H_0' is given by the formula

$$W_k = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ell_*(X_i)\right)^T I_*^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ell_*(X_i)\right). \tag{2.3}$$

Obviously, due to the classical multivariate central limit theorem the statistic W_k converges in distribution, under H_0, to the chi-square distribution with k degrees of freedom. Note that $I_*^{-1} = I_k + v(J - v^T v)^{-1} v^T$. By (A2) and the Bessel inequality, $J - v^T v$ is positive and nonincreasing in k.

The statistic W_k given in (2.3) depends on unknown nuisance parameters. To apply it to testing H_0, we have to estimate the parameters $\mu, F_s, f_s'/f_s, v$ and J appearing in it. We make natural assumptions on the estimators of these quantities.

(A3) Let $\widehat{\mu}$ be a $\sqrt{n}$-consistent estimator of μ (under $P_{\mu f_s}$), i.e., such that for every $\varepsilon > 0$ and each f_s there exists a positive constant M satisfying the condition $P_{\mu f_s}(\sqrt{n}|\widehat{\mu} - \mu| \geqslant M) < \varepsilon$ for sufficiently large n.

Let $Z = (X_1 - \mu, \ldots, X_n - \mu, \mu - X_1, \ldots, \mu - X_n)$ denote the pooled sample. Then the empirical distribution function $\mathcal{F}_{ns}$ of Z is an estimator of F_s.

(A4) Let $\overline{f_s'/f_s}$ be an estimator of f_s'/f_s based on Z, which is an odd function and satisfies

$$\int_{\mathbb{R}} \left(\overline{\frac{f_s'}{f_s}}(x) - \frac{f_s'}{f_s}(x)\right)^2 f_s(x)\, dx \xrightarrow{P_{\mu f_s}} 0 \quad \text{as } n \to \infty.$$

(A5) Let $\overline{v}$ and $\overline{J}$ be consistent (under $P_{\mu f_s}$) estimators of v and J, respectively, based on Z.

Let $\widehat{\mu}_d$ be a discretization of $\widehat{\mu}$ (see, e.g., Bickel et al. [2], p. 44). Since μ is unknown, the estimators $\mathcal{F}_{ns}$, $\overline{f_s'/f_s}$, $\overline{v}$ and $\overline{J}$ introduced above cannot be applied directly. So, denote by $\widehat{Z} = (X_1 - \widehat{\mu}_d, \ldots, X_n - \widehat{\mu}_d, \widehat{\mu}_d - X_1, \ldots, \widehat{\mu}_d - X_n)$ the estimated pooled sample Z. Then we may apply the empirical distribution function $\widehat{\mathcal{F}}_{ns}$ of $\widehat{Z}$ as an estimator of F_s. Next, let us put

$$\mathcal{N}_1 = \{1, \ldots, \lfloor n/2 \rfloor, n+1, \ldots, \lfloor (3n)/2 \rfloor\} \quad \text{and} \quad \mathcal{N}_2 = \{1, \ldots, 2n\} \setminus \mathcal{N}_1.$$

If $\overline{f_s'/f_s}$ is an estimator of f_s'/f_s satisfying (A4), then consider the two estimators of this score function obtained when replacing Z by one of the two parts of $\widehat{Z}$ and denoted by

$$\left(\widehat{\frac{f_s'}{f_s}}\right)_j \quad \text{when based on } \{\widehat{Z}_i;\ i \in \mathcal{N}_j\}, \quad j = 1, 2.$$

Then we get two estimators of ℓ_* of the form

$$\widehat{\ell}_{*j} = \Psi\big(\widehat{\mathcal{F}}_{ns}(\cdot - \widehat{\mu}_d)\big) + \frac{\widehat{v}}{\widehat{J}}\left(\widehat{\frac{f_s'}{f_s}}\right)_j (\cdot - \widehat{\mu}_d), \quad j = 1, 2, \tag{2.4}$$

where $\widehat{v},\ \widehat{J}$ are based on $\widehat{Z}$ and satisfy (A5) when replacing $\widehat{Z}$ by Z.

Now, set

$$\widehat{I_*^{-1}} = I_k + \widehat{v}(\widehat{J} - \widehat{v}^T\widehat{v})^{-1}\widehat{v}^T, \tag{2.5}$$

which is an estimator of I_*^{-1},

$$\overline{\widehat{\ell}_*} = \frac{1}{\sqrt{n}}\Big(\sum_{i=1}^{\lfloor n/2 \rfloor} \widehat{\ell}_{*2}(X_i) + \sum_{i=\lfloor n/2 \rfloor+1}^{n} \widehat{\ell}_{*1}(X_i)\Big) \tag{2.6}$$

and the corresponding estimated score statistic

$$\widehat{W}_k = \overline{\widehat{\ell}_*}^T\, \widehat{I_*^{-1}}\, \overline{\widehat{\ell}_*} \tag{2.7}$$

for testing H_0' in the family $\mathcal{G}_k$. We take the specific form (2.6) of the estimator of ℓ^* in order to prove (2.10), which guarantees that $\widehat{W}_k$ is asymptotically distribution-free (cf. Theorem 2.1 below).

It is well known that the choice of k is crucial to the behaviour of a test based on a score statistic. So, to adapt this choice to the data we apply a Schwarz type selection rule (cf. Schwarz [28] or, e.g., Inglot et al. [17])

$$S = \min\{1 \leqslant k \leqslant d(n) : \widehat{W}_k - k \log n = \max_{1 \leqslant j \leqslant d(n)} (\widehat{W}_j - j \log n)\}. \tag{2.8}$$

As an alternative to S, we take the less conservative selection rule L, which was introduced in Inglot and Janic [15] and applied to testing symmetry in Inglot et al. [17]. We recall its definition adapted to our present need. Let $1 \leqslant D < d(n)$ be a fixed number not depending on n, and let δ_n be a small positive number. Define the thresholds $c_{jn},\ j = 1, \ldots, D$, to be the solutions of the equations

$$1 - \Phi(c_{jn}) = \frac{1}{2}\left(\delta_n D^{-1} \binom{d(n)}{j}^{-1}\right)^{1/j},$$

where Φ denotes the standard normal distribution function. Next, consider the standardized vector $\mathcal{L} = (\widehat{I_*^{-1}})^{1/2}\overline{\widehat{\ell}_*}$ for $k = d(n)$, with $\widehat{I_*^{-1}}$ and $\overline{\widehat{\ell}_*}$ defined in (2.5) and

(2.6), and order the squares of its components from the smallest to the largest, obtaining $\mathcal{L}^2_{(1)}, \dots, \mathcal{L}^2_{(d(n))}$, and consider the event

$$E_n = \{\mathcal{L}^2_{(d(n))} \geqslant c^2_{1n}\} \cup \ldots \cup \{\mathcal{L}^2_{(d(n)-D+1)} \geqslant c^2_{Dn}\}.$$

Then define the data dependent penalty

$$\pi(j, n) = j \log n \cdot \mathbf{1}_{E^c_n} + 2j \cdot \mathbf{1}_{E_n},$$

where $\mathbf{1}_{E_n}$ denotes the indicator of the set E_n, and E^c_n denotes the complement of E_n, and define the corresponding selection rule L:

$$L = \min\big\{1 \leqslant k \leqslant d(n) : \widehat{W}_k - \pi(k, n) = \max_{1 \leqslant j \leqslant d(n)} \big(\widehat{W}_j - \pi(j, n)\big)\big\}. \tag{2.9}$$

Taking into account all the above considerations, we can apply $\widehat{W}_S$ and $\widehat{W}_L = \widehat{W}_L(D, \delta_n)$ as statistics of upper-tailed data driven tests for testing H_0.

The asymptotic behaviour of $\widehat{W}_S$ and $\widehat{W}_L$ is a consequence of the following basic proposition proved in Section 6.

PROPOSITION 2.2. *Under* (A1), (A2) *and for any estimators of* μ, f'_s/f_s, v *and* J *such that* (A3)–(A5) *are fulfilled we have*

$$\overline{\widehat{\ell}_*} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ell_*(X_i) \xrightarrow{P_{\mu f_s}} 0 \quad \text{as } n \to \infty, \tag{2.10}$$

where $\overline{\widehat{\ell}_*}$ *is defined by* (2.4) *and* (2.6), *while* ℓ_* *is given in* (2.2).

Observe that the consistency of the estimators $\widehat{v}$ and $\widehat{J}$ has already been shown in the proof of Proposition 2.2. Hence, $\widehat{I^{-1}_*}$, given in (2.5), is a consistent estimator of I^{-1}_*. Now, the properties of weak convergence and (2.10) imply the main asymptotic result for $\widehat{W}_k$.

THEOREM 2.1. *Under* H_0 *and* (A1)–(A5) *the following holds*:

$$\widehat{W}_k \xrightarrow{\mathcal{D}} \chi^2_k,$$

where $\widehat{W}_k$ *is defined in* (2.7) *and* χ^2_k *denotes a random variable from the chi-square distribution with* k *degrees of freedom.*

COROLLARY 2.1. *Suppose* $d(n)$ *is a bounded sequence and* (A1)–(A5) *are fulfilled. Then under* H_0 *it follows that*

$$P_{\mu f_s}(S = 1) \to 1 \quad \text{and} \quad \widehat{W}_S \xrightarrow{\mathcal{D}} \chi^2_1 \quad \text{as } n \to \infty. \tag{2.11}$$

If, in addition, $\delta_n \to 0$ *is such that* $\log(1/\delta_n) = o(n)$, *then*

$$P_{\mu f_s}(L = S) \to 1 \quad \text{and} \quad \widehat{W}_L \xrightarrow{\mathcal{D}} \chi^2_1 \quad \text{as } n \to \infty. \tag{2.12}$$

P r o o f. For (2.11) it is enough to see that for $k > 1$ we have $P_{\mu f_s}(S = k) \leqslant P_{\mu f_s}(\widehat{W}_k > \log n)$, which tends to zero by Theorem 2.1. Using an analogous argument to that in the proof of Theorem 3.1 in Inglot et al. [17] and boundedness of $d(n)$ we see that $c_{Dn}^2 \to \infty$ and $P_{\mu f_s}(E_n) \leqslant 2^{d(n)} P_{\mu f_s}(\widehat{W}_{d(n)} \geqslant c_{Dn}^2)$, which, again by Theorem 2.1, tends to zero. Hence $P_{\mu f_s}(L = S) \to 1$ and (2.12) follows. ■

Corollary 2.1 states that the data driven tests of H_0 based on $\widehat{W}_S$ and $\widehat{W}_L$ are asymptotically distribution-free. Since the convergence in (2.11) and (2.12) is rather slow, we shall use simulated critical values. As we show empirically in Section 4 they behave quite stably for moderate sample sizes when null densities f_s change as long as (A1) is fulfilled. For further discussion of this question see Subsection 4.1.

3. TESTS SPECIFICATION

In this section we specify all estimators needed to calculate the data driven test statistics $\widehat{W}_S$ and $\widehat{W}_L$ and to get our new tests of symmetry which we shall study empirically and which we recommend as good solutions for testing symmetry.

As an orthonormal system we take a sequence $b_1, b_3, \ldots$ of the odd Legendre polynomials. Since for typical null distributions $b_1 \circ F_s$ is strongly correlated with f_s'/f_s, resulting in large value of v_1 and negligibility of the first component of ℓ_*, we omit b_1 and set $\Psi = b = (b_3, \ldots, b_{2k+1})^T$ in (2.4).

The usual estimator of the median is the sample median. However, it has a large variance if $f_s(0)$ is close to zero (e.g., when f_s is bimodal). To overcome this problem, choose $q \in (0, 1/2]$ and consider the estimator

$$\widehat{\mu} = (X_{(\lfloor nq \rfloor)} + X_{(n - \lfloor nq \rfloor + 1)})/2, \tag{3.1}$$

where $X_{(i)}$ denotes the i-th order statistic of the original sample $X_1, \ldots, X_n$. It is easy to check that, under H_0, this estimator is unbiased. If we restrict our attention to null distributions F_s satisfying

$$f_s\big(F_s^{-1}(q)\big) > 0 \tag{E1}$$

then, by classical asymptotic results on L-statistics, $\widehat{\mu}$ is asymptotically normal, and therefore (A3) is fulfilled.

To estimate f_s'/f_s we use kernel estimators. To this end let K be a kernel satisfying

(E2) K is a symmetric probability density uniformly bounded by C, twice differentiable with $|K'| \leqslant CK,\ |K''| \leqslant CK$.

Additionally assume

(E3) γ_n, h_n are sequences of positive numbers such that $\gamma_n \to 0$, $h_n \to 0$ and $n\gamma_n^2 h_n^6 \to \infty$.

Consider a random bandwidth $\overline{h}_n$ based on Z and satisfying

$$\text{(E4)} \qquad \sqrt{n}\left(\frac{\overline{h}_n}{h_n} - 1\right) = O_{P_{\mu f_s}}(1).$$

Now, define nonparametric kernel estimators

$$\text{(3.2)} \qquad \overline{f_s}(x) = \gamma_n + \frac{1}{2n\overline{h}_n} \sum_{i=1}^{2n} K\left(\frac{x - Z_i}{\overline{h}_n}\right), \qquad \overline{f_s'}(x) = \frac{d}{dx}\left(\overline{f_s}(x)\right)$$

and, in consequence,

$$\text{(3.3)} \qquad \overline{\frac{f_s'}{f_s}}(x) = \frac{\overline{f_s'}(x)}{\overline{f_s}(x)}.$$

Then we have the following proposition.

PROPOSITION 3.1. *Under* H_0 *and* (E2)–(E4) *the estimator* $\overline{f_s'/f_s}$ *given in* (3.3) *satisfies* (A4).

Replacing in (3.2) and (3.3) Z by two parts of $\widehat{Z}$ we get estimators $(\widehat{f_s'/f_s})_j$, $j = 1, 2$, used in (2.4).

Plugging estimators of unknown quantities, based on Z, into the formula defining the column matrix v defined in (2.2) we obtain

$$\widetilde{v} = -\int_{\mathbb{R}} b\big(\mathcal{F}_{ns}(x)\big)\overline{f_s'}(x)dx$$

with $\overline{f_s'}$ as in (3.2). Since

$$\widetilde{v} = \sum_{i=1}^{2n-1} b\left(\frac{i}{2n}\right)\left(\overline{f_s}(Z_{(i)}) - \overline{f_s}(Z_{(i+1)})\right) + b(1)\left(\overline{f_s}(Z_{(1)}) + \overline{f_s}(Z_{(2n)})\right),$$

omitting the last summand we get an estimator of v, we shall apply in the sequel, of the form

$$\text{(3.4)} \qquad \overline{v} = \sum_{i=1}^{2n-1} b\left(\frac{i}{2n}\right)\left(\overline{f_s}(Z_{(i)}) - \overline{f_s}(Z_{(i+1)})\right).$$

The following proposition holds.

PROPOSITION 3.2. *Under* H_0 *and* (E2)–(E4), $\overline{v}$ *is a consistent estimator of* v.

Finally, we estimate the Fisher information J by a natural moment estimator

$$\text{(3.5)} \qquad \overline{J} = \frac{1}{2n} \sum_{i=1}^{2n} \left(\overline{\frac{f_s'}{f_s}}\right)^2 (Z_i) = \frac{1}{n} \sum_{i=1}^{n} \left(\overline{\frac{f_s'}{f_s}}\right)^2 (X_i - \mu).$$

PROPOSITION 3.3. *Under* H_0 *and* (E2)–(E4), $\overline{J}$ *is a consistent estimator of* J.

Proofs of Propositions 3.1, 3.2 and 3.3 are given in Section 6.

4. SIMULATION STUDY

In this section we present results of a simulation study for data driven tests based on $\widehat{W}_S$ and $\widehat{W}_L$. We discuss critical values and empirical powers of these tests. We restrict our attention to the sample size $n = 100$, $d(n) = 10$ and additionally, for the second test, to $D = 3$ and $\delta_n = 0.05$. For notational convenience, we shall denote here the corresponding tests by WS and WL, respectively. In simulations we took a typical significance level $\alpha = 0.05$. Every Monte Carlo experiment was repeated 10,000 times.

As an estimator of the median we took that given by (3.1) with $q = 0.3$. Such a choice is a compromise between unimodal and bimodal null densities and was adjusted empirically. For kernel estimators we took the standard Gaussian kernel $K = \phi$, $\gamma_n = 0.0001n^{-1/24}$ and $\overline{h}_n = 0.8\min\{0.8IQR, s\}n^{-3/20}$. Formally, to fulfil (E2) this kernel requires some modification which, however, does not influence the simulation results. So, we omit the precise description. We also neglected splitting the sample into two parts as described in (2.4) and estimated f_s'/f_s from the whole pooled sample.

4.1. Critical values. It is well known that data driven score statistics slowly converge to their asymptotic distribution. We expect a similar behaviour of WS and WL. Therefore, we determine empirical critical values. We took five symmetric distributions with finite Fisher information (with bounded support as well as distributed over the whole real line) including four unimodal distributions and one bimodal distribution. The results are shown in Table 1. By t_n we have denoted t-Student's distribution with n degrees of freedom. For description of LC(ρ, θ) see Subsection 4.2.

It can be seen that for unimodal smooth densities empirical critical values are quite stable, which attests that WS and WL are distribution-free (under the assumptions). However, when the density becomes less smooth (like Beta$(3, 3)$), the corresponding critical value takes larger values. Moreover, for bimodal distributions (like LC$(0.5, 0.75)$) the estimator $\widehat{\mu}$ has a large variance resulting in a much larger critical value. Taking this into account we recommend for a practical use an average of the five simulated critical values from Table 1 as the (fixed) critical values equal to 6.239 and 7.661 for WS and WL, respectively. Such a choice ensures that empirical sizes of the tests will fluctuate reasonably around the nominal level 0.05 when varying a null distribution, and for typical smooth, unimodal null densities will be a little bit less than 0.05.

TABLE 1. Simulated critical values of WS and WL for different symmetric distributions. $\alpha = 0.05$, $n = 100$, $d(100) = 10$; 10,000 MC runs

	Normal	t_3	Cauchy	Beta(3,3)	LC(0.5,0.75)	**Average**
WS	5.520	5.071	5.702	6.328	8.572	**6.239**
WL	6.786	6.056	6.686	7.952	10.827	**7.661**

To check what happens when the assumption (A1) is violated we simulated empirical size of WS for the Beta$(2,2)$ distribution with the infinite Fisher information (note that Beta(ξ,ξ), $\xi > 2$, have already finite Fisher information) obtaining 0.063 and for the uniform distribution (a discontinuous density) obtaining 0.210. On the other hand, when (A1) is fulfilled but f_s is bimodal with $f_s(0)$ close to zero we simulated empirical size of WS for LC$(0.5, 0.65)$ obtaining 0.089. Additionally, when $f_s(0)$ is equal to zero we considered the mixture of the chi-square distribution with six degrees of freedom with its reflection about zero and obtained much larger value 0.14. It is worth noting that Cabilio and Masaro [3] test also has troubles with keeping the nominal size when (A1) does not hold but for bimodal distributions behaves crazy getting a "size" 0.75 for the above-described two-sided chi-square distribution with six degrees of freedom. Note also that many popular families of asymmetric distributions like the Tukey family (see the next subsection) do not satisfy (A1). For example, Tukey distributions have finite Fisher information when positive parameters λ_3 and λ_4 are less than 1 and at least one of them is less than 1/2.

4.2. Power behaviour. To compare empirical powers of the new tests WS and WL we select three tests as their competitors:

• the Cabilio and Masaro [3] test, based on the difference between sample mean and sample median, denoted here by CM;

• the Mira [23] test, based on Bonferroni measure of skewness, denoted here by $Mira$;

• the Ekström and Jammalamadaka [6] test, based on sample spacings, denoted here by EJ.

To make our comparison more informative we also include two powerful tests designed for the case of known center of symmetry: the hybrid test of Modarres and Gastwirth [24], denoted here by MG, and the data driven test with a Schwarz type selection rule, studied recently by Inglot et al. [17], and denoted here by NS.

We took two families of asymmetric distributions frequently appearing in the literature. The first one is the Generalized Lambda Family, denoted here by Lambda (cf., e.g., Cabilio and Masaro [3]), described by two shape parameters λ_3 and λ_4. As our test statistics are location and scale invariant we omit giving parameters λ_1 (location) and λ_2 (scale) for distributions from the Lambda family. The second one is the Generalized Tukey-Lambda Family (see Freimer et al. [10]), denoted here by Tukey, described again by two shape parameters λ_3 and λ_4 (and with $\lambda_1 = 0$ and $\lambda_2 = 1$). Besides the alternatives described above we took several others which have been considered in Inglot et al. [17] or Józefczyk [21]. Below, we describe them for convenience of the reader, dividing into three groups according to a structure of asymmetry.

Let $\chi^2_k(x)$ denote the density of the chi-square distribution with k degrees of freedom, $\beta_{(\xi,\eta)}(x)$, $\xi, \eta > 0$, the density of the beta distribution, $\phi(x)$ the standard normal density function, and U a random variable uniformly distributed over $[0, 1]$.

Additionally, define the density function $\mathrm{en}(x)$ by the formula

$$\mathrm{en}(x) = c[\phi(x+1)\mathbf{1}_{(-\infty,-1)}(x) + \phi(0)\mathbf{1}_{[-1,1]}(x) + \phi(x-1)\mathbf{1}_{(1,\infty)}(x)]$$

with $c = \big(1 + 2\phi(0)\big)^{-1}$. Moreover, ρ is a parameter belonging to $[0,1]$ while in the two cases the letter m denotes the median of the underlying beta distribution.

- Alternatives with dominating asymmetry in the tails:

Notation	**Description of a random variable or a density**
Lambda(λ_3, λ_4)	$X = \mathrm{sgn}(\lambda_3)[U^{\lambda_3} - (1-U)^{\lambda_4}],\ \ \lambda_3 \cdot \lambda_4 > 0;$
Tukey(λ_3, λ_4)	$X = (U^{\lambda_3} - 1)/\lambda_3 - \big((1-U)^{\lambda_4} - 1\big)/\lambda_4,\ \ \lambda_3, \lambda_4 > 0;$
Ra(θ)	$f(x) = \theta^{-2} x \exp\{-x^2/2\theta^2\},\ \ x \geqslant 0,\ \theta > 0;$
ChiSq(θ)	$f(x) = \chi^2_\theta(x),\ \ x \in \mathbb{R},\ \theta = 4, 5, \ldots;$
N-Fechner(θ)	$f(x) = \phi\big(x/(1+\theta)\big)\mathbf{1}_{(-\infty,0]}(x) + \phi\big(x/(1-\theta)\big)\mathbf{1}_{(0,\infty)}(x),$ $x \in \mathbb{R},\ \theta \in (-1,1);$
EV(θ)	$f(x) = \exp\{(x-\theta) - \exp(x-\theta)\},\ \ x \in \mathbb{R},\ \theta \in \mathbb{R};$
Beta($2, \theta$)	$f(x) = \beta_{(2,\theta)}(x),\ \ x \in [0,1],\ \theta > 1;$
F(θ)	$f(x) = 0.5 + 2x\theta^{-2}(\theta - \lvert x\rvert)\mathbf{1}_{(\lvert x\rvert<\theta)},\ \ x \in [-1,1],\ \theta \in [0,1].$

- Alternatives with asymmetry in the tails and in the center:

Notation	**Density**
LC(ρ, θ)	$f(x) = \rho\phi(x - \theta/\rho) + (1-\rho)\phi\big(x + \theta/(1-\rho)\big),\ \ x \in \mathbb{R},\ \theta > 0;$
NB3(ρ, θ)	$f(x) = \rho\phi(x) + (1-\rho)\beta_{(3,3)}(x+\theta),\ \ x \in \mathbb{R},\ \theta \in [0,1];$
NC(ρ, θ)	$f(x) = \rho\phi(x) + (1-\rho)\big(1/\big[\pi\big(1 + (x+\theta)^2\big)\big]\big),\ \ x \in \mathbb{R},\ \theta \in \mathbb{R};$
Sin(θ, j)	$f(x) = 0.5 + \theta\sin(\pi j x),\ \ x \in [-1,1],\ \theta \in [-0.5, 0.5],\ j \geqslant 1;$
BiBeta(θ)	$f(x) = 0.5\,\big(\beta_{(2,\theta)}(x-1) + \beta_{(2,2)}(x)\big),\ \ x \in [0,2],\ \theta > 1;$
BiChiSq(θ)	$f(x) = 0.5\,\big(\chi^2_\theta(-x) + \chi^2_6(x)\big),\ \ x \in \mathbb{R},\ \theta = 2, 3, \ldots$

- Alternatives with asymmetry only in the center:

Notation	**Density**
NB(ρ, θ)	$f(x) = \rho\phi(x) + (1-\rho)\beta_{(2,\theta)}(x+m),\ \ x \in \mathbb{R},\ \theta > 1;$
ENB(ρ, θ)	$f(x) = \rho\,\mathrm{en}(x) + (1-\rho)\beta_{(2,\theta)}\big(x + \frac{1}{2}\big),\ \ x \in \mathbb{R},\ \theta > 1;$
MB(ρ, θ)	$f(x) = \frac{\rho}{2}\beta_{(2,3)}(x-1) + \frac{\rho}{2}\beta_{(3,2)}(x) + (1-\rho)\beta_{(2,\theta)}(x+m-1),$ $x \in [0,2],\ \theta > 1.$

Note that the symmetric part of Beta$(2,\theta)$ has finite Fisher information for $\theta > 2$ while the symmetric parts of F(θ) and Sin(θ, j) are the uniform distribution on $[-1,1]$ strongly violating (A1). Moreover, the alternatives Beta$(2,\theta)$, NB(ρ,θ), MB(ρ,θ) and BiBeta(θ), $1 < \theta \leqslant 2$, have continuous densities with infinite Fisher information as Beta$(2,2)$ does (cf. the remark in the previous subsection concerning Beta$(2,2)$). As was said previously, Tukey(λ_3, λ_4) has finite Fisher information if both positive parameters are less than 1 (continuous density) and at least one of them is less than 1/2. A similar property takes place for Lambda(λ_3, λ_4) with positive parameters, but for negative parameters Fisher information is always finite.

Finally, note that densities of BiBeta(θ) and BiChiSq(θ) take value zero at $x = 0$, which in effect causes poor null behaviour of all tests under consideration.

TABLE 2. Comparison of powers and average powers (in %) of WS, WL, CM, $Mira$, EJ, NS and MG. $\alpha = 0.05$, $n = 100$, $d(100) = 10$; 10,000 MC runs. The Lambda family

Type	λ_3	λ_4	WS	WL	CM	$Mira$	EJ	NS	MG
7	1.4	0.25	91	92	57	46	50	100	100
8	0.00007	0.1	100	100	97	96	82	100	100
9	0.025213	0.094029	61	52	55	53	44	74	85
10	−0.0075	−0.03	83	76	77	77	60	89	96
11	−0.13	−0.16	7	6	8	8	11	9	10
12	−0.1	−0.18	32	24	38	39	31	39	49
13	−0.001	−0.13	100	100	100	100	89	100	100
14	−0.0001	−0.17	100	100	100	100	89	100	100
Average power			**71.8**	**68.8**	**66.5**	**64.9**	**57.0**	**76.4**	**80.0**

In Table 2 we present powers of the tests WS, WL, CM, $Mira$, EJ and also NS and MG (for which data have been centered by the true median) for some alternatives from the Lambda family as considered, e.g., in Cabilio and Masaro [3]. It can be observed that in almost all cases CM dominates $Mira$ and EJ. So, in the next tables we shall restrict ourselves only to comparison of the new tests with CM and with NS and MG. It is also easily seen that WS behaves better than CM for the considered cases from the Lambda family. On average, WS loses ca. 8% in power to actually the best test (when a true median is known) which is MG. Moreover, the loss of WL with respect to WS is about 3%.

In Table 3 we show the powers of WS, WL, CM and also NS and MG for the alternatives with dominating asymmetry in the tails which have been described above.

TABLE 3. Comparison of empirical powers (in %) of WS, WL, CM, NS and MG. $\alpha = 0.05$, $n = 100$, $d(100) = 10$; 10,000 MC runs. Asymmetry in the tails

Distribution	WS	WL	CM	NS	MG
Tukey(10, 0.9)	42	38	19	49	67
Tukey(4, 6.5)	65	57	64	70	80
Lambda(0.025213, 0.094029)	61	52	55	74	85
Lambda(−0.1, −0.18)	32	24	38	39	49
Ra(2)	49	42	32	56	74
ChiSq(9)	65	58	52	75	89
N-Fechner(0.4)	42	33	34	54	68
EV(0.6)	66	57	56	72	69
Beta(2, 1.2)	69	64	35	85	94
Average power	**54.7**	**47.2**	**42.8**	**63.7**	**75.0**
F(0.4)	92	88	80	82	66

As could be expected, MG detects the alternatives from Table 3 with very high powers. Moreover, we can observe that WS and WL lose on average with respect to NS ca. 9% and 16%, respectively. The results for F(0.4), presented in the last row, are not reliable since sizes of the new tests as well as CM for symmetric part of this alternative are equal to 0.21 and 0.12, respectively, and are far from the nominal level 0.05. However, if one transforms X_i's onto the real line by taking $Y_i = \text{tg}\,\frac{\pi}{2}X_i$, then a character of asymmetry preserves (the function a remains unchanged), powers of WS and WL become 63 and 68, respectively, while the size decreases to the nominal 0.05. It corresponds well to powers of NS and MG for this alternative and to the rest of Table 3.

The results for the alternatives from the second group are shown in Table 4. For these alternatives the data driven tests provide superior power. WS and WL lose on average with respect to NS ca. 11% and 14%, respectively. Contrary to CM (cf. NC(0.4, 3)), WS and WL do not have weak points. Moreover, note that here MG is distinctly worse than our new data driven tests although there is a need of estimation of an unknown median.

TABLE 4. Comparison of empirical powers (in %) of WS, WL, CM, NS and MG. $\alpha = 0.05$, $n = 100$, $d(100) = 10$; 10,000 MC runs. Asymmetry in the tails and in the center

Distribution	WS	WL	CM	NS	MG
LC(0.7, 0.6)	70	62	66	81	69
NB3(0.8, 0.25)	47	40	57	67	63
NC(0.4, 3)	80	85	14	81	43
Average power	**65.7**	**62.3**	**45.7**	**76.3**	**58.3**
Sin(0.5, 3)	59	71	3	100	56
Sin(0.5, 8)	39	56	16	44	44
BiBeta(2)	10	10	75	5	4
BiChiSq(6)	9	9	72	5	4
BiBeta(8)	68	62	67	100	95
BiChiSq(2)	64	62	66	100	99

Results for the alternative Sin(θ, j), similarly to those for F(θ), seem to be not reliable. A different case of bimodal distributions is presented in the last four rows of Table 4. The last two rows show powers while the third and fourth from below show sizes for the two considered alternatives. Here large "powers" of CM correspond to even greater "sizes" of this test illustrating that CM is helpless in distinguishing between null and alternative hypotheses when $f_s(0)$ is close to zero. But both data driven tests behave quite reasonable, also in comparison with NS and MG.

TABLE 5. Comparison of empirical powers (in %) of WS, WL, CM, NS and MG. $\alpha = 0.05$, $n = 100$, $d(100) = 10$; 10,000 MC runs. Asymmetry only in the center

Distribution	WS	WL	CM	NS	MG
NB(0.15, 5)	28	34	11	27	14
ENB(0.4, 1.2)	28	27	24	66	23
MB(0.2, 6)	27	28	14	27	11
Average power	**27.7**	**29.7**	**16.3**	**40.0**	**16.0**

Finally, in Table 5 we present the powers for the alternatives with asymmetry only in the center. These alternatives are hard to detect for all tests but data driven tests become the best ones. CM is much worse and sometimes breaks down (cf. NB$(0.15, 5)$).

5. DISCUSSION, CONCLUSIONS AND EXAMPLE

For better insight into an ability of compared tests to detect various types of alternatives we calculated an average power of compared tests for twelve alternatives, six from Table 3 (from the third to eighth rows) and six from Tables 4 and 5, all satisfying the assumption (A1), and for which a value of the corresponding null density at zero is far from zero. We have obtained 49.6% for WS, 45.2% for WL, 37.8% for CM, 59.9% for NS, and 54.8% for MG. This together with Table 2 illustrates that WS and WL behave comparably well. A more difficult testing problem, they have to face, leads to a quite reasonable loss (ca. 10–14%) in power in comparison with the case when nothing has to be estimated (NS). Simultaneously, all test procedures taken into account, and especially CM, are significantly worse, both on average and for most individual alternatives. Moreover, the new data driven tests keep stable power in each case under the model and are relatively robust when violating assumptions. The results presented in Section 4 suggest to apply WS when we expect asymmetry on tails. In the cases when tails are rather symmetric we recommend to apply WL.

To illustrate how new tests apply to an analysis of real data we used that from Table II in Doksum [4] presenting survival times (in days) of 72 guinea pigs that received a dose of tubercle bacilli. For these data the statistics $\widehat{W}_S$ and $\widehat{W}_L$ take the same value 10.754. Empirical critical values, calculated in the same way as described in Subsection 4.1, for the sample size $n = 72$ are equal to 6.821 and 8.897, respectively. Hence both tests reject the null hypothesis of symmetry. Since the symmetric part of the density under consideration is unimodal and smooth (cf. Fig. 5.1 in Ghosh [12], confirming such a claim), we simulated p-values using the standard normal distribution, as a null distribution, obtaining 0.017 and 0.030, respectively.

6. PROOFS

In this section we present proofs of all statements formulated in Sections 2 and 3.

Hints for the proof of Proposition 2.1. A standard argument shows that under (A1) and (A2) the score vector for the family $\mathcal{G}_k$ has the following components:

$$\ell_\vartheta = \Psi\big(f_s(\cdot - \mu)\big), \quad \ell_\mu = -\frac{f_s'}{f_s}(\cdot - \mu), \quad \ell_{f_s} = \frac{2}{\sqrt{f_s(\cdot - \mu)}}.$$

The component ℓ_{f_s} is the superposition of the translation by μ and the operator of multiplication by $2/\sqrt{f_s}$ on the domain of symmetric functions in $L_2(P_{\mu f_s})$ orthogonal to $1/\sqrt{f_s}$. The component ℓ_ϑ is antisymmetric about μ, hence orthogonal to the range of ℓ_{f_s} consisting of symmetric functions about μ. Thus ℓ_*, which is the residue of the orthogonal projection of ℓ_ϑ onto the subspace spanned by the remaining components, has the form (2.2).

P r o o f o f P r o p o s i t i o n 2.2. The proof goes similarly to that of Theorem 1 in Inglot and Ledwina [19]. First we introduce an auxiliary notation. Consider a set

$$\Upsilon = \{(\nu_n) : \nu_n = \mu + t_n/\sqrt{n},\ (t_n) \text{ bounded}\}$$

of deterministic sequences corresponding to realizations of discretized estimator $\widehat{\mu}_d$ of μ. For $(\nu_n) \in \Upsilon$ put $Z_{\nu_n} = (X_1 - \nu_n, \ldots, X_n - \nu_n, \nu_n - X_1, \ldots, \nu_n - X_n)$ and denote by $\mathcal{F}_{\nu_n ns}$ the empirical distribution function of Z_{ν_n}. In particular, $Z = Z_\mu$, i.e., it corresponds to a constant sequence $\nu_n = \mu$. Moreover, denote by $\widehat{v}_{\nu_n}$, $\widehat{J}_{\nu_n}$ and $(\widehat{f_s'/f_s})_{\nu_n j}$, $j = 1, 2$, the estimators applied in (2.4) but based on Z_{ν_n} instead of on $\widehat{Z}$, and denote by $\widehat{\ell}_{*\nu_n j}$, $j = 1, 2$, the estimators of ℓ_* given by (2.4) but based again on two parts of Z_{ν_n}. Finally, let us put

$$\ell_{*\nu_n} = \Psi\big(F_s(\cdot - \nu_n)\big) + \frac{v}{J}\frac{f_s'}{f_s}(\cdot - \nu_n) \text{ a.e.},$$

in which nothing is estimated and only the true median μ is replaced by an arbitrary $(\nu_n) \in \Upsilon$.

Observe that

$$\begin{aligned}(6.1)\quad \overline{\widehat{\ell}_*} - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \ell_*(X_i) \\ = \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor} \big(\widehat{\ell}_{*2}(X_i) - \ell_*(X_i)\big) + \frac{1}{\sqrt{n}}\sum_{i=\lfloor n/2\rfloor+1}^{n} \big(\widehat{\ell}_{*1}(X_i) - \ell_*(X_i)\big).\end{aligned}$$

By an exchangeability argument it is enough to show that one, say the first, term on the right-hand side of (6.1) converges to zero in probability. Arguing as in Theorem 2, p. 44, in Bickel et al. [2] it is enough to do this replacing $\widehat{\ell}_{*2}$ by $\widehat{\ell}_{*\nu_n 2}$ with arbitrary $(\nu_n) \in \Upsilon$. To this end take any $(\nu_n) \in \Upsilon$ and consider the following decomposition:

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor} \big(\widehat{\ell}_{*\nu_n 2}(X_i) - \ell_*(X_i)\big) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor} \big(\ell_{*\nu_n}(X_i) - \ell_*(X_i)\big) \\ &+ \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor} \big(\widehat{\ell}_{*\nu_n 2}(X_i) - \ell_{*\nu_n}(X_i)\big) = \mathcal{R}_1 + \mathcal{R}_2.\end{aligned}$$

The convergence of $\mathcal{R}_1$ in probability $P_{\mu f_s}$ to zero follows immediately from Theorem 2.3 of Schick [27] by setting $f_s'/f_s(\cdot-\mu)$ as his κ_μ and $\ell_*(\cdot-\mu)$ as his h_μ. Here we use the orthogonality relation

$$\int \ell_*(x-\mu)\frac{f_s'}{f_s}(x-\mu)f_s(x-\mu)dx = 0$$

for f_s'/f_s and ℓ_*, which is an obvious consequence of the definition of the efficient score vector ℓ^*. Since the sequences $(P^n_{\mu f_s})$ and $(P^n_{\nu_n f_s})$ of distributions are mutually contiguous, it is enough to prove that $\mathcal{R}_2$ converges to zero under $P_{\nu_n f_s}$. But then we can remove in all places an unnecessary deterministic shift $t_n/\sqrt{n}$ and get the equivalent condition

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor}\left(\widehat{\ell}_{*\mu 2}(X_i)-\ell_*(X_i)\right)\xrightarrow{P_{\mu f_s}} 0. \tag{6.2}$$

Now, by (2.2) and (2.4) the left-hand side of (6.2) can be written as

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor}\left[\Psi\big(\mathcal{F}_{ns}(X_i-\mu)\big)-\Psi\big(F_s(X_i-\mu)\big)\right]+\left[\frac{\overline{v}}{\overline{J}}-\frac{v}{J}\right]\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor}\frac{f_s'}{f_s}(X_i-\mu)$$
$$+\frac{\overline{v}}{\overline{J}}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor n/2\rfloor}\left[\left(\overline{\frac{f_s'}{f_s}}\right)_2(X_i-\mu)-\frac{f_s'}{f_s}(X_i-\mu)\right]\right)=\mathcal{R}_3+\mathcal{R}_4+\mathcal{R}_5.$$

We apply Proposition A.2 of Inglot et al. [17] to $\mathcal{R}_3$ after observing that $\mathcal{R}_3$ equals $\big(1/\sqrt{2}+o(1)\big)(\widehat{\Phi}-\widehat{\Phi}_0)$ in the notation of that paper. So, this term tends to zero in probability. The remainder term $\mathcal{R}_4$ tends to zero due to (A5) and boundedness in $L_2(P_{\mu f_s})$ of the second factor. The first factor in $\mathcal{R}_5$ is bounded in probability again by (A5). The conditional expectation (under $P_{\mu f_s}$), with respect to the second part of Z, of the second factor in $\mathcal{R}_5$ equals zero since by (A4) both f_s'/f_s and its estimator are odd functions. In effect, the conditional variance (under $P_{\mu f_s}$), with respect to the second part of Z, of the second factor in $\mathcal{R}_5$ is equal to

$$\big(1/2+o(1)\big)\int_{\mathbb{R}}\left(\overline{\frac{f_s'}{f_s}}(x)-\frac{f_s'(x)}{f_s(x)}\right)^2 f_s(x)\,dx,$$

which tends to zero in probability by (A4). This completes the proof. ■

Proof of Proposition 3.1. The proof goes similarly to that of Lemma A.2 in Inglot and Ledwina [20]. By symmetry of the kernel K and symmetry of Z it follows that $\overline{f_s}$ is an even function, which by (3.3) implies that $\overline{f_s'/f_s}$ is an odd function. To prove the second condition in (A4) introduce auxiliary kernel estimators

$$\widetilde{f}_s(x)=\gamma_n+\frac{1}{2nh_n}\sum_{i=1}^{2n}K\left(\frac{x-Z_i}{h_n}\right),\qquad \widetilde{f}_s'(x)=\frac{1}{2nh_n^2}\sum_{i=1}^{2n}K'\left(\frac{x-Z_i}{h_n}\right)$$

and $\widetilde{f'_s/f_s} = \widetilde{f}'_s/\widetilde{f}_s$, where γ_n and h_n are as in (E3). From Section 8 of Forrester et al. [9] and (E3) it follows that

$$\int_{\mathbb{R}} \left(\widetilde{\frac{f'_s}{f_s}}(x) - \frac{f'_s(x)}{f_s(x)} \right)^2 f_s(x)\,dx \xrightarrow{P_{\mu f_s}} 0 \quad \text{as } n \to \infty. \tag{6.3}$$

Observe that $\overline{h}_n|\overline{f'_s}(x)| \leqslant C\overline{f_s}(x)$ due to (E2) and $\widetilde{f}_s(x) \geqslant \gamma_n$. Hence, by (3.3),

$$\left| \frac{\overline{f'_s}}{\overline{f_s}} - \widetilde{\frac{f'_s}{f_s}} \right| \leqslant \left| \frac{\overline{f'_s}}{\overline{f_s}} \right| \frac{|\overline{f_s} - \widetilde{f}_s|}{\widetilde{f}_s} + \frac{|\overline{f'_s} - \widetilde{f}'_s|}{\widetilde{f}_s} \leqslant C \frac{|\overline{f_s} - \widetilde{f}_s|}{\gamma_n \overline{h}_n} + \frac{|\overline{f'_s} - \widetilde{f}'_s|}{\gamma_n}. \tag{6.4}$$

From (E2) and (E4) for some ξ_i's between h_n and $\overline{h}_n$ we have

$$\begin{aligned} \overline{f_s}(x) - \widetilde{f}_s(x) &= \frac{h_n - \overline{h}_n}{2nh_n\overline{h}_n} \sum_{i=1}^{2n} K\left(\frac{x - Z_i}{\overline{h}_n} \right) \\ &+ \frac{1}{2nh_n} \sum_{i=1}^{2n} K'\left(\frac{x - Z_i}{\xi_i} \right) \frac{x - Z_i}{\xi_i} \frac{\xi_i(h_n - \overline{h}_n)}{h_n\overline{h}_n} = O_{P_{\mu f_s}}\left(\frac{1}{h_n\sqrt{n}} \right) \end{aligned} \tag{6.5}$$

uniformly in x. Similarly we obtain

$$|\overline{f'_s}(x) - \widetilde{f}'_s(x)| = O_{P_{\mu f_s}}\left(\frac{1}{h_n^2\sqrt{n}} \right) \tag{6.6}$$

uniformly in x. Now, (6.4)–(6.6) imply

$$\int_{\mathbb{R}} \left(\frac{\overline{f'_s}}{\overline{f_s}}(x) - \widetilde{\frac{f'_s}{f_s}}(x) \right)^2 f_s(x)\,dx = O_{P_{\mu f_s}}\left(\frac{1}{n\gamma_n^2 h_n^4} \right).$$

By (E3) and (6.3) the condition in (A4) follows. ∎

P r o o f o f P r o p o s i t i o n 3.2. First observe that $\overline{f_s}(Z_{(1)}) + \overline{f_s}(Z_{(2n)}) \to 0$ in probability. Hence, by the definition of $\overline{v}$ it is enough to prove consistency of $\widetilde{v}$. To this end write

$$v - \widetilde{v} = \int_{\mathbb{R}} \big[b\big(\mathcal{F}_{ns}(x)\big) - b\big(F_s(x)\big) \big] \overline{f'_s}(x)dx + \int_{\mathbb{R}} b\big(F_s(x)\big) [\overline{f'_s}(x) - f'_s(x)]dx. \tag{6.7}$$

Since b is a vector of Lipschitz functions, the Euclidean norm of the first integral in (6.7) can be estimated by

$$\frac{C}{\overline{h}_n} |b'(1)|_k \sup_{x \in \mathbb{R}} |\mathcal{F}_{ns}(x) - F_s(x)|,$$

where $|y|_k = (y_1^2 + \ldots + y_k^2)^{1/2}$ is the Euclidean norm of the vector y, which clearly goes to zero in probability by (E3) and (E4). Calculating the second integral in (6.7) by parts we see that it can be replaced by

$$(6.8)\quad -\int_{\mathbb{R}} b'\big(F_s(x)\big)[\overline{f_s}(x) - f_s(x)]f_s(x)dx$$
$$= -\int_{\mathbb{R}} b'\big(F_s(x)\big)[\overline{f_s}(x) - \widetilde{f}_s(x)]f_s(x)dx - \int_{\mathbb{R}} b'\big(F_s(x)\big)[\widetilde{f}_s(x) - f_s(x)]f_s(x)dx$$
$$= \mathcal{R}_6 + \mathcal{R}_7.$$

Recall that b' is a vector of bounded functions. Thus, $\mathcal{R}_6$ goes to zero in probability by (6.5) and (E3). By the Schwarz inequality, $(\mathcal{R}_7)^2$ can be estimated by $|b'(1)|_k^2 \int_{\mathbb{R}}[\widetilde{f}_s(x) - f_s(x)]^2 f_s(x)dx$ which by (E3) tends to zero in probability. This proves the consistency of $\widetilde{v}$. ■

Proof of Proposition 3.3. From (6.4)–(6.6) we get

$$\left|\left(\frac{\overline{f_s'}}{f_s}\right)^2 - \left(\frac{\widetilde{f_s'}}{f_s}\right)^2\right| = \left|\frac{\overline{f_s'}}{f_s} + \frac{\widetilde{f_s'}}{f_s}\right|\left|\frac{\overline{f_s'}}{f_s} - \frac{\widetilde{f_s'}}{f_s}\right| = O_{P_{\mu f_s}}\left(\frac{1}{\sqrt{n}\gamma_n h_n^3}\right)$$

uniformly in x, which by (E3) tends to zero in probability. Since by the law of large numbers we have

$$\frac{1}{2n}\sum_{i=1}^{2n}\left(\frac{f_s'}{f_s}\right)^2(Z_i) \xrightarrow{P_{\mu f_s}} J,$$

to prove the consistency of $\overline{J}$, it is enough to check

$$(6.9)\qquad \frac{1}{2n}\sum_{i=1}^{2n}\left[\left(\frac{\widetilde{f_s'}}{f_s}\right)^2(Z_i) - \left(\frac{f_s'}{f_s}\right)^2(Z_i)\right] \xrightarrow{P_{\mu f_s}} 0.$$

The summands in (6.9) have the same distribution. Therefore, (6.9) will follow if we show that

$$E\left|\left(\frac{\widetilde{f_s'}}{f_s}\right)^2(Z_1) - \left(\frac{f_s'}{f_s}\right)^2(Z_1)\right| \to 0.$$

Since $E(f_s'/f_s)^2(Z_1) = J$ does not depend on n, a routine argument shows that it is enough to prove

$$E\left(\frac{\widetilde{f_s'}}{f_s}(Z_1) - \frac{f_s'}{f_s}(Z_1)\right)^2 \to 0.$$

Taking the conditional expectation with respect to Z_1, we see that the last condition reduces to

$$(6.10)\qquad E\left(\int_{\mathbb{R}}\left(\frac{\widetilde{f}_s'(x) + (2nh_n^2)^{-1}L_1(x)}{\widetilde{f}_s(x) + (2nh_n)^{-1}L(x)} - \frac{f_s'(x)}{f_s(x)}\right)^2 f_s(x)\,dx\right) \to 0,$$

where $L_1(x) = K'(0) + K'\big((2x)/h_n\big) - K'\big((x - Z_1)/h_n\big) - K'\big((x + Z_1)/h_n\big)$ and $L(x) = K(0) + K\big((2x)/h_n\big) - K\big((x - Z_1)/h_n\big) - K\big((x + Z_1)/h_n\big)$ are random variables which are bounded, uniformly in x, by $4C^2$. Consequently, due to (E3), (6.10) follows from (6.3), and the proof is complete. ■

REFERENCES

[1] P. K. Bhattacharya, J. L. Gastwirth, and A. L. Wright, *Two modified Wilcoxon tests for symmetry about an unknown location parameter*, Biometrika 69 (1982), pp. 377–382.

[2] P. J. Bickel, C. A. Klaassen, Y. Ritov, and J. A. Wellner, *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins Univ. Press, Baltimore 1993.

[3] P. Cabilio and J. Masaro, *A simple test of symmetry about an unknown median*, Canad. J. Statist. 24 (1996), pp. 349–361.

[4] K. A. Doksum, *Measures of location and asymmetry*, Scand. J. Statist. 2 (1975), pp. 11–22.

[5] K. A. Doksum, G. Fenstad, and R. Aaberge, *Plots and tests for symmetry*, Biometrika 64 (1977), pp. 473–487.

[6] M. Ekström and S. R. Jammalamadaka, *An asymptotically distribution-free test of symmetry*, J. Statist. Plann. Inference 137 (2007), pp. 799–810.

[7] M. Ekström and S. R. Jammalamadaka, *A general measure of skewness*, Statist. Probab. Lett. 82 (2012), pp. 1559–1568.

[8] M. Fernandes, E. F. Mendes, and O. Scaillet, *Testing for symmetry and conditional symmetry using asymmetric kernels*, HEC Genève DP 2011.01 and Swiss Finance Institute DP 2011.32, 2011.

[9] J. Forrester, W. Hooper, H. Peng, and A. Schick, *On the construction of efficient estimators in semiparametric models*, Statist. Decisions 21 (2003), pp. 109–138.

[10] M. Freimer, G. Kollia, G. S. Mudholkar, and C. T. Lin, *A study of the generalized Tukey lambda family*, Comm. Statist. Theory Methods 17 (1988), pp. 3547–3567.

[11] J. L. Gastwirth, *On the sign test of symmetry*, J. Amer. Math. Soc. 66 (1971), pp. 821–823.

[12] K. Ghosh, *A new nonparametric test of symmetry*, in: *Advances in Directional and Linear Statistics*, M. T. Wells and A. SenGupta (Eds.), Physica-Verlag HD, Springer, Berlin–Heidelberg 2011, pp. 69–83.

[13] M. K. Gupta, *An asymptotically nonparametric test of symmetry*, Ann. Math. Statist. 38 (1967), pp. 849–866.

[14] H. E. T. Holgerson, *A modified skewness measure for testing asymmetry*, Comm. Statist. Simulation Comput. 39 (2010), pp. 335–346.

[15] T. Inglot and A. Janic, *How powerful are data driven score tests for uniformity*, Appl. Math. (Warsaw) 36 (2009), pp. 375–395.

[16] T. Inglot and A. Janic, *Data driven tests for univariate symmetry about unknown median*, Technical Report, Institute of Mathematics and Computer Science, Wrocław University of Technology, 2014.

[17] T. Inglot, A. Janic, and J. Józefczyk, *Data driven tests for univariate symmetry*, Probab. Math. Statist. 32 (2) (2012), pp. 323–358.

[18] T. Inglot, A. Janic, and J. Józefczyk, *Data driven tests for univariate symmetry about unknown median. A simulation study*, Technical Report No. I-18/2013/P-049, Institute of Mathematics and Computer Science, Wrocław University of Technology, 2013.

[19] T. Inglot and T. Ledwina, *Data-driven score tests for homoscedastic linear regression model: Asymptotic results*, Probab. Math. Statist. 26 (1) (2006), pp. 41–61.

[20] T. Inglot and T. Ledwina, *Data driven score tests for homoscedastic linear regression model*, Preprint 665, Institute of Mathematics of the Polish Academy of Sciences, 2006.

[21] J. Józefczyk, *Data driven score tests for univariate symmetry based on non-smooth functions*, Probab. Math. Statist. 32 (2) (2012), pp. 301–322.

[22] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer, 2005.
[23] A. Mira, *Distribution-free test for symmetry based on Bonferroni's measure*, J. Appl. Stat. 26 (1999), pp. 959–972.
[24] R. Modarres and J. L. Gastwirth, *Hybrid test for the hypothesis of symmetry*, J. Appl. Stat. 25 (1998), pp. 777–783.
[25] R. H. Randles, M. A. Flinger, G. E. Policello II, and D. A. Wolfe, *An asymptotically distribution-free test for symmetry versus asymmetry*, J. Amer. Statist. Assoc. 75 (1980), pp. 168–172.
[26] G. Sansone, *Orthogonal Functions*, Interscience, New York 1959.
[27] A. Schick, *On asymptotic differentiability of averages*, Statist. Probab. Lett. 51 (2001), pp. 15–23.
[28] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. 6 (1978), pp. 461–464.
[29] T. Zheng and J. L. Gastwirth, *On bootstrap tests of symmetry about an unknown median*, J. Data Sci. 8 (2010), pp. 397–412.

Tadeusz Inglot
Institute of Mathematics and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
54-404 Wrocław, Poland
E-mail: Tadeusz.Inglot@pwr.edu.pl

Alicja Janic
Institute of Mathematics and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
54-404 Wrocław, Poland
E-mail: Alicja.Janic@pwr.edu.pl

Received on 15.3.2014;
revised version on 19.5.2014