Optimal Rank Tests for Symmetry against Edgeworth-type Alternatives

Delphine Cassart, Marc Hallin and Davy Paindaveine

Abstract We are constructing, for the problem of univariate symmetry (with respect to specified or *un*specified location), a class of signed-rank tests achieving optimality against the family of asymmetric (local) alternatives considered in Hallin et al. (2011). Those alternatives are based on a non-Gaussian generalization of classical first-order Edgeworth expansions indexed by a measure of skewness such that *(i)* location, scale and skewness play well-separated roles (diagonality of the corresponding information matrices), and *(ii)* the classical tests based on the Pearson-Fisher coefficient of skewness are optimal in the vicinity of Gaussian densities. Asymptotic distributions are derived under the null and under local alternatives. Asymptotic relative efficiencies are computed and, in most cases, indicate that the proposed rank tests significantly outperform their traditional competitors.

1 Introduction

The assumption of symmetry is among the most important and fundamental ones in statistics. This importance explains the variety of existing parametric as well as nonparametric testing procedures of the null hypothesis of symmetry in an i.i.d. sample $X_1, \ldots, X_n$; see Hollander (1988) for a classical survey.

Traditional parametric tests of the null hypothesis of symmetry—the hypothesis under which $X_1 - \theta \stackrel{d}{=} -(X_1 - \theta)$ for some location (automatically, the population

Delphine Cassart
ECARES, Université Libre de Bruxelles, 1050 Bruxelles, Belgium, e-mail: dcassart@gmail.com

Marc Hallin
ECARES, Université Libre de Bruxelles, 1050 Bruxelles, Belgium and ORFE, Princeton University, Princeton, NJ 08544, USA, e-mail: mhallin@ulb.ac.be

Davy Paindaveine
ECARES and Department of Mathematics, Université Libre de Bruxelles, 1050 Bruxelles, Belgium, e-mail: dpaindav@ulb.ac.be

median) $\theta \in \mathbb{R}$, where $\stackrel{d}{=}$ stands for equality in distribution—are based on third-order moments. Write $m_k^{(n)} := m_k^{(n)}(\bar{X}^{(n)})$ for the sample moment of order k, where

$$m_k^{(n)}(\theta) := n^{-1} \sum_{i=1}^{n} (X_i - \theta)^k$$

and $\bar{X}^{(n)} := n^{-1} \sum_{i=1}^{n} X_i$. When the location θ is specified, the classical test statistic is

$$S_1^{(n)} := n^{1/2} m_3^{(n)}(\theta) / (m_6^{(n)}(\theta))^{1/2}, \tag{1}$$

with, under finite sixth-order moments, asymptotically standard normal null distribution. When θ is unspecified, the classical test is based on the empirical coefficient of skewness $b_1^{(n)} := m_3^{(n)}/s_n^3$, where $s_n := (m_2^{(n)})^{1/2}$ stands for the empirical standard error in a sample of size n. More precisely, that test relies on the asymptotic standard normal distribution (still under finite moments of order six) of

$$S_2^{(n)} := n^{1/2} m_3^{(n)} / (m_6^{(n)} - 6 s_n^2 m_4^{(n)} + 9 s_n^6)^{1/2}. \tag{2}$$

This test is generally considered a Gaussian test. Cassart et al. (2011) indeed show that it is locally asymptotically optimal, in the Le Cam sense, for the null hypothesis of i.i.d. Gaussian observations with unspecified location and scale, against asymmetric alternatives described by a first-order Edgeworth expansion of the form

$$\phi(x-\theta) + n^{-1/2} \xi (x-\theta) \phi(x-\theta)((x-\theta)^2 - \kappa), \tag{3}$$

where κ(=3) is the Gaussian kurtosis coefficient, θ a location parameter, and ξ a measure of skewness; see Chapter XVI of Feller (1971) for a concise introduction to those expansions, the idea of which goes back to Edgeworth (1905). Cassart et al. (2011) further show that the local experiments associated with (3) enjoy the appealing property that location, scale, and skewness play well-separated roles, in the sense that the Fisher information matrix associated with the triple $(\theta,\ \sigma\ \xi)$ is diagonal.

Extending that Gaussian approach to a broad class of symmetric densities by embedding the null hypothesis of i.i.d.-ness with density

$$f:\ x \mapsto f(x) := \sigma^{-1} f_1((x-\theta)/\sigma)$$

(f_1 symmetric with respect to the origin) into a family of locally asymmetric alternatives based on an adequate generalization of (3), Cassart et al. (2011) show that those families (for fixed f) are uniformly locally asymptotically normal (ULAN) under mild assumptions on f_1. The resulting locally asymptotically optimal tests are derived and studied in detail.

It should be insisted, however, that considering alternatives of this (generalized) Edgeworth type by no means implies any assumption on the form of the asymmetry present in the data. Those local expansions are just considered as a way of producing non-Gaussian alternatives to the classical test (2). And simulations (see Cassart et al. 2011) indicate that the resulting tests perform quite well against other types of asymmetries, such as the skew-normal and skew-t densities investigated, for instance, by Azzalini and Capitanio (2003).

Yet, these tests all are of a parametric nature, while symmetry, very typically, is a nonparametric hypothesis, enjoying a rich group invariance structure, through which classical maximal invariance arguments naturally bring *signs* and *signed-ranks* into the picture.

The most popular nonparametric signed-rank tests of symmetry (with respect to any specified location—without loss of generality, the origin) are the *sign test*, based on the binomial distribution of the number of negative signs in a sample of size n, and the *Wilcoxon signed-rank test*, based on the exact or asymptotic null distribution of

$$S_W^{(n)} := n^{-1/2} \sum_{i=1}^{n} s_i R_{+,i}^{(n)}$$

(or a linear transformation thereof, such as $n^{-1/2} \sum_{i=1}^{n} I[s_i = 1] R_{+,i}^{(n)}$), where $s_1, \ldots, s_n$ denote the signs, and $R_{+,1}^{(n)}, \ldots, R_{+,n}^{(n)}$ the ranks of absolute values in a sample of size n. These tests are not optimal in any satisfactory sense against asymmetry: actually, they are locally asymptotically most powerful against symmetry-preserving location shifts—under otherwise unspecified density for the sign test, under logistic densities for the Wilcoxon one. Moreover, the sign test is completely insensitive to nonsymmetric alternatives preserving the median. There is no way such tests can be adapted to an unspecified-location context (*alignment* for the location nuisance here produces, at probability level α, a test with trivial asymptotic power α).

Another signed-rank test based on signs only is the runs test proposed by McWilliams (1990) and further investigated by Henze (2003). If not completely invariant, this test has low sensitivity against location shifts; however, it does not address any well-identified alternative and does not exploit the ranks themselves. The *triples test* by Randles et al. (1980) is location-invariant, and also based on signs. Those signs, though, are those of quantities of the form $X_i + X_j - 2X_k$, $1 \leq i < j \leq n$ and $i \neq j \neq k$, which do not follow from any concept of group invariance and are not distribution-free; optimality properties, if any, are unclear.

To the best of our knowledge, the problem of constructing optimal rank-based tests of symmetry only has been touched in Cassart et al. (2008) and Ley and Paindaveine (2009), which are both focusing on quite specific alternatives (Fechner and Ferreira-Steel types, respectively). The objective of this paper is to construct signed-rank tests that are optimal against the Edgeworth-type alternatives defined in Cassart et al. (2011). The proposed tests are distribution-free (asymptotically so in case of an unspecified location θ) under the null hypothesis of symmetry, and therefore remain valid under very mild distributional assumptions (for the specified-location case, they are valid in the absence of *any* distributional assumption).

For instance, the normal-score signed-rank test rejects the null hypothesis of symmetry with respect to (specified) θ for large values of

$$\underset{\sim}{T}^{(n)}_{\phi_1}(\theta) := \left(n\gamma^{(n)}_{\phi_1}\right)^{-1/2} \sum_{i=1}^{n} s_i(\theta)\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\theta)}{2(n+1)}\Big) \times\Big(\Big(\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\theta)}{2(n+1)}\Big)\Big)^2-3\Big),$$

where Φ denotes the standard normal distribution function, $s_i(\theta)$ is the sign of $Z_i(\theta) := X_i - \theta$, $R^{(n)}_{+,i}(\theta)$ the rank of $|Z_i(\theta)|$ among $|Z_1(\theta)|,\ldots,|Z_n(\theta)|$, and

$$\gamma^{(n)}_{\phi_1} := n^{-1}\sum_{r=1}^{n}\left(\Phi^{-1}\Big(\frac{n+1+r}{2(n+1)}\Big)\Big(\Big(\Phi^{-1}\Big(\frac{n+1+r}{2(n+1)}\Big)\Big)^2-3\Big)\right)^2$$

a standardizing constant. That test is distribution-free under the null hypothesis of symmetry with respect to θ, asymptotically equivalent to the test based on $b_1^{(n)}$ under Gaussian densities, and hence asymptotically most powerful against local Edgeworth alternatives of the form (3) with $\xi > 0$. And, under a very broad class of non-Gaussian densities (containing, among many others, all Student and power-exponential ones), the ARE (see Section 3.4) of this signed-rank test is strictly larger than one with respect to the traditional test based on $b_1^{(n)}$.

The problem we are considering throughout is that of testing the null hypothesis of symmetry. In the notation of Section 1, ξ (see (4) for a more precise definition) is thus the parameter of interest; the location θ and the standardized null symmetric density f_1 either are specified or play the role of nuisance parameters, whereas the scale σ (not necessarily a standard error) always is a nuisance.

The paper is organized as follows. In Section 2.1, we briefly describe the Edgeworth-type families of local alternatives we are considering. Section 2.2 restates the local asymptotic normality (with respect to location, scale, and asymmetry parameters) result of Cassart et al. (2011) that provides the main theoretical tool of the paper. Signed-rank versions of the corresponding central sequences for asymmetry are defined in Section 2.3. In Section 3.1, we propose nonparametric signed-rank (hence distribution-free) versions of the optimal procedures obtained in Cassart et al. (2011) for specified location θ. The case of unspecified θ is treated in Section 3.2, and requires the delicate estimation of a cross-information quantity of the same type as those appearing in the asymptotic variances of R-estimators (see Cassart et al. 2010). That estimation problem is discussed in some detail in Section 4. The van der Waerden, Wilcoxon and Laplace versions of the signed-rank statistics are described in Section 3.3, and Section 3.4 provides asymptotic relative efficiencies of signed-rank tests with respect to the classical ones based on (1) and (2), indicating the superiority of the former.

2 A class of locally asymptotically normal families of asymmetric distributions.

2.1 Families of asymmetric densities based on Edgeworth approximations.

Denote by $\boldsymbol{X}^{(n)} := (X_1^{(n)}, \ldots, X_n^{(n)})$, $n \in \mathbb{N}$ an i.i.d. n-tuple of observations with common density f. The null hypotheses we are interested in are

(a) the hypothesis $\mathscr{H}_\theta^{(n)}$ of symmetry with respect to specified location $\theta \in \mathbb{R}$: under $\mathscr{H}_\theta^{(n)}$, the X_i's have density function $x \mapsto f(x) := \sigma^{-1} f_1((x-\theta)/\sigma)$ (all densities are over the real line, with respect to the Lebesgue measure), for some unspecified $\sigma \in \mathbb{R}_0^+$ and f_1 in the class of standardized symmetric densities

$$\mathscr{F}_0 := \left\{ f_1 \ : \ f_1(-z) = f_1(z) \text{ and } \int_{-1}^{1} f_1(z)\, dz = 0.5 \right\}.$$

The scale parameter σ (associated with the symmetric density f) we are considering here thus is not the standard error, but the median of the absolute deviations $|X_i - \theta|$; this avoids making any moment assumptions on f;

(b) the hypothesis $\mathscr{H}^{(n)} := \bigcup_{\theta \in \mathbb{R}} \mathscr{H}_\theta^{(n)}$ of symmetry with respect to unspecified location.

As explained in the introduction, efficient testing requires the definition of families of asymmetric alternatives exhibiting some adequate structure, such as local asymptotic normality, at the null hypothesis of symmetry. For a selected class of densities f enjoying the required regularity assumptions, we therefore are embedding the null hypothesis of symmetry into families of distributions indexed by $\theta \in \mathbb{R}$ (location), $\sigma \in \mathbb{R}_0^+$ (scale), and a parameter $\xi \in \mathbb{R}$ characterizing asymmetry. More precisely, consider the class $\mathscr{F}_1$ of densities f_1 satisfying

(i) (symmetry and standardization) $f_1 \in \mathscr{F}_0$;

(ii) (absolute continuity) there exists $\dot{f}_1$ such that, for all $z_1 < z_2$,

$$f_1(z_2) - f_1(z_1) = \int_{z_1}^{z_2} \dot{f}_1(z)\, dz;$$

(iii) (strong unimodality) $z \mapsto \phi_{f_1}(z) := -\dot{f}_1(z)/f_1(z)$ is monotone increasing, or the difference of two monotone increasing functions, and

(iv) (finite Fisher information) the integral

$$\mathscr{K}(f_1) := \int_{-\infty}^{+\infty} z^4 \phi_{f_1}^2(z) f_1(z) dz$$

is finite, hence also, under (iii) above, the integrals

$$\mathscr{I}(f_1) := \int_{-\infty}^{+\infty} \phi_{f_1}^2(z) f_1(z) dz \quad \text{and} \quad \mathscr{J}(f_1) := \int_{-\infty}^{+\infty} z^2 \phi_{f_1}^2(z) f_1(z) dz;$$

(v) there exists $\beta > 0$ such that

$$\int_a^\infty f_1(z)\, dz = O(a^{-\beta}) \quad \text{as } a \to \infty \quad \text{and} \quad \phi_{f_1}(z) = o(z^{\beta/2-2}) \quad \text{as } z \to \infty.$$

That class $\mathscr{F}_1$ thus consists of all symmetric standardized densities f_1 that are absolutely continuous, strongly unimodal (that is, log-concave), and have finite information $\mathscr{I}(f_1)$ for location, $\mathscr{J}(f_1)$ for scale and, as we shall see, $\mathscr{K}(f_1)$ for asymmetry, with tails satisfying (v). Assumption (iii) is not required for ULAN, but for the asymptotic representation of the rank-based statistics associated with the score function ϕ_{f_1}.

For all $f_1 \in \mathscr{F}_1$, denote by $\kappa(f_1) := \mathscr{J}(f_1)/\mathscr{I}(f_1)$ the ratio of information for scale and information for location. For Gaussian densities ($f_1 = \phi_1$), $\kappa(\phi_1) = 3$ reduces to kurtosis and, as we shall see, $\kappa(f_1)$ can be interpreted as a *generalized kurtosis coefficient*. Finally, write $\mathrm{P}^{(n)}_{\theta,\sigma,\xi;f_1}$ for the probability distribution of $\mathbf{X}^{(n)}$ when the X_i's are i.i.d. with density

$$\begin{aligned} f(x) = {} & \sigma^{-1} f_1\left(\frac{x-\theta}{\sigma}\right) \\ & - \frac{\xi}{\sigma} \dot{f}_1\left(\frac{x-\theta}{\sigma}\right)\left(\left(\frac{x-\theta}{\sigma}\right)^2 - \kappa(f_1)\right) I[|x-\theta| \le \sigma|z^*|] \\ & + \frac{\mathrm{sign}(\xi)}{\sigma} f_1\left(\frac{x-\theta}{\sigma}\right)\Big[I[x-\theta > \mathrm{sign}(-\xi)\sigma|z^*|] - I[x-\theta < \mathrm{sign}(\xi)\sigma|z^*|]\Big]. \end{aligned} \tag{4}$$

Here $\theta \in \mathbb{R}$ and $\sigma \in \mathbb{R}_0^+$ clearly are location and scale parameters, $\xi \in \mathbb{R}$ is a measure of skewness, $\kappa(f_1)$ (strictly positive for $f_1 \in \mathscr{F}_1$) the generalized kurtosis coefficient just defined, and z^* the unique (for ξ small enough; unicity follows from the monotonicity of $\phi_{f_1}(z)$) solution of $f_1(z^*) = \xi \dot{f}_1(z^*)((z^*)^2 - \kappa(f_1))$. The function f defined in (4) is indeed a probability density (nonnegative, integrating up to one), since it is obtained by adding and subtracting the same probability mass

$$\frac{|\xi|}{\sigma} \int_\theta^\infty \min\left(\dot{f}_1\left(\frac{x-\theta}{\sigma}\right)\left(\left(\frac{x-\theta}{\sigma}\right)^2 - \kappa(f_1)\right),\ f_1\left(\frac{x-\theta}{\sigma}\right)\right) dx$$

on both sides of θ (according to the sign of ξ). Note that $\xi > 0$ implies $f(x) = 0$ for $x - \theta < -\sigma|z^*|$ and $f(x) = 2\sigma^{-1} f_1((x-\theta)/\sigma)$ for $x - \theta > \sigma|z^*|$. Moreover, the density $x \mapsto f(x)$ is continuous whenever $\dot{f}_1(x)$ is, vanishes for $x \le \theta + \sigma z^*$ if $\xi > 0$, for $x \ge \theta + \sigma z^*$ if $\xi < 0$, and is left- or right-skewed according as $\xi < 0$ or $\xi > 0$. As for z^*, it tends to $-\infty$ as $\xi \downarrow 0$, to ∞ as $\xi \uparrow 0$; in the Gaussian case, it is easy to check that $|z^*| = O(|\xi|^{-1/3})$ as $\xi \to 0$.

The intuition behind this class of alternatives is that, in the Gaussian case, equation (4), with $\xi = n^{-1/2}\tau$ yields (for $x \in [\theta \pm \sigma z^*]$) the first-order Edgeworth ex-

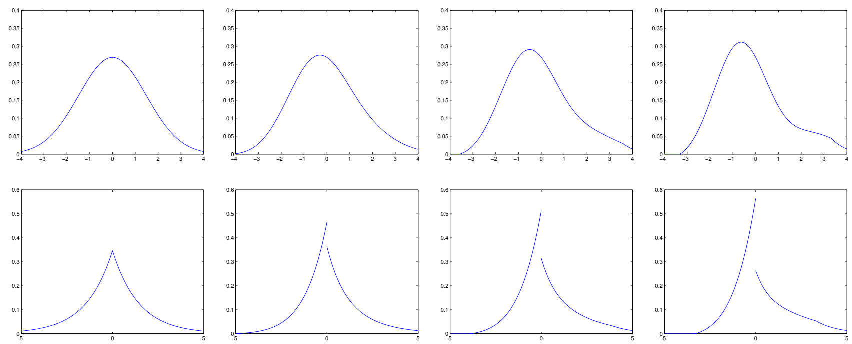

Fig. 1 *Graphical representations of the Gaussian ($f_1 = \phi_1$) and double-exponential ($f_1 = f_{\mathscr{L}}$) Edgeworth families (4), for asymmetry parameter values* $\xi = 0,\ 0.05,\ 0.10,$ *and* 0.15.

pansion of the density of the standardized mean of an i.i.d. n-tuple of variables with third-order moment $6\tau\sigma^3$ (where standardization is based on the median σ of absolute deviations from θ). For a "local" value of ξ, of the form $n^{-1/2}\tau$, (4) thus describes the type of deviation from symmetry that corresponds to the classical central-limit context. Hence, if a Gaussian density is justified as resulting from the additive combination of a large number of small independent symmetric variables, the locally asymmetric f results from the same additive combination, of independent, but slightly skewed observations. We show in Cassart et al. (2011) that the locally optimal test in such case is the traditional test based on $b_1^{(n)}$.

Besides the Gaussian one, interesting special cases of (4) are obtained in the vicinity of

(i) the Student distributions with $\nu > 2$ degrees of freedom, with standardized densities

$$f_1(z) = f_{t_\nu}(z) := C_{t_\nu}(1 + a_\nu z^2/\nu)^{-(\nu+1)/2},$$

$$\mathscr{I}(f_1) = a_\nu(\nu+1)/(\nu+3), \quad \mathscr{J}(f_1) = 3(\nu+1)/(\nu+3),$$

and

$$\mathscr{K}(f_1) = 15\nu(\nu+1)/a_\nu(\nu-2)(\nu+3);$$

the corresponding Gaussian values (density $\phi_1(z) := (a/2\pi)^{1/2}\exp(-az^2/2)$) are obtained by taking limits as $\nu \to \infty$: $\mathscr{I}(\phi_1) = a \approx 0.4549$, $\mathscr{J}(\phi_1) = 3$ and $\mathscr{K}(\phi_1) = 15/a$;

(ii) the logistic distributions, with standardized density

$$f_1(z) = f_{\text{Log}}(z) := \sqrt{b}\exp(-\sqrt{b}z)/(1+\exp(-\sqrt{b}z))^2,$$

$$\mathscr{I}(f_1) = b/3, \quad \mathscr{J}(f_1) = (12+\pi^2)/9, \quad \text{and} \quad \mathscr{K}(f_1) = \pi^2(120+7\pi^2)/45b;$$

(iii) the double-exponential (or Laplace) distributions, with standardized density

$$f_1(z) = f_{\mathscr{L}}(z) := (1/2d)\exp(-|z|/d),$$

$$\mathscr{I}(f_1) = 1/d^2, \quad \mathscr{J}(f_1) = 2, \quad \text{and} \quad \mathscr{K}(f_1) = 24d^2;$$

(iv) the power-exponential distributions, with standardized densities

$$f_1(z) = f_{\exp_\eta}(z) := C_{\exp_\eta}\exp(-(g_\eta z)^{2\eta}),\ \eta \in \mathbb{N}_0,$$

$$\mathscr{I}(f_1) = 2g_\eta^2\eta\frac{\Gamma(2-1/2\eta)}{\Gamma(1+1/2\eta)}, \quad \mathscr{J}(f_1) = 1+2\eta, \textit{and} \ \ \mathscr{K}(f_1) = 2g_\eta\frac{\eta}{\Gamma(1+1/2\eta)}$$

(the positive constants $C_{t_\nu}, C_{\exp_\eta}, a_\nu, a, b, d$, and g_η are such that $f_1 \in \mathscr{F}_1$). Figure 1 provides graphical representations of some densities in the Gaussian ($f_1 = \phi_1$) and double-exponential ($f_1 = f_{\mathscr{L}}$) Edgeworth families (4), respectively. In the Gaussian case, the skewed densities are continuous, while the double-exponential case, due to the discontinuity of $\dot{f}_{\mathscr{L}}(x)$ at $x = 0$, exhibits a discontinuity at the origin.

2.2 Uniform local asymptotic normality (ULAN).

The main technical tool in our derivation of optimal tests is the uniform local asymptotic normality (ULAN), with respect to $\boldsymbol{\vartheta} := (\theta, \sigma, \xi)'$, at any $(\theta, \sigma, 0)'$, of the parametric families

$$\mathscr{P}_{f_1}^{(n)} := \bigcup_{\sigma>0} \mathscr{P}_{\sigma;f_1}^{(n)} := \bigcup_{\sigma>0} \left\{ \mathrm{P}_{\theta,\sigma,\xi;f_1}^{(n)} \,|\, \theta \in \mathbb{R},\ \xi \in \mathbb{R} \right\}, \tag{5}$$

where $f_1 \in \mathscr{F}_1$. More precisely, we are using the following result, which is proved in Cassart et al. (2011).

Proposition 1 *(ULAN) For any $f_1 \in \mathscr{F}_1$, $\theta \in \mathbb{R}$, and $\sigma \in \mathbb{R}_0^+$, the family $\mathscr{P}_{f_1}^{(n)}$ is ULAN at $(\theta, \sigma, 0)'$, with (writing Z_i for $Z_i^{(n)}(\theta, \sigma) := \sigma^{-1}(X_i^{(n)} - \theta)$) central sequence*

$$\boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta}) =: \begin{pmatrix} \Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;2}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}) \end{pmatrix} = n^{-1/2}\sum_{i=1}^{n} \begin{pmatrix} \sigma^{-1}\phi_{f_1}(Z_i) \\ \sigma^{-1}(\phi_{f_1}(Z_i)Z_i - 1) \\ \phi_{f_1}(Z_i)\left(Z_i^2 - \kappa(f_1)\right) \end{pmatrix} \tag{6}$$

and full-rank information matrix

$$\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}) = \begin{pmatrix} \sigma^{-2}\mathscr{I}(f_1) & 0 & 0 \\ 0 & \sigma^{-2}(\mathscr{J}(f_1) - 1) & 0 \\ 0 & 0 & \gamma(f_1) \end{pmatrix}, \tag{7}$$

where $\gamma(f_1) := \mathscr{K}(f_1) - \mathscr{J}^2(f_1)/\mathscr{I}(f_1)$. In other words, for any sequence $\boldsymbol{\vartheta}^{(n)}$ of the form $(\theta^{(n)}, \sigma^{(n)}, 0)'$ such that $\boldsymbol{\vartheta}^{(n)} - \boldsymbol{\vartheta} = O(n^{-1/2})$ for some $\boldsymbol{\vartheta} = (\theta, \sigma, 0)'$, *as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\theta^{(n)},\sigma^{(n)},0;f_1}$,

$$\log \frac{d\mathrm{P}^{(n)}_{\theta^{(n)},\sigma^{(n)},n^{1/2}\tau^{(n)};f_1}}{d\mathrm{P}^{(n)}_{\theta^{(n)},\sigma^{(n)},0;f_1}} = \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}^{(n)}) - \frac{1}{2}\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)$$

for any bounded sequence $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^3$*, and* $\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}^{(n)})$ *is asymptotically normal with mean* $\mathbf{0}$ *and covariance matrix* $\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})$.

The diagonal form of the information matrix $\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})$ confirms that location, scale, and skewness, in the parametric family (5), play distinct and well separated roles. Note that orthogonality between the scale and skewness components of $\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta})$ automatically follows from the symmetry of f_1, while for location and skewness, this orthogonality is a consequence of the definition of $\kappa(f_1)$. The Gaussian versions of (6) and (7) are

$$\boldsymbol{\Delta}^{(n)}_{\phi_1}(\boldsymbol{\vartheta}) = n^{-1/2}\sum_{i=1}^{n}\begin{pmatrix} a\sigma^{-1}Z_i \\ \sigma^{-1}(aZ_i^2-1) \\ aZ_i(Z_i^2-\frac{3}{a}) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma}_{\phi_1}(\boldsymbol{\vartheta}) = \begin{pmatrix} a\sigma^{-2} & 0 & 0 \\ 0 & 2\sigma^{-2} & 0 \\ 0 & 0 & 6/a \end{pmatrix},$$

respectively (recall that $a \approx 0.4549$).

2.3 *Signed-rank versions of the central sequence.*

As mentioned in the introduction, the hypothesis of symmetry enjoys strong group invariance features. The null hypothesis $\mathscr{H}^{(n)}_{\theta}$ of symmetry with respect to θ indeed is generated by the group $\mathscr{G}^{(n)}_{\theta},\circ$ of all transformations $\mathscr{G}_h$ of $\mathbb{R}^n$ such that $\mathscr{G}_h(x_1,\ldots,x_n) := (h(x_1),\ldots,h(x_n))$, where $\lim_{x\to\infty} h(x) = \infty$, and $x \mapsto h(x)$ is continuous, monotone increasing, and skew-symmetric with respect to θ (i.e., such that $h(\theta - z) = -h(\theta + z)$). A maximal invariant for that group is known to be the vector $(s_1(\theta),\ldots,s_n(\theta))$ of signs along with the vector $(R^{(n)}_{+,1}(\theta),\ldots,R^{(n)}_{+,n}(\theta))$ of ranks, where $s_i(\theta)$ is the sign of $(X_i - \theta)$ and $R^{(n)}_{+,i}(\theta)$ the rank of $|X_i - \theta|$ among $|X_1 - \theta|,\ldots,|X_n - \theta|$.

General results on semiparametric efficiency (Hallin and Werker 2003) indicate that, in such context, the expectation of the central sequence $\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta})$ conditional on those signs and ranks yields a version of the semiparametrically efficient (at f_1 and $\boldsymbol{\vartheta}$) central sequence. The only component of the central sequence $\boldsymbol{\Delta}^{(n)}_{f_1}$ which is used in this section is the ξ-component $\Delta^{(n)}_{f_1;3}$, with *signed-rank version* (a terminology justified in Proposition 2)

$$\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta) := n^{-1/2}\sum_{i=1}^{n} s_i(\theta)\phi_{f_1}\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)\Big(\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)^2 - \kappa(f_1)\Big),$$

where $F_{1+} : (x) \mapsto (2F_1(x)-1)I[x \geq 0]$ and F_1 denote the distribution functions of $|Z_i|$ and Z_i, respectively, when Z_i has density f_1. Later on, however, we also will need the signed-rank version

$$\underset{\sim}{\Delta}^{(n)}_{f_1;1}(\theta) := n^{-1/2}\sum_{i=1}^{n} s_i(\theta)\phi_{f_1}\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)$$

of the θ-component. The following results follow from the classical Hájek theory for linear signed-rank statistics (see, e.g., Chapter 3 of Puri and Sen 1985).

Proposition 2 *Let $f_1 \in \mathscr{F}_1$ and $g_1 \in \mathscr{F}_0$. Then,*

(i) under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, as $n \to \infty$,

$$\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta) = n^{-1/2}\sum_{i=1}^{n} \phi_{f_1}\Big(F_1^{-1}(G_1(Z^{(n)}_i(\theta,\sigma)))\Big) \times\Big(\Big(F_1^{-1}(G_1(Z^{(n)}_i(\theta,\sigma)))\Big)^2 - \kappa(f_1)\Big) + o_{L^2}(1),$$

and hence, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;f_1}$, $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta) = \Delta^{(n)}_{f_1;3}(\theta,\sigma,0) + o_{\mathrm{P}}(1)$;

(ii) under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta)$ has mean zero and variance

$$\underset{\sim}{\gamma}^{(n)}_{f_1} := n^{-1}\sum_{r=1}^{n} \phi^2_{f_1}\Big(F^{-1}_{1+}\Big(\frac{r}{n+1}\Big)\Big)\Big(\Big(F^{-1}_{1+}\Big(\frac{r}{n+1}\Big)\Big)^2 - \kappa(f_1)\Big)^2 \tag{8}$$

tending to $\gamma(f_1)$ as $n \to \infty$.

Note that Part (i) of this result entails that $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta)$, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, with $g_1 \in \mathscr{F}_0$, is asymptotically equivalent (in the mean square sense) to a random variable of the form $n^{-1/2}\sum_{i=1}\eta_i$, where the η_i's are i.i.d. with mean zero (since $u \mapsto \phi_{f_1}(u)u^2$ and $u \mapsto \phi_{f_1}(u)$ are skew-symmetric) and variance $\gamma(f_1)$. Consequently, $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta)$, still under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, is also asymptotically normal with mean zero and variance $\gamma(f_1)$.

3 Rank-based tests for symmetry.

3.1 Optimal signed-rank tests of symmetry: specified location.

Proposition 2 immediately yields a distribution-free signed-rank test of the hypothesis of symmetry with respect to a specified location θ. With the notation of Theorem 2, consider the rank-based test statistic

$$\underset{\sim}{T}^{(n)}_{f_1}(\theta) := \left(\underset{\sim}{\gamma}^{(n)}_{f_1}\right)^{-1/2} \underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta) \tag{9}$$

$$= \left(n\underset{\sim}{\gamma}^{(n)}_{f_1}\right)^{-1/2} \sum_{i=1}^{n} s_i(\theta)\phi_{f_1}\left(F^{-1}_{1+}\left(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\right)\right)\left(\left(F^{-1}_{1+}\left(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\right)\right)^2 - \kappa(f_1)\right).$$

Define the *cross-information coefficients*

$$\mathscr{I}(f_1,g_1) := \int_0^1 \phi_{f_1}\left(F_1^{-1}(u)\right)\phi_{g_1}\left(G_1^{-1}(u)\right)du,$$

$$\mathscr{J}(f_1,g_1) := \int_0^1 (F_1^{-1}(u))^2\phi_{f_1}\left(F_1^{-1}(u)\right)\phi_{g_1}\left(G_1^{-1}(u)\right)du, \text{ and}$$

$$\mathscr{K}(f_1,g_1) := \int_0^1 (F_1^{-1}(u))^2(G_1^{-1}(u))^2\phi_{f_1}\left(F_1^{-1}(u)\right)\phi_{g_1}\left(G_1^{-1}(u)\right)du,$$

and denote by

$$\underset{\sim}{\mathscr{F}}_{f_1} := \{g_1 \in \mathscr{F}_0 | \mathscr{I}(f_1,g_1) < \infty,\ \mathscr{J}(f_1,g_1) < \infty, \text{ and } \mathscr{K}(f_1,g_1) < \infty\}$$

the class of densities for which those integrals exist and are finite.

Proposition 3 *Let $f_1 \in \mathscr{F}_1$. Then,*

(i) $\underset{\sim}{T}^{(n)}_{f_1}(\theta)$ *is asymptotically normal, with mean zero under* $\bigcup_{g_1\in\mathscr{F}_0}\bigcup_{\sigma\in\mathbb{R}_0^+}\{\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}\}$, *mean*

$$\frac{\tau}{\gamma^{1/2}(f_1)}\left[\mathscr{K}(f_1,g_1) - \mathscr{J}(f_1,g_1)\kappa(g_1) - \mathscr{J}(g_1,f_1)\kappa(f_1) + \mathscr{I}(f_1,g_1)\kappa(f_1)\kappa(g_1)\right]$$

under $\bigcup_{\sigma\in\mathbb{R}_0^+}\{\mathrm{P}^{(n)}_{\theta,\sigma,n^{-1/2}\tau;g_1}\}$, $g_1 \in \underset{\sim}{\mathscr{F}}_{f_1}$, *and variance one under both;*

(ii) the sequence of tests rejecting the hypothesis $\mathscr{H}^{(n)}_\theta := \bigcup_{g_1\in\mathscr{F}_0}\bigcup_{\sigma\in\mathbb{R}_0^+}\{\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}\}$ *of symmetry with respect to* θ *whenever* $\underset{\sim}{T}^{(n)}_{f_1}(\theta)$ *exceeds the* $(1-\alpha)$ *standard normal quantile* z_α *is locally asymptotically most powerful, at asymptotic level* α, *against* $\bigcup_{\xi>0}\bigcup_{\sigma\in\mathbb{R}_0^+}\{\mathrm{P}^{(n)}_{\theta,\sigma,\xi;f_1}\}$.

Only asymptotic critical values are reported in Part (ii) of the proposition, but exact (or simulated) ones of course also can be considered, as the test is entirely distribution-free. The two-sided version also readily follows.

3.2 Optimal signed-rank tests of symmetry: unspecified location.

The unspecified-location case is more difficult and, to the best of our knowledge, so far only has been considered in Cassart et al. (2008). When θ is unspecified, a consistent estimator $\hat{\theta}$ has to be substituted for θ, yielding *aligned* signs $s_i(\hat{\theta})$ and *aligned* ranks $R^{(n)}_{+,i}(\hat{\theta})$. This substitution, however, has an impact on the distribution-freeness of $\underset{\sim}{T}^{(n)}_{f_1}$, and on its asymptotic distribution. That impact, as we shall see, requires a delicate asymptotic analysis.

As usual in this context, denoting by $\mathrm{P}^{(n)}_{\lambda}$ a sequence of probability measures indexed by a parameter λ, consider a sequence of estimators $\hat{\lambda}^{(n)}$ of λ satisfying, under $\mathrm{P}^{(n)}_{\lambda}$, the following assumptions:

(C1) (root-n consistency) $\hat{\lambda}^{(n)} - \lambda = O_{\mathrm{P}}(n^{-1/2})$, and
(C2) (local discreteness) the number of possible values of $\hat{\lambda}^{(n)}$ in balls with $O(n^{-1/2})$ radius centered at λ is bounded as $n \to \infty$.

An estimator $\lambda^{(n)}$ satisfying (C1) but not (C2) is easily discretized by letting, for some arbitrary $c > 0$, $\lambda^{(n)}_{\#} := (cn^{1/2})^{-1}\mathrm{sign}(\lambda^{(n)})\lceil cn^{1/2}|\lambda^{(n)}|\rceil$, which satisfies both (C1) and (C2). Subscripts $_{\#}$ in the sequel are used for estimators ($\hat{\theta}_{\#}$, $\hat{\sigma}_{\#}$, ...) satisfying (C1) and (C2). It should be noted, however, that (C2) has no implications in practice, where n is fixed, as the discretization constant c can be chosen arbitrarily large.

For all $\kappa \in \mathbb{R}^+_0$, define

$$\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\kappa;\theta) := n^{-1/2}\sum_{i=1}^{n} s_i(\theta)\phi_{f_1}\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)\Big(\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)^2 - \kappa\Big);$$

for $\kappa = \kappa(f_1)$, $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\kappa;\theta)$ and $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\theta)$ coincide. The joint distribution, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, of

$$\begin{pmatrix} \underset{\sim}{\Delta}^{(n)}_{f_1;3}(\kappa;\theta) \\ \Delta^{(n)}_{g_1;1}(\theta,\sigma,0) \end{pmatrix} = n^{-1/2}\sum_{i=1}^{n}\begin{pmatrix} \phi_{f_1}\Big(F_1^{-1}(G_1(Z^{(n)}_i(\theta,\sigma)))\Big)\Big(\Big(F_1^{-1}(G_1(Z^{(n)}_i(\theta,\sigma)))\Big)^2 - \kappa\Big) \\ \frac{1}{\sigma}\phi_{g_1}(Z_i(\theta,\sigma)) \end{pmatrix} + o_{\mathrm{P}}(1),$$

(an asymptotic representation similar to that of Part (i) of Proposition 2 clearly holds) is asymptotically normal, with mean zero and covariance matrix

$$\begin{pmatrix} \underset{\sim}{\gamma}^{\kappa}(f_1) & \delta^{\kappa}(f_1,g_1) \\ \delta^{\kappa}(f_1,g_1) & \sigma^{-2}\mathscr{I}(g_1) \end{pmatrix}, \tag{10}$$

where

$$\underset{\sim}{\gamma}^{\kappa}(f_1) := \mathscr{K}(f_1) - 2\kappa\mathscr{J}(f_1) + \kappa^2\mathscr{I}(f_1)$$

and

$$\delta^{\kappa}(f_1,g_1) := \sigma^{-1}(\mathscr{J}(f_1,g_1) - \kappa\mathscr{I}(f_1,g_1)).$$

It immediately follows from (10) and Le Cam's third Lemma that root-n perturbations of θ, hence also the replacement of θ by a root-n consistent estimator, do affect the asymptotic centering of $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\kappa;\theta)$—unless the covariance $\delta^{\kappa}(f_1,g_1)$ is zero, that is, unless

$$\kappa = \kappa(f_1, g_1) := \mathscr{J}(f_1, g_1)/\mathscr{I}(f_1, g_1).$$

Let $\underset{\sim}{\kappa}^{(n)}(f_1;\theta)$ be a consistent (under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$) estimator of $\kappa(f_1, g_1)$. Since the mapping from κ to the values of $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\kappa;\theta)$ is continuous, $\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\underset{\sim}{\kappa}^{(n)}(f_1;\theta);\theta)$ also can be expected to be asymptotically insensitive to root-n perturbations of θ. This is indeed the case, and the same reasoning as in Section 3.2.2 of Cassart et al. (2011) yields, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, the asymptotic equivalence

$$\underset{\sim}{\Delta}^{(n)}_{f_1;3}(\underset{\sim}{\kappa}^{(n)}(f_1;\hat{\theta}_{\#});\hat{\theta}_{\#}) - \underset{\sim}{\Delta}^{(n)}_{f_1;3}(\underset{\sim}{\kappa}^{(n)}(f_1;\theta);\theta) = o_{\mathrm{P}}(1)$$

for any estimator $\hat{\theta}_{\#}$ of θ satisfying (C1) and (C2) and any density g_1 in

$$\underset{\sim}{\mathscr{F}}^{*}_{f_1} := \{g_1 \in \mathscr{F}_0 \mid \mathscr{I}(g_1) < \infty,\ \mathscr{I}(f_1, g_1) < \infty, \text{ and } \mathscr{J}(f_1, g_1) < \infty\}.$$

Obtaining a consistent estimator $\underset{\sim}{\kappa}^{(n)}(f_1;\theta)$ of $\underset{\sim}{\kappa}(f_1, g_1)$, which is a ratio of expected values, taken under unspecified density g, of variables that themselves depend on g, however, is delicate. A general method is proposed in Cassart et al. (2010), which we describe in Section 4. Defining

$$\underset{\sim}{T}^{(n)*}_{f_1}(\theta) := \left(n\,\underset{\sim}{\gamma}^{(n)*}_{f_1}\right)^{-1/2} \times \sum_{i=1}^{n} s_i(\theta)\phi_{f_1}\left(F^{-1}_{1+}\left(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\right)\right)\left(\left(F^{-1}_{1+}\left(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\right)\right)^2 - \underset{\sim}{\kappa}^{(n)}(f_1;\theta)\right), \tag{11}$$

where (convergence under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, as $n\to\infty$)

$$\begin{aligned}\underset{\sim}{\gamma}^{(n)*}_{f_1} &:= n^{-1}\sum_{r=1}^{n}\phi^2_{f_1}\left(F^{-1}_{1+}\left(\frac{r}{n+1}\right)\right)\left(\left(F^{-1}_{1+}\left(\frac{r}{n+1}\right)\right)^2 - \underset{\sim}{\kappa}^{(n)}(f_1;\theta)\right)^2 \\ &= \mathscr{K}(f_1) - 2\underset{\sim}{\kappa}(f_1,g_1)\mathscr{J}(f_1) + \underset{\sim}{\kappa}^2(f_1,g_1)\mathscr{I}(f_1) + o(1),\end{aligned}$$

we thus have established the following result.

Proposition 4 *Let $f_1 \in \mathscr{F}_1$. Then,*

(i) $\underset{\sim}{T}^{(n)*}_{f_1}(\hat{\theta}_{\#})$ is asymptotically normal, with mean zero under the null hypothesis $\bigcup_{g_1\in\underset{\sim}{\mathscr{F}}^{*}_{f_1}}\bigcup_{\theta\in\mathbb{R}}\bigcup_{\sigma\in\mathbb{R}^+_0}\{\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}\}$ of symmetry with respect to unspecified θ, mean

$$\tau\frac{\mathscr{K}(f_1,g_1) - \mathscr{J}(g_1,f_1)\underset{\sim}{\kappa}(f_1,g_1)}{\left(\mathscr{K}(f_1) - 2\underset{\sim}{\kappa}(f_1,g_1)\mathscr{J}(f_1) + \underset{\sim}{\kappa}^2(f_1,g_1)\mathscr{I}(f_1)\right)^{1/2}}$$

under the local alternative $\mathrm{P}^{(n)}_{\theta,\sigma,n^{-1/2}\tau;g_1}$, *where* $g_1 \in \underset{\sim}{\mathscr{F}}^{*}_{f_1}$, *and variance one under both;*

(ii) the sequence of tests rejecting the null hypothesis $\bigcup_{g_1\in \underset{\sim}{\mathscr{F}}^{*}_{f_1}}\bigcup_{\theta\in\mathbb{R}}\bigcup_{\sigma\in\mathbb{R}^+_0}\{\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}\}$ *of symmetry with respect to unspecified* θ *whenever* $\underset{\sim}{T}^{(n)*}_{f_1}(\hat{\theta}_{\#})$ *exceeds the* $(1-\alpha)$ *standard normal quantile* z_α *is locally asymptotically most powerful, at asymptotic level* α, *against* $\bigcup_{\xi>0}\bigcup_{\theta\in\mathbb{R}}\bigcup_{\sigma\in\mathbb{R}^+_0}\{\mathrm{P}^{(n)}_{\theta,\sigma,\xi;f_1}\}$.

3.3 The van der Waerden, Wilcoxon and Laplace tests of symmetry.

Important particular cases of (9) and (11) are the Laplace (sign test or double-exponential scores), Wilcoxon (logistic scores) and van der Waerden (normal scores) tests, which are optimal at double-exponential, logistic, and normal distributions, respectively.

The van der Waerden tests are based on $f_1 = \phi_1$, with

$$F^{-1}_{1+}(u) = \phi_{f_1}(F^{-1}_{1+}(u)) = a^{-1/2}\Phi^{-1}\Big(\frac{u+1}{2}\Big),$$

where Φ is the standard normal distribution function. The specified-location test statistic (9) then reduces to

$$\underset{\sim}{T}^{(n)}_{\mathrm{vdW}}(\theta) := \Big(n\underset{\sim}{\gamma}^{(n)}_{\phi_1}\Big)^{-1/2} \times \sum_{i=1}^{n} s_i(\theta)\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\theta)}{2(n+1)}\Big)\Bigg(\Big(\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\theta)}{2(n+1)}\Big)\Big)^2-3\Bigg),$$

where

$$\underset{\sim}{\gamma}^{(n)}_{\phi_1} := n^{-1}\sum_{r=1}^{n}\Bigg(\Phi^{-1}\Big(\frac{n+1+r}{2(n+1)}\Big)\Bigg)^2\Bigg(\Big(\Phi^{-1}\Big(\frac{n+1+r}{2(n+1)}\Big)\Big)^2-3\Bigg)^2.$$

The corresponding unspecified-location test statistic (11) takes the form

$$\underset{\sim}{T}^{(n)*}_{\mathrm{vdW}}(\hat{\theta}) := \Big(n\,\underset{\sim}{\gamma}^{(n)*}_{\phi_1}\Big)^{-1/2}\sum_{i=1}^{n} s_i(\hat{\theta})\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\hat{\theta})}{2(n+1)}\Big) \times\Bigg(\Big(\Phi^{-1}\Big(\frac{n+1+R^{(n)}_{+,i}(\hat{\theta})}{2(n+1)}\Big)\Big)^2-\underset{\sim}{\kappa}^{(n)}(\phi_1;\hat{\theta})\Bigg),$$

where

$$\underset{\sim}{\gamma}^{(n)*}_{\phi_1} := n^{-1}\sum_{r=1}^{n}\left(\Phi^{-1}\left(\frac{n+1+r}{2(n+1)}\right)\right)^2\left(\left(\Phi^{-1}\left(\frac{n+1+r}{2(n+1)}\right)\right)^2-\underset{\sim}{\kappa}^{(n)}(\phi_1;\hat{\theta})\right)^2.$$

In the Wilcoxon case (logistic density), one easily checks that

$$F_{1+}^{-1}(u)=b^{-1/2}\log\frac{1+u}{1-u}\quad\text{and}\quad \phi_{f_1}(F_{1+}^{-1}(u))=b^{-1/2}u.$$

Therefore, (9) and (11) reduce to

$$\underset{\sim}{T}^{(n)}_{\mathrm{W}}(\theta) := \left(n\underset{\sim}{\gamma}^{(n)}_{f_{\mathrm{Log}}}\right)^{-1/2}\sum_{i=1}^{n}s_i(\theta)R^{(n)}_{+,i}(\theta)\left(\left(\log\frac{n+1+R^{(n)}_{+,i}(\theta)}{n+1-R^{(n)}_{+,i}(\theta)}\right)^2-\frac{12+\pi^2}{3}\right),$$

and

$$\underset{\sim}{T}^{(n)*}_{\mathrm{W}}(\hat{\theta}) := \left(n\underset{\sim}{\gamma}^{(n)*}_{f_{\mathrm{Log}}}\right)^{-1/2}\sum_{i=1}^{n}s_i(\hat{\theta})R^{(n)}_{+,i}(\hat{\theta})\left(\left(\log\frac{n+1+R^{(n)}_{+,i}(\hat{\theta})}{n+1-R^{(n)}_{+,i}(\hat{\theta})}\right)^2-\underset{\sim}{\kappa}^{(n)}(f_{\mathrm{Log}};\hat{\theta})\right),$$

where

$$\underset{\sim}{\gamma}^{(n)}_{f_{\mathrm{Log}}} := n^{-1}\sum_{r=1}^{n}r^2\left(\left(\log\frac{n+1+r}{n+1-r}\right)^2-\frac{12+\pi^2}{3}\right)^2$$

and

$$\underset{\sim}{\gamma}^{(n)*}_{f_{\mathrm{Log}}} := n^{-1}\sum_{r=1}^{n}r^2\left(\left(\log\frac{n+1+r}{n+1-r}\right)^2-\underset{\sim}{\kappa}^{(n)}(f_{\mathrm{Log}};\hat{\theta})\right)^2,$$

respectively.

As for the Laplace-score version of (9), it is associated with the double-exponential density $f_1 = f_{\mathscr{L}}$. One easily obtains

$$F_{1+}^{-1}(u)=-d\log(1-u)\quad\text{and}\quad \phi_{f_1}(F_{1+}^{-1}(u))=1/d,$$

hence

$$\underset{\sim}{T}^{(n)}_{\mathscr{L}}(\theta) := \left(n\underset{\sim}{\gamma}^{(n)}_{f_{\mathscr{L}}}\right)^{-1/2}\sum_{i=1}^{n}s_i(\theta)\left(\left(\log\left(1-\frac{R^{(n)}_{+,i}(\theta)}{n+1}\right)\right)^2-2\right),$$

where

$$\underset{\sim}{\gamma}^{(n)}_{f_{\mathscr{L}}} := n^{-1}\sum_{r=1}^{n}\left(\left(\log\left(1-\frac{r}{n+1}\right)\right)^2-2\right)^2.$$

The unspecified-location test statistic (11) is derived along the same lines as previously, yielding

$$\underset{\sim}{T}^{(n)*}_{\mathscr{L}}(\hat{\theta}) := \left(n\underset{\sim}{\gamma}^{(n)*}_{f_{\mathscr{L}}}\right)^{-1/2}\sum_{i=1}^{n}s_i(\hat{\theta})\left(\left(\log\left(1-\frac{R^{(n)}_{+,i}(\hat{\theta})}{n+1}\right)\right)^2-\underset{\sim}{\kappa}^{(n)}(f_{\mathscr{L}};\hat{\theta})\right),$$

with

$$\underset{\sim}{\gamma}^{(n)*}_{f_{\mathscr{L}}} := n^{-1} \sum_{r=1}^{n} \left(\left(\log\left(1 - \frac{r}{n+1}\right) \right)^2 - \underset{\sim}{\kappa}^{(n)}(f_{\mathscr{L}}; \hat{\theta}) \right)^2 .$$

3.4 Asymptotic relative efficiencies.

The asymptotic shifts provided in Propositions 3 and 4, together with their pseudo-Gaussian counterparts in Propositions 3.2 and 3.5 of Cassart et al. (2011) allow us to compute ARE values for the tests based on $\underset{\sim}{T}^{(n)}_{f_1}(\theta)$ and $\underset{\sim}{T}^{(n)*}_{f_1}(\hat{\theta}_{\#})$ with respect to their classical counterparts, based on $m_3^{(n)}(\theta)$ (or $S_1^{(n)}$; see (1)) and $b_1^{(n)}$ (or $S_2^{(n)}$; see (2)), respectively. Those ARE values are obtained as the squared ratios of the corresponding local shifts, for various densities g_1. The pseudo-Gaussian tests, hence also our ARE values, require finite sixth-order moments. But signed-rank tests of course remain valid without such assumption and, whenever g_1 has infinite moment of order six, the asymptotic relative efficiency under g_1 of any signed-rank test with respect to its pseudo-Gaussian competitor can be considered as infinite.

Proposition 5 *Let $f_1 \in \mathscr{F}_1$; denoting by μ_k the moment of order k of g_1, assume that $\mu_6 < \infty$.*

(i) The asymptotic relative efficiency under $g_1 \in \mathscr{F}_{\phi_1} \cap \underset{\sim}{\mathscr{F}}_{f_1}$ of the specified-location signed-rank test based on $\underset{\sim}{T}^{(n)}_{f_1}(\theta)$ with respect to the classical procedure based on $m_3^{(n)}(\theta)$ (see (1)) is

$$ARE_{g_1}\big(\underset{\sim}{T}^{(n)}_{f_1}(\theta)/m_3^{(n)}(\theta)\big) = \frac{\Big(\mathscr{K}(f_1,g_1) - \mathscr{J}(f_1,g_1)\kappa(g_1) - \mathscr{J}(g_1,f_1)\kappa(f_1) + \mathscr{I}(f_1,g_1)\kappa(f_1)\kappa(g_1)\Big)^2}{\gamma(f_1)\Big(5\mu_4 - 3\kappa(g_1)\mu_2\Big)^2/\mu_6}.$$

(ii) The asymptotic relative efficiency under $g_1 \in \mathscr{F}_{\phi_1} \cap \underset{\sim}{\mathscr{F}}_{f_1}$ of the specified-location signed-rank test based on $\underset{\sim}{T}^{(n)}_{f_1}(\theta)$ with respect to the classical procedure based on $b_1^{(n)}$ (see (2)) is

$$ARE_{g_1}(\underset{\sim}{T}^{(n)}_{f_1}(\theta)/b_1^{(n)}) = \frac{\Big(\mathscr{K}(f_1,g_1) - \mathscr{J}(g_1,f_1)\kappa(g_1) - \mathscr{J}(g_1,f_1)\kappa(f_1) + \mathscr{I}(f_1,g_1)\kappa(f_1)\kappa(g_1)\Big)^2}{\gamma(f_1)\big(5\mu_4 - 9\mu_2^2\big)^2/\big(\mu_6 - 6\mu_2\mu_4 + 9\mu_2^3\big)}.$$

(iii) *The asymptotic relative efficiency under* $g_1 \in \mathscr{F}_{\phi_1} \cap \underset{\sim}{\mathscr{F}}^*_{f_1}$ *of the unspecified-location signed-rank test based on* $\underset{\sim}{T}^{(n)*}_{f_1}(\hat{\theta}_\#)$ *with respect to the classical procedure based on* $b_1^{(n)}$ *(see (2)) is*

$$ARE_{g_1}(\underset{\sim}{T}^{(n)*}_{f_1}(\hat{\theta}_\#)/b_1^{(n)}) = \frac{\left(\mathscr{K}(f_1,g_1) - \mathscr{J}(g_1,f_1)\underset{\sim}{\kappa}(f_1,g_1)\right)^2\left(\mu_6 - 6\mu_2\mu_4 + 9\mu_2^3\right)}{\left(\mathscr{K}(f_1) - 2\underset{\sim}{\kappa}(f_1,g_1)\mathscr{J}(f_1) + \underset{\sim}{\kappa}^2(f_1,g_1)\mathscr{I}(f_1)\right)\left(5\mu_4 - 9\mu_2^2\right)^2}.$$

Numerical values of those AREs, under Student (6.5, 8, 10, and 20 degrees of freedom), normal, power-exponential (exponents 2, 3, and 5), and logistic densities are displayed in Tables 1 and 2. Those values are quite high, particularly so under heavy tails (see, for instance, the Student density with 6.5 degrees of freedom). Also, the AREs of the van der Waerden tests are uniformly larger than or equal to one. The tests with power-exponential scores however are not performing as well under Student and logistic densities. On the other hand, under the power-exponential density with exponent 2, the classical procedure based on $m_3^{(n)}(\theta)$ has no power at all, yielding infinite ARE values (the shift in the denominator is zero) in Table 1.

The van der Waerden tests thus appear as a very attractive alternative to the classical tests—all the more so that their validity does not require the actual density g to have finite moments of order 6; recall, however, that $g_1 \in \underset{\sim}{\mathscr{F}}^*_{f_1}$ (here, $\underset{\sim}{\mathscr{F}}^*_{\phi_1}$) is still needed in the unspecified-location case.

4 Estimation of cross-information quantities.

4.1 Consistent estimation of $\mathscr{I}(f_1,g_1)$ *and* $\mathscr{J}(f_1,g_1)$.

Implementing the unspecified-location rank-based tests of Section 3.2 thus requires consistent estimation of $\underset{\sim}{\kappa}(f_1,g_1)$, that is, consistent estimation of the cross-information quantities $\mathscr{I}(f_1,g_1)$ and $\mathscr{J}(f_1,g_1)$. The cross-information for location $\mathscr{I}(f_1,g_1)$ is a familiar quantity in classical rank-based inference. It explicitly appears, indeed, in the asymptotic powers of traditional rank and signed-rank tests for location, and in the asymptotic variance of the corresponding R-estimators. In a different context (R-estimation of shape), its counterpart also plays a central role in the construction of one-step R-estimators (Hallin, Oja, and Paindaveine 2006).

Estimating $\mathscr{I}(f_1,g_1)$, however, is all but straightforward. No empirical version of this expectation is available, as it involves the unknown score ϕ_{g_1} associated with the unspecified density g_1. Various methods have been proposed for estimating $\mathscr{I}(f_1,g_1)$ in connection with R-estimation. Some of them (Lehmann 1963;

	actual density g_1								
score f_1	$f_{t_{6.5}}$	f_{t_8}	$f_{t_{10}}$	$f_{t_{20}}$	ϕ_1	$f_{\mathscr{E}_2}$	$f_{\mathscr{E}_3}$	$f_{\mathscr{E}_5}$	f_{Log}
$f_{t_{6.5}}$	4.6923	1.8304	1.3556	1.0350	0.9374	0.9593	1.0669	0.9055	1.1843
	3.9000	1.8338	1.6019	1.7313	2.3436	∞	3.6063	0.3666	1.3948
f_{t_8}	4.6848	1.8333	1.3618	1.0462	0.9542	1.0170	1.1670	1.0646	1.1849
	3.8938	1.8367	1.6091	1.7501	2.3855	∞	3.9446	0.4309	1.3955
$f_{t_{10}}$	4.6648	1.8308	1.3636	1.0540	0.9679	1.0733	1.2707	1.2459	1.1826
	3.8771	1.8342	1.6113	1.7630	2.4197	∞	4.2951	0.5044	1.3928
$f_{t_{20}}$	4.5765	1.8074	1.3543	1.0612	0.9904	1.2008	1.5258	1.7512	1.1670
	3.8037	1.8108	1.6003	1.7752	2.4759	∞	5.1574	0.7089	1.3745
ϕ_1	4.3988	1.7494	1.3199	1.0510	1.0000	1.3402	1.8394	2.4753	1.1304
	3.6560	1.7526	1.5596	1.7580	2.5000	∞	6.2175	1.0020	1.3313
$f_{\mathscr{E}_2}$	2.5240	1.0455	0.8207	0.7145	0.7515	1.7834	3.3911	7.2764	0.6740
	2.0978	1.0475	0.9698	1.1952	1.8788	∞	11.4624	2.9454	0.7938
$f_{\mathscr{E}_3}$	1.3575	0.5802	0.4699	0.4391	0.4988	1.6399	3.6878	9.7508	0.3776
	1.1283	0.5812	0.5552	0.7345	1.2470	∞	12.4654	3.9471	0.4448
$f_{\mathscr{E}_5}$	0.3853	0.1770	0.1541	0.1685	0.2245	1.1766	3.2606	11.0283	0.1208
	0.3202	0.1773	0.1820	0.2819	0.5611	∞	11.0214	4.4642	0.1423
f_{Log}	4.6839	1.8311	1.3593	1.0439	0.9528	1.0132	1.1739	1.1232	1.1864
	3.8930	1.8345	1.6062	1.7462	2.3820	∞	3.9680	0.4547	1.3973

Table 1 AREs, under Student ($f_{t_{6.5}}$, f_{t_8}, $f_{t_{10}}$ and $f_{t_{20}}$), normal (ϕ_1), power-exponential ($f_{\mathscr{E}_2}$, $f_{\mathscr{E}_3}$, $f_{\mathscr{E}_5}$, with exponents 2, 3, and 5), and logistic (f_{Log}) densities, of various signed-rank tests (based on Student, van der Waerden, power-exponential, and Wilcoxon scores), with respect to the pseudo-Gaussian test (based on $b_1^{(n)}$, see Proposition 5(ii); first line) and with respect to the classical test of skewness (based on $m_3^{(n)}$, see Proposition 5(i); second line), for testing symmetry about a specified location θ.

Sen 1966) involve comparisons of lengths of confidence intervals. Some others (Kraft and van Eeden 1972, Antille 1974, or Jurečková and Sen 1996, page 321) rely on the asymptotic linearity property of rank statistics. More elaborated approaches involve kernel estimates of g_1—hence cannot be expected to perform well under small and moderate sample sizes. Such kernel methods have been considered, for Wilcoxon scores, by Schweder (1975) (see also Cheng and Serfling 1981, Bickel and Ritov 1988, and Fan 1991) and, in a more general setting, in Section 4.5 of Koul (2002).

All these methods, however, involve quite arbitrary choices (choice of a confidence level for confidence intervals; choice of an arbitrary $O(n^{-1/2})$ perturbation for the method based on asymptotic linearity; choice of a kernel and a bandwidth for the estimation of g_1). Although they do not affect consistency and asymptotic efficiency, such choices may have a dramatic impact on finite sample results. As for kernel methods, they require large sample sizes and are kind of antinomic to the spirit of rank-based methods: if densities are to be estimated, indeed, using them all the way by inserting estimated scores into the parametric tests of Cassart et al. (2011) seems more coherent than considering ranks.

	actual density g_1								
score f_1	$f_{t_{6.5}}$	f_{t_8}	$f_{t_{10}}$	$f_{t_{20}}$	ϕ_1	$f_{\mathscr{E}_2}$	$f_{\mathscr{E}_3}$	$f_{\mathscr{E}_5}$	f_{Log}
$f_{t_{6.5}}$	4.6923	2.7856	1.3628	1.0580	0.9902	1.3769	2.0874	3.7623	1.1853
f_{t_8}	4.6915	1.8333	1.3634	1.0591	0.9919	1.3974	2.1401	3.9103	1.1850
$f_{t_{10}}$	4.6895	2.7846	1.3636	1.0601	0.9938	1.4173	2.1906	4.0506	1.1848
$f_{t_{20}}$	4.6803	2.7795	1.3623	1.0612	0.9978	1.4627	2.3052	4.3671	1.1839
ϕ_1	4.6547	2.7642	1.3558	1.0589	1	1.5129	2.4359	4.7336	1.1799
$f_{\mathscr{E}_2}$	4.1723	1.6254	1.2084	0.9448	0.8993	1.7834	3.5036	8.4598	1.0536
$f_{\mathscr{E}_3}$	3.7109	1.4365	1.0629	0.8255	0.7833	1.7052	3.6878	10.1646	0.9311
$f_{\mathscr{E}_5}$	3.0659	1.1736	0.8610	0.6594	0.6200	1.4453	3.4284	11.0283	0.7591
f_{Log}	4.6878	2.7828	1.3618	1.0585	0.9926	1.3728	2.0678	3.6975	1.1864

Table 2 AREs, under Student ($f_{t_{6.5}}$, f_{t_8}, $f_{t_{10}}$ and $f_{t_{20}}$), normal (ϕ_1), power-exponential ($f_{\mathscr{E}_2}$, $f_{\mathscr{E}_3}$, $f_{\mathscr{E}_5}$, with exponents 2, 3, and 5), and logistic (f_{Log}) densities, of various signed-rank tests (based on Student, van der Waerden, power-exponential, and Wilcoxon scores), with respect to the pseudo-Gaussian test (based on $b_1^{(n)}$, see Proposition 5(iii)), for testing symmetry with respect to unspecified location θ.

The approach proposed in Hallin, Oja, and Paindaveine (2006) is of an entirely different nature; the basic idea consists in solving a local linearized likelihood equation. In the present setting, this estimator of $\mathscr{I}(f_1, g_1)$ is constructed as follows. Denoting by $\hat{\theta}$ and $\widehat{\sigma}$ root-n consistent (under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, $g_1 \in \underset{\sim}{\mathscr{F}}{}^*_{f_1}$) estimators of θ and σ, respectively (the median and the median of absolute deviations constitute appropriate choices), by $\hat{\theta}_\#$ and $\widehat{\sigma}_\#$ their discretized versions, and by $\underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}$ a discretized version of $\underset{\sim}{\Delta}{}^{(n)}_{f_1;1}$, let, for any $\beta > 0$,

$$\underset{\sim}{\theta}{}^{(n)}_*(\beta) := \hat{\theta}_\# + n^{-1/2}\beta\,\widehat{\sigma}^2_\#\underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\hat{\theta}_\#). \tag{12}$$

Choosing a further arbitrary discretization constant $c > 0$, put $\beta_\ell := \ell/c$, $\ell \in \mathbb{N}$, and define

$$\beta_1^- := \min\{\beta_\ell \mid \underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\underset{\sim}{\theta}{}^{(n)}_*(\beta_{\ell+1}))\underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\hat{\theta}_\#) < 0\}, \qquad \beta_1^+ := \beta_1^- + \frac{1}{c},$$

and

$$\beta_1^* := \beta_1^- + \frac{1}{c}\,\frac{\underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\underset{\sim}{\theta}{}^{(n)}_*(\beta_1^-))}{\underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\underset{\sim}{\theta}{}^{(n)}_*(\beta_1^-)) - \underset{\sim}{\Delta}{}^{(n)}_{f_1;1;\#}(\underset{\sim}{\theta}{}^{(n)}_*(\beta_1^+))}.$$

Then,

$$\mathscr{I}^{(n)}(f_1) := (\beta_1^*)^{-1} = \mathscr{I}(f_1, g_1) + o_{\mathrm{P}}(1)$$

under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, $g_1 \in \underset{\sim}{\mathscr{F}}^*_{f_1}$, as $n\to\infty$. Moreover, $\underset{\sim}{\theta}^{(n)}_{f_1} := \underset{\sim}{\theta}^{(n)}_*(\beta^*_1)$ is an efficient (at $\mathrm{P}^{(n)}_{\theta,\sigma,0;f_1}$—efficiency here is in the parametric sense) R-estimator of θ. A proof for this can be obtained by parallelling that of Section 4.2 in Hallin, Oja, and Paindaveine (2006). The same claim, however, also follows from a more general result by Cassart et al. (2010) which we also need for the estimation of $\mathscr{J}(f_1,g_1)$.

The estimation method just described for $\mathscr{I}(f_1,g_1)$ indeed does not apply to $\mathscr{J}(f_1,g_1)$ which, contrary to $\mathscr{I}(f_1,g_1)$, is not associated with any optimal one-step R-estimation procedure (it does not follow as a coefficient of the covariance, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$, of any component $\underset{\sim}{\Delta}^{(n)}_{f_1;\ell}(\boldsymbol{\vartheta})$ of the rank-based version of the central sequence with the corresponding component $\Delta^{(n)}_{g_1;\ell}(\boldsymbol{\vartheta})$ of $\boldsymbol{\Delta}^{(n)}_{g_1}(\boldsymbol{\vartheta})$). Fortunately, Proposition 2.1 of Cassart et al. (2010) applies, yielding the desired estimator.

That proposition requires the existence of a rank-based statistic $\underset{\sim}{S}^{(n)}(\theta)$ satisfying, under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$ and for all σ, the following two conditions (Assumption (A) of Cassart et al. (2010)).

(D1) The sequence $\underset{\sim}{S}^{(n)}(\theta)$, $n\in\mathbb{N}$ is tight and bounded away from zero, that is, for all $\varepsilon>0$, there exist δ_ε, M_ε and N_ε such that, for all $n>N_\varepsilon$,

$$\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}[\delta_\varepsilon \le |\underset{\sim}{S}^{(n)}(\theta)| \le M_\varepsilon] \ge 1-\varepsilon. \tag{13}$$

(D2) For any θ, there exists $\Upsilon(\theta)\neq 0$ such that, for any bounded sequence $t^{(n)}$,

$$\underset{\sim}{S}^{(n)}(\theta+n^{-1/2}t^{(n)}) = \underset{\sim}{S}^{(n)}(\theta) - t^{(n)}\mathscr{J}(g)\Upsilon^{-1}(\theta) + o_{\mathrm{P}}(1) \quad \text{as } n\to\infty, \tag{14}$$

where the mapping $\theta\mapsto\Upsilon(\theta)$ is continuous.

In the present context, we propose

$$\underset{\sim}{S}^{(n)}(\theta) := \underset{\sim}{S}^{(n)}_{f_1}(\theta) := n^{-1/2}\sum_{i=1}^n s_i^{(n)}(\theta)\phi_{f_1}\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)\Big(F^{-1}_{1+}\Big(\frac{R^{(n)}_{+,i}(\theta)}{n+1}\Big)\Big)^2.$$

It follows from the classical Hájek results for linear signed-rank statistics (see, e.g., Chapter 3 of Puri and Sen 1985) that $\underset{\sim}{S}^{(n)}_{f_1}(\theta)$ is asymptotically centered normal, so that (13) is satisfied, while (14) holds, in view of asymptotic linearity, for $g_1\in\underset{\sim}{\mathscr{F}}^*_{f_1}$, with $\Upsilon(\theta)=\sigma\mathscr{J}^{-1}(f_1,g_1)$.

As in the estimation of $\mathscr{I}(f_1,g_1)$, root-n consistent preliminary estimators $\hat\theta$ and $\widehat\sigma$ of θ and σ are required as well; $\hat\theta$ moreover should be such that (Assumption (B) of Cassart et al. (2010)), under $\mathrm{P}^{(n)}_{\theta,\sigma,0;g_1}$ and for all σ, $\underset{\sim}{S}^{(n)}_{f_1}(\hat\theta)$ is not $o_{\mathrm{P}}(1)$—a condition which is trivially satisfied by the empirical median. All assumptions of Proposition 2.1 of Cassart et al. (2010) then hold, which entails the desired consistency of the estimator we now describe.

Proceeding as in (12) above, let (with the usual discretized versions $\hat\theta_\#$ and $\widehat\sigma_\#$)

$$\underset{\sim}{\theta}^{(n)}_{**}(\beta) := \hat\theta_\# + n^{-1/2}\beta\,\widehat\sigma_\#\underset{\sim}{S}^{(n)}_{f_1;\#}(\hat\theta_\#),$$

$$\beta_2^- := \min\{\beta_\ell \mid \underset{\sim}{S}^{(n)}_{f_1;\#}(\underset{\sim}{\theta}^{(n)}_{**}(\beta_{\ell+1}))\, \underset{\sim}{S}^{(n)}_{f_1;\#}(\hat{\theta}_\#) < 0\}, \quad \text{and} \quad \beta_2^+ := \beta_2^- + \frac{1}{c}.$$

Defining

$$\beta_2^* := \beta_2^- + \frac{1}{c}\, \frac{\underset{\sim}{S}^{(n)}_{f_1;\#}(\underset{\sim}{\theta}^{(n)}_{**}(\beta_2^-))}{\underset{\sim}{S}^{(n)}_{f_1;\#}(\underset{\sim}{\theta}^{(n)}_{**}(\beta_2^-)) - \underset{\sim}{S}^{(n)}_{f_1;\#}(\underset{\sim}{\theta}^{(n)}_{**}(\beta_2^+))},$$

the estimator $\mathscr{J}^{(n)}(f_1) := (\beta_2^*)^{-1}$ then is such that

$$\mathscr{J}^{(n)}(f_1) = \mathscr{J}(f_1, g_1) + o_{\mathrm{P}}(1) \quad \text{under } \mathrm{P}^{(n)}_{\theta,\sigma,0;g_1},\ g_1 \in \underset{\sim}{\mathscr{F}}^*_{f_1}, \ \text{as } n \to \infty,$$

as desired.

4.2 Practical implementation.

As usual, all discretizations in the construction of $\mathscr{I}^{(n)}(f_1)$ and $\mathscr{J}^{(n)}(f_1)$ are required for the purpose of asymptotic statements, but can be dispensed with in applications, where n remains fixed. The practical versions of $\mathscr{I}^{(n)}(f_1)$ and $\mathscr{J}^{(n)}(f_1)$ therefore are

$$\mathscr{I}^{(n)}(f_1) := (\beta_1^*)^{-1} \quad \text{and} \quad \mathscr{J}^{(n)}(f_1) := (\beta_2^*)^{-1},$$

respectively, where

$$\beta_1^* := \inf\left\{\beta > 0 \mid \underset{\sim}{\Delta}^{(n)}_{f_1;1}(\hat{\theta} + n^{-1/2}\beta\widehat{\sigma}^2\underset{\sim}{\Delta}^{(n)}_{f_1;1}(\hat{\theta}))\, \underset{\sim}{\Delta}^{(n)}_{f_1;1}(\hat{\theta}) \le 0\right\}$$

and

$$\beta_2^* := \inf\left\{\beta > 0 \mid \underset{\sim}{S}^{(n)}_{f_1}(\hat{\theta} + n^{-1/2}\beta\widehat{\sigma}\, \underset{\sim}{S}^{(n)}_{f_1}(\hat{\theta}))\, \underset{\sim}{S}^{(n)}_{f_1}(\hat{\theta}) \le 0\right\},$$

and follow from adopting "large" values of the discretizing constants (or, letting c tend to infinity). The ratio $\mathscr{J}^{(n)}(f_1)/\mathscr{I}^{(n)}(f_1)$ then provides the estimator $\underset{\sim}{\kappa}^{(n)}(f_1;\theta)$ of $\underset{\sim}{\kappa}(f_1, g_1)$ required in the definition (11) of the test statistic of Proposition 4.

Acknowledgements Marc Hallin and Davy Paindaveine acknowledge the support of the IAP research network grant P7/06 of the Belgian federal government (Belgian Science Policy (BELSPO)). Davy Paindaveine also acknowledges the support of an A.R.C. contract from the Communauté Française de Belgique.

References

1. Antille, A. (1974). A linearized version of the Hodges-Lehmann estimator. *Annals of Statistics* **2**, 1308-1313.
2. Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew *t* distribution. *Journal of the Royal Statistical Society Series B* **65**, 367-389.
3. Bickel, P.J. and Ritov, Y. (1988). Estimating integrated squared density derivatives. *Sankhya A* **50**, 381-393.
4. Cassart, D., Hallin, M. and Paindaveine, D. (2008). Optimal detection of Fechner-asymmetry. *Journal of Statistical Planning and Inference* **138**, 2499-2525.
5. Cassart, D., Hallin, M. and Paindaveine, D. (2010). On the estimation of cross-information quantities in R-estimation. In J. Antoch, M. Hušková and P.K. Sen, Editors, *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková*, I.M.S. Monographs-Lecture Notes, 35-45.
6. Cassart, D., Hallin, M. and Paindaveine, D. (2011). A class of optimal tests for symmetry based on local Edgeworth approximations. *Bernoulli* **17**, 1063-1094.
7. Cheng, K.F. and Serfling, R.J. (1981). On estimation of a class of efficiency-related parameters. *Scandinavian Actuarial Journal* **8**, 83-92.
8. Fan, J. (1991). On the estimation of quadratic functionals. *Annals of Statistics* **19**, 1273-1294.
9. Edgeworth, F.Y. (1905). The law of error. *Transactions of the Cambridge Philosophical Society* **20**, 36-65 and 113-141.
10. Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, Vol. 2, New York : J. Wiley.
11. Hallin, M., Oja, H., and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape: II Optimal R-estimation of shape. *Annals of Statistics* **34**, 2757-2789.
12. Hallin, M. and Werker, B.J.M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9**, 137-165.
13. Henze, N. (1993). On the consistency of a test for symmetry based on a runs statistic. *Journal of Nonparametric Statistics* **3**, 195-199.
14. Hollander, M. (1988). Testing for symmetry. In Johnson, N.L. and S. Kotz, *Encyclopedia of Statistical Sciences* Vol. 9. J. Wiley, New York, pp. 211-216.
15. Jurečková, J. and Sen, P.K. (1996). *Robust Statistical Procedures : Asymptotics and Interrelations*. New York : Wiley.
16. Koul, H.L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models*, 2nd edition. New York: Springer Verlag.
17. Kraft, C.H. and van Eeden, C. (1972). Linearized rank estimates and signed-rank estimates for the general linear hypothesis. *The Annals of Mathematical Statistics* **43**, 42-57.
18. Lehmann, E.L. (1963). Nonparametric confidence intervals for a shift parameter. *The Annals of Mathematical Statistics* **34**, 1507-1512.
19. Ley, C. and Paindaveine, D. (2009). Le Cam optimal tests for symmetry against Ferreira and Steel's general skewed distributions. *Journal of Nonparametric Statistics* **21**, 943-967.
20. McWilliams, T.P. (1990). A distribution-free test for symmetry based on a runs statistic. *Journal of the American Statistical Association* **85**, 1130-1133.
21. Puri, M.L. and Sen, P.K. (1985). *Nonparametric Methods in General Linear Models.* New York : J. Wiley.
22. Randles, R.H., Fligner, M.A., Policello, G.E., and Wolfe, D.A. (1980). An asymptotically distribution-free test for symmetry versus asymmetry. *Journal of the American Statistical Association* **75**, 168-172.
23. Schweder, T. (1975). Window estimation of the asymptotic variance of rank estimators of location. *Scandinavian Journal of Statistics* **2**, 113-126.
24. Sen, P.K. (1966). On a distribution-free method of estimating asymptotic efficiency of a class of nonparametric tests. *The Annals of Mathematical Statistics* **37**, 1759-1770.