

A class of optimal tests for symmetry based on local Edgeworth approximations

DELPHINE CASSART¹, MARC HALLIN^{1,2,3} and DAVY PAINDAVEINE^{1,4}

¹*Institut de Recherche en Statistique, ECARES, Université Libre de Bruxelles, Brussels, Belgium*

²*Académie Royale de Belgique, Brussels, Belgium*

³*CentER, Tilburg University, Tilburg, The Netherlands*

⁴*Département de Mathématique, Université Libre de Bruxelles, Brussels, Belgium*

The objective of this paper is to provide, for the problem of univariate symmetry (with respect to specified or unspecified location), a concept of optimality, and to construct tests achieving such optimality. This requires embedding symmetry into adequate families of asymmetric (local) alternatives. We construct such families by considering non-Gaussian generalizations 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.

Keywords: Edgeworth expansion; local asymptotic normality; locally asymptotically most powerful tests; skewed densities; tests for symmetry

1. Introduction

1.1. Testing for symmetry

Symmetry is one of the most important and fundamental structural assumptions in statistics, playing a major role, for instance, in the identifiability of location or intercept under nonparametric conditions: see [2,14,15]. This importance explains the huge variety of existing testing procedures of the null hypothesis of symmetry in an i.i.d. sample X_1, \dots, X_n ; see [8] for a survey.

Traditional tests of the null hypothesis of symmetry – the hypothesis under which $X_1 - \theta \stackrel{d}{=} -(X_1 - \theta)$ for some location $\theta \in \mathbb{R}$, with $\stackrel{d}{=}$ standing for equality in distribution – are based on standardized empirical third-order moments. Let $m_k^{(n)}(\theta) := n^{-1} \sum_{i=1}^n (X_i - \theta)^k$ and $m_k^{(n)} := m_k^{(n)}(\bar{X}^{(n)})$, where $\bar{X}^{(n)} := n^{-1} \sum_{i=1}^n X_i$. When the location θ is specified, the test statistic is

$$S_1^{(n)}(\theta) := n^{1/2} m_3^{(n)}(\theta) / (m_6^{(n)}(\theta))^{1/2}, \quad (1.1)$$

the null distribution of which, under finite sixth-order moments, is asymptotically standard normal. When θ is unspecified, the classical test is based on the empirical coefficient of skewness

$$b_1^{(n)} := m_3^{(n)} / s_n^3, \quad (1.2)$$

where $s_n := (m_2^{(n)})^{1/2}$ stands for the empirical standard error in a sample of size n . More precisely, this test relies on the asymptotic standard normal distribution (still under finite moments of order six) of

$$S_2^{(n)} := n^{1/2} m_3^{(n)} / (m_6^{(n)} - 6s_n^2 m_4^{(n)} + 9s_n^6)^{1/2}, \quad (1.3)$$

which, under Gaussian densities, asymptotically reduces to $\sqrt{n/6} b_1^{(n)}$.

These two tests are generally considered as Gaussian procedures, although they do not require any Gaussian assumptions and despite the fact that none of them can be considered optimal in any Gaussian sense, since asymmetric alternatives clearly cannot belong to a Gaussian universe. Despite the long history of the problem, the optimality features of those classical procedures thus are all but clear, and optimality issues, in that fundamental problem, remain essentially unexplored.

The main objective of this paper is to provide this classical testing problem with a concept of optimality that confirms practitioners' intuition (i.e., justifying the $b_1^{(n)}$ -based Gaussian practice), and to construct tests achieving such optimality. This requires embedding the null hypothesis of symmetry into adequate families of asymmetric alternatives. We therefore define local (in the LeCam sense) alternatives indexed by location, scale and a measure of skewness in such a way that:

- (i) Location, scale, and skewness play well-separated roles (diagonality of the corresponding information matrices).
- (ii) The traditional tests based on $b_1^{(n)}$ (more precisely, based on $S_2^{(n)}$ given in (1.3)) become locally and asymptotically optimal in the vicinity of Gaussian densities.

As we shall see, part (ii) of this objective is achieved by considering local first-order Edgeworth approximations of the form

$$\phi(x - \theta) + n^{-1/2} \xi(x - \theta) \phi(x - \theta) ((x - \theta)^2 - \kappa), \quad (1.4)$$

where ϕ as usual stands for the standard normal density, $\kappa (= 3)$ is the Gaussian kurtosis coefficient, θ is a location parameter, and ξ a measure of skewness. Adequate modifications of (1.4), playing similar roles in the vicinity of non-Gaussian standardized symmetric reference densities f_1 , are proposed in (2.2).

The resulting tests of symmetry (for specified as well as for unspecified location θ) are valid under a broad class of symmetric densities, and parametrically efficient at the reference (standardized) density f_1 . Of particular interest are the pseudo-Gaussian tests (associated with a Gaussian reference density; see Proposition 3.6), which require finite moments of order six and appear to be asymptotically equivalent (under their specified- θ version as well as under the unspecified- θ one) to the test (1.3) based on $b_1^{(n)}$, and the Laplace tests (associated with a double-exponential reference density; see Proposition 3.6), which only require moments of order four and are closely related with the tests against Fechner asymmetry derived in [3].

These tests are of a parametric nature. Since the null hypothesis of symmetry enjoys a rich group invariance structure, classical maximal invariance arguments naturally bring *signs* and *signed ranks* into the picture. Such a nonparametric approach is adopted in a companion paper [4], where we construct signed-rank versions of the parametrically efficient tests proposed here.

These signed-rank tests are distribution-free (asymptotically so in case of an unspecified location θ) under the null hypothesis of symmetry, and therefore remain valid under much milder distributional assumptions (for the specified location case, they are valid in the absence of *any* distributional assumption).

The main technical tool throughout the paper is LeCam's asymptotic theory of statistical experiments and the properties of locally asymptotically normal (LAN) families. LAN has become a standard tool in asymptotics: see [12] or Chapters 6–9 of [17] for details. Log-likelihoods in a LAN family with k -dimensional parameter ϑ admit local quadratic approximations of the form $\tau' \Delta_{\vartheta}^{(n)} - \frac{1}{2} \tau' \Gamma_{\vartheta} \tau$, where the random vector $\Delta_{\vartheta}^{(n)}$, called a *central sequence*, is asymptotically normal $\mathcal{N}(\Gamma_{\vartheta} \tau, \Gamma_{\vartheta})$ under sequences of parameter values of the form $\vartheta + n^{-1/2} \tau$ (*local alternatives*). Let $\phi^*(\Delta)$ be an optimal test (uniformly most powerful, maximin, most stringent, ...) in the *Gaussian shift model* describing a hypothetical observation Δ with distribution in the family $\{\mathcal{N}(\Gamma_{\vartheta} \tau, \Gamma_{\vartheta}) \mid \tau \in \mathbb{R}^k\}$ (Γ_{ϑ} specified). Those families are extremely simple, and optimal tests in that context are well known (see, e.g., Section 11.9 of [11]). Then, the sequence $\phi^*(\Delta_{\vartheta}^{(n)})$ is a sequence of *locally asymptotically optimal* (locally asymptotically *uniformly most powerful, maximin, most stringent, ...*) tests for the original problem – where optimality is based on the local convergence of risk functions to the risk functions of Gaussian shift experiments. More analytical characterizations can be found, for instance, in [5].

1.2. Outline of the paper

The problem we are considering throughout is that of testing the null hypothesis of symmetry. In the notation of Section 1.1, ξ (see (2.2) 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 describe the Edgeworth-type families of local alternatives, extending (1.4), that we are considering. Section 2.2 establishes the local and asymptotic normality (with respect to location, scale and the asymmetry parameters) result to be used throughout; actually, we establish a slightly stronger version of LAN, called ULAN (uniform LAN), which allows us to handle the problems related with estimated nuisance parameters. The classical LeCam theory then is used in Section 3.1 for developing asymptotically optimal procedures for testing symmetry ($\xi = 0$), with specified or unspecified location θ but specified standardized symmetric density f_1 . The more realistic case of an unspecified f_1 is treated in Section 3.2, where we obtain versions of the optimal (at given f_1) tests that remain valid under $g_1 \neq f_1$, for specified (Section 3.2.1) and unspecified (Section 3.2.2) location θ , respectively. The particular case of pseudo-Gaussian procedures (which are optimal for Gaussian f_1 but valid under any symmetric density with finite moments of order six) is studied in detail in Section 3.3 and their relation with classical tests of symmetry is discussed. We also show that the Laplace tests (which are optimal for double-exponential f_1 but valid under any symmetric density with finite fourth-order moment) are closely related to the Fechner-type tests derived in [3]. The finite-sample performances of these tests are investigated via simulations in Section 4, where they are applied to the classical skew-normal and skew- t densities.

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

2.1. Families of asymmetric densities based on Edgeworth approximations

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

- (a) The hypothesis $\mathcal{H}_\theta^{(n)}$ of symmetry with respect to specified location $\theta \in \mathbb{R}$: under $\mathcal{H}_\theta^{(n)}$, the X_i 's have density function

$$x \mapsto f(x) := \sigma^{-1} f_1((x - \theta)/\sigma) \tag{2.1}$$

(all densities are over the real line, with respect to the Lebesgue measure) for some unspecified $\sigma \in \mathbb{R}_0^+$, where f_1 belongs to the class of standardized symmetric densities

$$\mathcal{F}_0 := \left\{ h_1 : h_1(-z) = h_1(z) \text{ for all } z \in \mathbb{R} \text{ and } \int_{-\infty}^1 h_1(z) dz = 0.75 \right\}.$$

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

- (b) The hypothesis $\mathcal{H}^{(n)} := \bigcup_{\theta \in \mathbb{R}} \mathcal{H}_\theta^{(n)}$ of symmetry with respect to unspecified location (and scale): there exist (θ, σ) such that the X_i 's have density (2.1).

In both cases, the standardized density f_1 may be specified (Section 3.1) or not (Sections 3.2–3.4). The specified- f_1 problem, however, mainly serves as a preparation for the more realistic unspecified- f_1 one.

As explained in the introduction, a characterization of efficient testing requires the definition of families of asymmetric alternatives exhibiting some adequate structure, such as local asymptotic normality, at the null. 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 \mathcal{F}_1 of densities f_1 satisfying:

- (i) (symmetry and standardization) $f_1 \in \mathcal{F}_0$;
- (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;
- (iv) (finite Fisher information) $\mathcal{K}(f_1) := \int_{-\infty}^{+\infty} z^4 \phi_{f_1}^2(z) f_1(z) dz$, hence also, under strong unimodality,

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

- are finite;
- (v) (polynomial tails)

$$\int_y^\infty f_1(z) dz = O(y^{-\beta}) \quad \text{as } y \rightarrow \infty$$

for some $\beta > 0$ and

$$\phi_{f_1}(z) = o(z^{\beta/2-2}) \quad \text{as } z \rightarrow \infty.$$

That class \mathcal{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 $\mathcal{I}(f_1)$ and $\mathcal{J}(f_1)$ for location and scale, and, as we shall see, $\mathcal{K}(f_1)$ for asymmetry, with tails satisfying (v).

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

$$f(x) = \sigma^{-1} f_1\left(\frac{x-\theta}{\sigma}\right) - \xi \sigma^{-1} \dot{f}_1\left(\frac{x-\theta}{\sigma}\right) \left(\left(\frac{x-\theta}{\sigma}\right)^2 - \kappa(f_1) \right) I[|x-\theta| \leq \sigma|z^*|] \\ - \text{sign}(\xi) \sigma^{-1} f_1\left(\frac{x-\theta}{\sigma}\right) \{ I[x-\theta > \text{sign}(-\xi)\sigma|z^*|] - I[x-\theta < \text{sign}(\xi)\sigma|z^*|] \}. \tag{2.2}$$

Here $\theta \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ clearly are location and scale parameters, $\xi \in \mathbb{R}$ is a measure of skewness, $\kappa(f_1)$ (strictly positive for $f_1 \in \mathcal{F}_1$) the generalized kurtosis coefficient just defined and z^* the unique (for ξ small enough; unicity follows from the monotonicity of ϕ_{f_1}) solution of $f_1(z^*) = \xi \dot{f}_1(z^*)((z^*)^2 - \kappa(f_1))$. The function f defined in (2.2) is indeed a probability density (non-negative, 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, $x \mapsto f(x)$ is continuous whenever $\dot{f}_1(x)$ is; vanishes for $x \leq \theta + \sigma z^*$ if $\xi > 0$, for $x \geq \theta + \sigma z^*$ 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 $\xi \rightarrow 0$.

The intuition behind this class of alternatives is that, in the Gaussian case, (2.2), with $\xi = n^{-1/2}\tau$ yields (for $x \in [\theta \pm \sigma z^*]$) the first-order Edgeworth development 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 “small” value of the asymmetry parameter ξ , of the form $n^{-1/2}\tau$, (2.2) 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 shocks, the locally asymmetric f results from the same additive combination of independent but slightly skew shocks. As we shall see, the locally optimal test in such a case happens to be the traditional test based on $b_1^{(n)}$ (see (1.2)).

Besides the Gaussian one (with standardized density $\phi_1(z) := \sqrt{a/2\pi} \exp(-az^2/2)$), interesting special cases of (2.2) are obtained in the vicinity of:

- (i) The double-exponential or Laplace distributions, with standardized density

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

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

- (ii) The logistic distributions, with standardized density

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

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

- (iii) The power-exponential distributions, with standardized densities

$$f_1(z) = f_{\text{exp}_\eta}(z) := C_{\text{exp}_\eta} \exp(-(g_\eta z)^{2\eta}),$$

$$\eta \in \mathbb{N}_0, \mathcal{I}(f_1) = 2g_\eta^2 \eta \Gamma(2 - 1/2\eta) / \Gamma(1 + 1/2\eta), \mathcal{J}(f_1) = 1 + 2\eta \text{ and } \mathcal{K}(f_1) = 2g_\eta \eta / \Gamma(1 + 1/2\eta) \text{ (the positive constants } C_{\text{exp}_\eta}, a, b, d \text{ and } g_\eta \text{ are such that } f_1 \in \mathcal{F}_1).$$

Although not strongly unimodal, the Student distributions with $\nu > 2$ degrees of freedom also can be considered here (strong unimodality indeed is essentially used as a sufficient condition for the existence of z^* in (2.2) – an existence that can be checked directly in the Student case). Standardized Student densities take the form

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

with $\mathcal{I}(f_1) = a_\nu(\nu + 1)/(\nu + 3)$, $\mathcal{J}(f_1) = 3(\nu + 1)/(\nu + 3)$ and $\mathcal{K}(f_1) = 15\nu(\nu + 1)/a_\nu(\nu - 2)(\nu + 3)$ (C_{t_ν} and a_ν are normalizing constants). Note that the corresponding Gaussian values, namely $\mathcal{I}(\phi_1) = a \approx 0.4549$, $\mathcal{J}(\phi_1) = 3$ and $\mathcal{K}(\phi_1) = 15/a$, are obtained by taking limits as $\nu \rightarrow \infty$.

Figures 1 and 2 provide graphical representations of some densities in the Gaussian ($f_1 = \phi_1$) and double-exponential ($f_1 = f_{\mathcal{L}}$) Edgeworth families (2.2), respectively. In the Gaussian case, the skewed densities are continuous, while the double-exponential ones, due to the discontinuity of $\hat{f}_{\mathcal{L}}(x)$ at $x = 0$, exhibit 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 $(\theta, \sigma, 0)'$, of the parametric families

$$\mathcal{P}_{f_1}^{(n)} := \bigcup_{\sigma > 0} \mathcal{P}_{\sigma; f_1}^{(n)} := \bigcup_{\sigma > 0} \{P_{\theta, \sigma, \xi; f_1}^{(n)} \mid \theta \in \mathbb{R}, \xi \in \mathbb{R}\}, \tag{2.3}$$

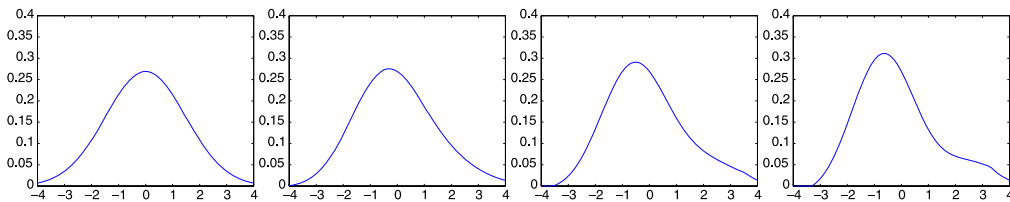


Figure 1. Graphical representation of the Gaussian Edgeworth family (2.2) ($f_1 = \phi_1$), for $\xi = 0, 0.05, 0.10$ and 0.15 .

where $f_1 \in \mathcal{F}_1$. More precisely, the following result holds (see the [Appendix](#) for proof).

Proposition 2.1 (ULAN). For any $f_1 \in \mathcal{F}_1$, $\theta \in \mathbb{R}$, and $\sigma \in \mathbb{R}_0^+$, the family $\mathcal{P}_{f_1}^{(n)}$ is ULAN at $(\theta, \sigma, 0)'$, with (writing Z_i for $Z_i^{(n)}(\theta, \sigma) := \sigma^{-1}(X_i^{(n)} - \theta)$ and ϕ_{f_1} for $-\dot{f}_1/f_1$) central sequence

$$\begin{aligned} \Delta_{f_1}^{(n)}(\boldsymbol{\vartheta}) &=: \begin{pmatrix} \Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;2}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}) \end{pmatrix} \\ &= n^{-1/2} \sum_{i=1}^n \begin{pmatrix} \sigma^{-1} \phi_{f_1}(Z_i) \\ \sigma^{-1} (\phi_{f_1}(Z_i) Z_i - 1) \\ \phi_{f_1}(Z_i) (Z_i^2 - \kappa(f_1)) \end{pmatrix} \end{aligned} \tag{2.4}$$

and full-rank information matrix

$$\Gamma_{f_1}(\boldsymbol{\vartheta}) = \begin{pmatrix} \sigma^{-2} \mathcal{I}(f_1) & 0 & 0 \\ 0 & \sigma^{-2} (\mathcal{J}(f_1) - 1) & 0 \\ 0 & 0 & \gamma(f_1) \end{pmatrix}, \tag{2.5}$$

where $\gamma(f_1) := \mathcal{K}(f_1) - \mathcal{J}^2(f_1)/\mathcal{I}(f_1)$.

More precisely, for any $\boldsymbol{\vartheta}^{(n)} := (\theta^{(n)}, \sigma^{(n)}, 0)'$ such that $\theta^{(n)} - \theta = O(n^{-1/2})$ and $\sigma^{(n)} - \sigma = O(n^{-1/2})$, and for any bounded sequence $\boldsymbol{\tau}^{(n)} = (t^{(n)}, s^{(n)}, \tau^{(n)})' \in \mathbb{R}^3$, we have, under $\mathbf{P}_{\boldsymbol{\vartheta}^{(n)}; f_1}^{(n)}$,

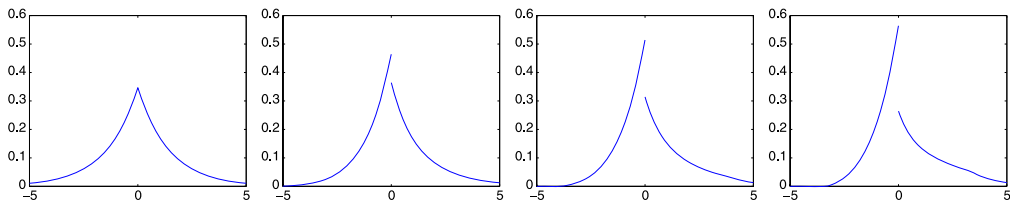


Figure 2. Graphical representation of the double-exponential Edgeworth family (2.2) ($f_1 = f_{\mathcal{L}}$), for $\xi = 0, 0.05, 0.10$ and 0.15 .

as $n \rightarrow \infty$,

$$\begin{aligned} \Lambda_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)}/f_1; f_1}^{(n)} &:= \log \left(\frac{d\mathbf{P}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)}/f_1}^{(n)}}{d\mathbf{P}_{\boldsymbol{\vartheta}^{(n)}/f_1}^{(n)}} \right) \\ &= \boldsymbol{\tau}^{(n)'} \boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta}^{(n)}) - \frac{1}{2} \boldsymbol{\tau}^{(n)'} \boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}) \boldsymbol{\tau}^{(n)} + o_{\mathbb{P}}(1) \end{aligned}$$

and

$$\boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta}^{(n)}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})).$$

The diagonal form of the information matrix $\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})$ confirms that location, scale and skewness, in the parametric family (2.3), play distinct and well-separated roles. The practical consequences of that orthogonality (in the sense of information) property, as we shall see, are twofold:

- (a) The fact that location and scale are unspecified has no cost in terms of efficiency and under specified standardized density f_1 when testing for symmetry.
- (b) Substituting root- n -consistent (and, in principle, duly *discretized*: see assumption (C2) below) estimators for the true values has no impact on the asymptotic validity and local powers, under specified standardized density f_1 , of tests for symmetry.

Note that orthogonality between the scale and skewness components of $\boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta})$ 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 (2.4) and (2.5) are

$$\boldsymbol{\Delta}_{\phi_1}^{(n)}(\boldsymbol{\vartheta}) = n^{-1/2} \sum_{i=1}^n \begin{pmatrix} a\sigma^{-1}Z_i \\ \sigma^{-1}(aZ_i^2 - 1) \\ aZ_i \left(Z_i^2 - \frac{3}{a} \right) \end{pmatrix}$$

and

$$\boldsymbol{\Gamma}_{\phi_1}(\boldsymbol{\vartheta}) = \begin{pmatrix} a\sigma^{-2} & 0 & 0 \\ 0 & 2\sigma^{-2} & 0 \\ 0 & 0 & 6/a \end{pmatrix},$$

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

3. Locally asymptotically optimal tests

The various test statistics described in this section are listed, for easy reference, in Table 3, placed at the end of Section 5.

3.1. Locally asymptotically optimal tests: Specified density

For specified $f_1 \in \mathcal{F}_1$, consider the null hypothesis

$$\mathcal{H}_{\theta; f_1}^{(n)} := \bigcup_{\sigma \in \mathbb{R}_0^+} \{P_{\theta, \sigma, 0; f_1}^{(n)}\}$$

of symmetry with respect to some specified location θ , and the null hypothesis

$$\mathcal{H}_{f_1}^{(n)} := \bigcup_{\theta \in \mathbb{R}} \bigcup_{\sigma \in \mathbb{R}_0^+} \{P_{\theta, \sigma, 0; f_1}^{(n)}\}$$

of symmetry with respect to unspecified θ . ULAN and the diagonal structure of (2.5) imply that substituting discretized root- n -consistent estimators $\hat{\theta}$ and $\hat{\sigma}$ for the unknown θ and σ has no influence, asymptotically, on the ξ -part of the central sequence.

Recall that a sequence of estimators $\hat{\lambda}^{(n)}$ defined in a sequence of experiments $\{P_\lambda^{(n)} \mid \lambda \in \Lambda\}$ indexed by some parameter λ is *root- n -consistent* and *asymptotically discrete* if, under $P_\lambda^{(n)}$, as $n \rightarrow \infty$:

(C1) $\hat{\lambda}^{(n)} - \lambda = O_P(n^{-1/2})$.

(C2) The number of possible values of $\hat{\lambda}^{(n)}$ in balls with $O(n^{-1/2})$ radius centered at λ is bounded as $n \rightarrow \infty$.

An estimator $\lambda^{(n)}$ satisfying (C1) but not (C2) is easily discretized by letting, for some arbitrary constant $c > 0$, $\lambda_\#^{(n)} := (cn^{1/2})^{-1} \text{sign}(\lambda^{(n)}) \lceil cn^{1/2} |\lambda^{(n)}| \rceil$, which satisfies both (C1) and (C2). Subscripts $\#$ in the sequel are used for estimators $(\hat{\theta}_\#, \hat{\sigma}_\#, \dots)$ 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.

It follows from the diagonal form of the information matrix (2.5) that locally uniformly asymptotically most powerful tests of $\mathcal{H}_{\theta; f_1}^{(n)}$ (resp., of $\mathcal{H}_{f_1}^{(n)}$) can be based on $\Delta_{f_1; 3}^{(n)}(\theta, \hat{\sigma}_\#, 0)$ (resp., on $\Delta_{f_1; 3}^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#, 0)$), hence on $T_{f_1}^{(n)}(\theta, \hat{\sigma}_\#)$ (resp., on $T_{f_1}^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#)$), where

$$T_{f_1}^{(n)}(\theta, \sigma) := \frac{1}{\sqrt{n\gamma(f_1)}} \sum_{i=1}^n \phi_{f_1}(Z_i(\theta, \sigma))(Z_i^2(\theta, \sigma) - \kappa(f_1)). \tag{3.1}$$

Root- n -consistent (under the null hypothesis of symmetry) estimators of θ and σ that do not require any moment assumptions are, for instance, the medians $\hat{\theta} := \text{Med}(X_i^{(n)})$ and $\hat{\sigma} := \text{Med}(|X_i^{(n)} - \hat{\theta}|)$ of the $X_i^{(n)}$'s and of their absolute deviations from $\hat{\theta}$, respectively.

The following proposition then results from classical results on ULAN families (see, e.g., Chapter 11 of [11]).

Proposition 3.1. *Let $f_1 \in \mathcal{F}_1$. Then:*

- (i) $T_{f_1}^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#) = T_{f_1}^{(n)}(\theta, \sigma) + \text{op}(1)$ is asymptotically normal, with mean zero under $\mathbb{P}_{\theta, \sigma, 0; f_1}^{(n)}$, mean $\tau\gamma^{1/2}(f_1)$ under $\mathbb{P}_{\theta, \sigma, n^{-1/2}\tau; f_1}^{(n)}$ and variance one under both.
- (ii) The sequence of tests rejecting the null hypothesis of symmetry (with standardized density f_1) whenever $T_{f_1}^{(n)}(\theta, \hat{\sigma}_\#)$ (resp., $T_{f_1}^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#)$) exceeds the $(1 - \alpha)$ standard normal quantile z_α is locally asymptotically most powerful at asymptotic level α for $\mathcal{H}_{\theta; f_1}^{(n)}$ (resp., for $\mathcal{H}_{f_1}^{(n)}$) against $\bigcup_{\xi > 0} \bigcup_{\sigma \in \mathbb{R}_+^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_1}^{(n)}\}$ (resp., $\bigcup_{\xi > 0} \bigcup_{\theta \in \mathbb{R}} \bigcup_{\sigma \in \mathbb{R}_+^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_1}^{(n)}\}$).

This confirms that unspecified location θ and scale σ do not induce any loss of efficiency when the standardized density f_1 itself is specified.

The Gaussian version of (3.1) is

$$T_{\phi_1}^{(n)}(\theta, \sigma) := \sqrt{\frac{a^3}{6n}} \sum_{i=1}^n Z_i(\theta, \sigma) \left(Z_i^2(\theta, \sigma) - \frac{3}{a} \right) = \sqrt{\frac{a}{6n}} \sum_{i=1}^n (aZ_i^3(\theta, \sigma) - 3Z_i(\theta, \sigma));$$

thanks to the linearity of Gaussian scores, it easily follows from a traditional Slutsky argument that $\hat{\theta}$ and $\hat{\sigma}$ in $T_{\phi_1}^{(n)}(\hat{\theta}, \hat{\sigma})$ need not be discretized. Under Gaussian densities, both $T_{\phi_1}^{(n)}(\hat{\theta}, \hat{\sigma})$ and $T_{\phi_1}^{(n)}(\theta, \hat{\sigma})$ are asymptotically equivalent to $T_{\phi_1}^{(n)}(\bar{X}^{(n)}, \hat{\sigma}) = (na^3/6)^{1/2} m_3^{(n)}/\hat{\sigma}^3 = \sqrt{n/6} b_1^{(n)} + \text{op}(1)$, that is, to $S_2^{(n)}$ given in (1.3). The latter is thus locally asymptotically optimal under Gaussian assumptions, whether θ is specified or not, whereas the specified- θ test based on $m_3^{(n)}(\theta)/(m_6^{(n)}(\theta))^{1/2}$ (more precisely, on $S_1^{(n)}(\theta)$ given in (1.1)) is suboptimal. The fact that $m_3^{(n)}(\hat{\theta})$ yields a better performance than $m_3^{(n)}(\theta)$ under specified location θ (see the comments after Proposition 3.5 for a comparison of local powers) looks puzzling at first sight. The reason is that orthogonality, in the Fisher information sense, between asymmetry and location, is a “built-in” feature of Edgeworth families. Contrary to $S_2^{(n)}$, which is shift invariant, $m_3^{(n)}(\theta)$ and $S_1^{(n)}(\theta)$ are sensitive to location shifts. Therefore, tests based on $S_1^{(n)}(\theta)$ are “wasting away” some power on location alternatives (which are irrelevant when θ is specified), to the detriment of asymmetry alternatives.

Locally asymptotically maximin two-sided tests are easily derived along the same lines.

3.2. Locally asymptotically optimal tests: Unspecified density

The tests based on (3.1) achieve local and asymptotic optimality at correctly specified f_1 , which sets the parametric efficiency bounds for the problem, but has limited practical value, as these tests are not valid anymore under density $g_1 \neq f_1$. If Proposition 3.1 is to be adapted to the more realistic null hypotheses $\mathcal{H}_\theta^{(n)} := \bigcup_{g_1} \mathcal{H}_{\theta; g_1}^{(n)}$ and $\mathcal{H}^{(n)} := \bigcup_{g_1} \mathcal{H}_{g_1}^{(n)}$ under which the (symmetric) density remains unspecified, the test statistic $T_{f_1}^{(n)}$ needs to be adapted in order to cope with three problems: its centering under the null and $g_1 \neq f_1$ ($T_{f_1}^{(n)}$ in the previous section only had to be centered under density $g_1 = f_1$), its scaling (same remark) and (still under the null and $g_1 \neq f_1$) the impact of the substitution of estimators $\hat{\sigma}$ (and $\hat{\theta}$) for the unspecified values of σ (and θ) on its asymptotic distribution.

3.2.1. Specified location

Let us first assume that both θ and σ are specified. The test statistic $T_{f_1}^{(n)}$ in Proposition 3.1 is essentially a scaled version of $\Delta_{f_1;3}^{(n)}$. More generally, write $\Delta_{f_1;3}^{(n)}(\kappa)$ for $n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_i)(Z_i^2 - \kappa)$, where $\kappa \in \mathbb{R}_0^+$ denotes a strictly positive real number; an adequate sample-based value will be selected later on. Note that $\Delta_{f_1;3}^{(n)}(\kappa)$ remains centered under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$, irrespective of the choice of κ . Indeed, the functions $z \mapsto \phi_{f_1}(z)z^2$ and $z \mapsto \phi_{f_1}(z)$ are skew symmetric, and their expectations under any symmetric density are automatically zero – provided that they exist. The variance under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$ of $\Delta_{f_1;3}^{(n)}(\kappa)$ is then

$$\gamma_{g_1}^\kappa(f_1) := E_{g_1}[(\phi_{f_1}(Z_i)(Z_i^2 - \kappa))^2] = \mathcal{K}_{g_1}(f_1) - 2\kappa \mathcal{J}_{g_1}(f_1) + \kappa^2 \mathcal{I}_{g_1}(f_1),$$

where

$$\mathcal{I}_{g_1}(f_1) := \int_{-\infty}^{\infty} \phi_{f_1}^2(z)g_1(z) dz, \quad \mathcal{J}_{g_1}(f_1) := \int_{-\infty}^{\infty} z^2 \phi_{f_1}^2(z)g_1(z) dz$$

and (still, provided that those integrals exist)

$$\mathcal{K}_{g_1}(f_1) := \int_{-\infty}^{+\infty} z^4 \phi_{f_1}^2(z)g_1(z) dz.$$

We know from LeCam’s third lemma that, under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$, the impact on $\Delta_{f_1;3}^{(n)}(\kappa)$ of an estimated scale depends on the asymptotic joint distribution (under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$) of $\Delta_{f_1;3}^{(n)}(\kappa)$ and $\Delta_{g_1;2}^{(n)}$. More precisely, LeCam’s third lemma (see, e.g., page 90 of [17]) tells us that if, under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$,

$$\left(\log \left(\frac{\Delta_{f_1;3}^{(n)}(\kappa)}{\mathbb{d}\mathbb{P}_{\theta,\sigma+n^{-1/2}\tau,0;g_1}^{(n)}} \right) \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left(\left(\begin{matrix} \mu_1 \\ -\frac{1}{2}d_2^2 \end{matrix} \right), \begin{pmatrix} d_1^2 & d_{11} \\ d_{11} & d_2^2 \end{pmatrix} \right) \quad \text{as } n \rightarrow \infty, \quad (3.2)$$

then, under $\mathbb{P}_{\theta,\sigma+n^{-1/2}\tau,0;g_1}^{(n)}$, $\Delta_{f_1;3}^{(n)}(\kappa) \xrightarrow{\mathcal{L}} \mathcal{N}(\mu_1 + d_{11}, d_1^2)$. In the present context, Proposition 2.1 yields that

$$\log \left(\frac{\mathbb{d}\mathbb{P}_{\theta,\sigma+n^{-1/2}\tau,0;g_1}^{(n)}}{\mathbb{d}\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}} \right) = \tau \Delta_{g_1;2}^{(n)}(\theta, \sigma, 0) - \frac{1}{2} \tau^2 \sigma^{-2} (\mathcal{J}(g_1) - 1) + o_{\mathbb{P}}(1) \quad (3.3)$$

as $n \rightarrow \infty$, under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$; hence (still as $n \rightarrow \infty$ under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$),

$$\begin{aligned} \left(\begin{matrix} \Delta_{f_1;3}^{(n)}(\kappa) \\ \Delta_{g_1;2}^{(n)}(\theta, \sigma, 0) \end{matrix} \right) &= n^{-1/2} \sum_{i=1}^n \left(\begin{matrix} \phi_{f_1}(Z_i)(Z_i^2 - \kappa) \\ \sigma^{-1}(\phi_{g_1}(Z_i)Z_i - 1) \end{matrix} \right) \\ &\xrightarrow{\mathcal{L}} \mathcal{N} \left(\left(\begin{matrix} \mathcal{J}_{g_1}(f_1) \\ 0 \end{matrix} \right), \begin{pmatrix} d_1^2 & d_{11} \\ d_{11} & d_2^2 \end{pmatrix} \right), \end{aligned} \quad (3.4)$$

with $d_2^2 = \sigma^{-2}(\mathcal{J}(g_1) - 1)$ and $d_{11} = \tau\sigma^{-1} \int_{-\infty}^{\infty} \phi_{f_1}(z)(z^2 - \kappa)(\phi_{g_1}(z)z - 1)g_1(z) dz = 0$ (as the integral of a skew-symmetric function). We conclude that (3.2) holds with $d_{11} = 0$.

LeCam’s third lemma therefore shows that the effect on the asymptotic distribution of $\Delta_{f_1;3}^{(n)}(\kappa)$ of a $O(n^{-1/2})$ perturbation of σ is asymptotically nil; the asymptotic linearity result of Proposition A.1 and a classical argument on asymptotically discrete estimators (see, e.g., Lemma 4.4 in [10]) allow for extending this conclusion to the stochastic $O_P(n^{-1/2})$ perturbations induced by substituting a duly discretized root- n -consistent estimator $\hat{\sigma}_{\#}^{(n)}$ for σ . Such a substitution consequently does not affect the asymptotic behavior of $\Delta_{f_1;3}^{(n)}(\kappa)$.

For $f_1 \in \mathcal{F}_1$ and $g_1 \in \mathcal{F}_{f_1} := \{g_1 \in \mathcal{F}_1 : \mathcal{K}_{g_1}(f_1) < \infty\}$ (due to strong unimodality, $\mathcal{K}_{g_1}(f_1) < \infty$ also implies $\mathcal{I}_{g_1}(f_1) < \infty$ and $\mathcal{J}_{g_1}(f_1) < \infty$), let

$$\gamma^{(n)}(f_1) = \gamma^{(n)}(f_1, \theta, \sigma) := \mathcal{K}^{(n)}(f_1) - 2\kappa(f_1)\mathcal{J}^{(n)}(f_1) + \kappa^2(f_1)\mathcal{I}^{(n)}(f_1), \tag{3.5}$$

where

$$\mathcal{I}^{(n)}(f_1) = \mathcal{I}^{(n)}(f_1, \theta, \sigma) := n^{-1} \sum_{i=1}^n \phi_{f_1}^2(Z_i(\theta, \sigma)), \tag{3.6}$$

$$\mathcal{J}^{(n)}(f_1) = \mathcal{J}^{(n)}(f_1, \theta, \sigma) := n^{-1} \sum_{i=1}^n Z_i^2(\theta, \sigma)\phi_{f_1}^2(Z_i(\theta, \sigma)) \tag{3.7}$$

and

$$\mathcal{K}^{(n)}(f_1) = \mathcal{K}^{(n)}(f_1, \theta, \sigma) := n^{-1} \sum_{i=1}^n Z_i^4(\theta, \sigma)\phi_{f_1}^2(Z_i(\theta, \sigma)) \tag{3.8}$$

under $P_{\theta, \sigma, 0; g_1}^{(n)}$ are consistent estimates of $\mathcal{I}_{g_1}(f_1)$, $\mathcal{J}_{g_1}(f_1)$ and $\mathcal{K}_{g_1}(f_1)$, respectively. Now in practice, $\mathcal{I}^{(n)}(f_1)$, $\mathcal{J}^{(n)}(f_1)$ and $\mathcal{K}^{(n)}(f_1)$, hence $\gamma^{(n)}(f_1)$, cannot be computed from the observations and $Z_i(\theta, \hat{\sigma}_{\#})$ is to be substituted for $Z_i(\theta, \sigma)$ in (3.6)–(3.8), yielding $\gamma^{(n)}(f_1, \theta, \hat{\sigma}_{\#})$. This substitution in general requires a slight reinforcement of regularity assumptions. Along the same lines as above (LeCam’s third lemma and asymptotic linearity), we easily obtain that $\gamma^{(n)}(f_1, \theta, \hat{\sigma}_{\#}) - \gamma^{(n)}(f_1, \theta, \sigma)$ is $o_P(1)$ under $P_{\theta, \sigma, 0; g_1}^{(n)}$ provided that the asymptotic covariance of $\gamma^{(n)}(f_1, \theta, \sigma)$ and $\Delta_{g_1;2}^{(n)}$ is finite. A simple computation (and the strong unimodality of f_1 and g_1) shows that a sufficient condition for this is

$$g_1 \in \mathcal{F}_{f_1}^* := \left\{ h_1 \in \mathcal{F}_{f_1} : \int_{-\infty}^{\infty} z^5 \phi_{f_1}^2(z)\phi_{h_1}(z)h_1(z) dz < \infty \right\}. \tag{3.9}$$

Defining the test statistic

$$\hat{T}_{f_1}^{(n)}(\theta, \sigma) := \frac{1}{\sqrt{n\gamma^{(n)}(f_1, \theta, \sigma)}} \sum_{i=1}^n \phi_{f_1}(Z_i(\theta, \sigma))(Z_i^2(\theta, \sigma) - \kappa(f_1)) \tag{3.10}$$

and the cross-information quantities

$$\begin{aligned} \mathcal{I}_{g_1}(f_1, g_1) &:= \int_{-\infty}^{+\infty} \phi_{f_1}(z)\phi_{g_1}(z)g_1(z) dz, \\ \mathcal{J}_{g_1}(f_1, g_1) &:= \int_{-\infty}^{+\infty} z^2\phi_{f_1}(z)\phi_{g_1}(z)g_1(z) dz \end{aligned}$$

and

$$\mathcal{K}_{g_1}(f_1, g_1) := \int_{-\infty}^{+\infty} z^4\phi_{f_1}(z)\phi_{g_1}(z)g_1(z) dz$$

(which for $f_1 \in \mathcal{F}_1$ and $g_1 \in \mathcal{F}_{f_1}^*$ are finite because of Cauchy–Schwarz), we thus have the following result.

Lemma 3.1. *Let $f_1 \in \mathcal{F}_1$ and $g_1 \in \mathcal{F}_{f_1}^*$. Then:*

- (i) $\hat{T}_{f_1}^{(n)}(\theta, \hat{\sigma}_\#) = \hat{T}_{f_1}^{(n)}(\theta, \sigma) + o_p(1)$ is asymptotically normal, with mean zero under $\mathbf{P}_{\theta, \sigma, 0; g_1}^{(n)}$, mean

$$\tau \frac{\mathcal{K}_{g_1}(f_1, g_1) - \mathcal{J}_{g_1}(f_1, g_1)(\kappa(f_1) + \kappa(g_1)) + \mathcal{I}_{g_1}(f_1, g_1)\kappa(f_1)\kappa(g_1)}{[\mathcal{K}_{g_1}(f_1) - 2\mathcal{J}_{g_1}(f_1)\kappa(f_1) + \mathcal{I}_{g_1}(f_1)\kappa^2(f_1)]^{1/2}} \tag{3.11}$$

under $\mathbf{P}_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ and variance one under both.

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

The tests based on $\hat{T}_{f_1}^{(n)}(\theta, \hat{\sigma}_\#)$ enjoy all the validity (under $\mathcal{H}_\theta^{(n)}$) and optimality (against $\bigcup_{\xi > 0} \bigcup_{\sigma \in \mathbb{R}_0^+} \{\mathbf{P}_{\theta, \sigma, \xi; f_1}^{(n)}\}$) properties one can expect. However, a closer look reveals that they are quite unsatisfactory on one count: under $g_1 \neq f_1$, their behavior strongly depends on the arbitrary choice of the concept of scale (here, the median of absolute deviations).

Consider, for example, the Gaussian version of (3.10), which takes the form

$$\hat{T}_{\phi_1}^{(n)}(\theta, \sigma) = \frac{1}{\sqrt{n\gamma^{(n)}(\phi_1)}} \sum_{i=1}^n (aZ_i^3(\theta, \sigma) - 3Z_i(\theta, \sigma)),$$

where

$$\gamma^{(n)}(\phi_1) = \gamma^{(n)}(\phi_1, \theta, \sigma) = a^2\sigma^{-6}m_6^{(n)}(\theta) - 6a\sigma^{-4}m_4^{(n)}(\theta) + 9\sigma^{-2}m_2^{(n)}(\theta).$$

The test based on $\hat{T}_{\phi_1}^{(n)}(\theta, \hat{\sigma})$ (here again, Slutsky’s lemma allows for not discretizing $\hat{\sigma}$) is a pseudo-Gaussian test, hence optimal under Gaussian assumptions; the asymptotic shift (3.11) is $\tau\sqrt{6/a}$ under $P_{\theta, \sigma, n^{-1/2}\tau; \phi_1}^{(n)}$, and

$$\tau[5a\mu_4(g_1) - (9 + 3a\kappa(g_1))\mu_2(g_1) + 3\kappa(g_1)][a^2\mu_6(g_1) - 6a\mu_4(g_1) + 9\mu_2(g_1)]^{-1/2},$$

where $\mu_k(g_1) := \int_{-\infty}^{\infty} z^k g_1(z) dz$, under $P_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$. This asymptotic shift strongly depends on a , hence on our (arbitrary) choice of a scale parameter. Setting to one the standard deviation instead of the median of absolute deviations would significantly modify the local behaviour of $\hat{T}_{f_1}^{(n)}(\theta, \hat{\sigma}_{\#})$ as soon as $g_1 \neq f_1$. This does not affect optimality properties (which hold under f_1), but is highly undesirable.

Now, that unpleasant feature of $\hat{T}_{\phi_1}^{(n)}$ is entirely due to the choice of $\kappa = \kappa(f_1)$ as a (non-random) centering in (3.10). That choice was entirely motivated by asymptotic orthogonality considerations under $P_{\theta, \sigma, 0; f_1}^{(n)}$, and does not affect the validity of the test. It follows that replacing $\kappa(f_1)$ with any data-dependent sequence $\kappa^{(n)}$ such that $\kappa^{(n)} - \kappa(f_1) = o_p(1)$ under $P_{\theta, \sigma, 0; f_1}^{(n)}$ asymptotically has no impact on $\hat{T}_{f_1}^{(n)}(\theta, \sigma)$ under $P_{\theta, \sigma, 0; f_1}^{(n)}$. Let us show that this sequence $\kappa^{(n)}$ can be chosen in order to cancel the unpleasant dependence of the test statistic on the definition of scale.

Provided that

$$f_1 \in \mathcal{F}_1^{\circ} := \{h_1 \in \mathcal{F}_1 : z \mapsto \phi_{h_1}(z) \text{ is differentiable, with derivative } \dot{\phi}_{h_1}\},$$

integration by parts yields

$$\mathcal{I}_{g_1}(f_1, g_1) = \int_{-\infty}^{\infty} \dot{\phi}_{f_1}(z)g_1(z) dz$$

and

$$\mathcal{J}_{g_1}(f_1, g_1) = 2 \int_{-\infty}^{\infty} z\phi_{f_1}(z)g_1(z) dz + \int_{-\infty}^{\infty} z^2\dot{\phi}_{f_1}(z)g_1(z) dz.$$

Therefore, $\mathcal{I}_{g_1}(f_1, g_1)$, $\mathcal{J}_{g_1}(f_1, g_1)$ and $\kappa_{g_1}(f_1, g_1) := \mathcal{J}_{g_1}(f_1, g_1)/\mathcal{I}_{g_1}(f_1, g_1)$ under $P_{\theta, \sigma, 0; g_1}^{(n)}$ are consistently estimated by

$$\mathcal{I}^{(n)\circ}(f_1) = \mathcal{I}^{(n)\circ}(f_1, \theta, \sigma) := \frac{1}{n} \sum_{i=1}^n \dot{\phi}_{f_1}(Z_i(\theta, \sigma)),$$

$$\mathcal{J}^{(n)\circ}(f_1) = \mathcal{J}^{(n)\circ}(f_1, \theta, \sigma) := \frac{2}{n} \sum_{i=1}^n Z_i(\theta, \sigma)\phi_{f_1}(Z_i(\theta, \sigma)) + \frac{1}{n} \sum_{i=1}^n Z_i^2(\theta, \sigma)\dot{\phi}_{f_1}(Z_i(\theta, \sigma))$$

and

$$\kappa^{(n)\circ}(f_1) = \kappa^{(n)\circ}(f_1, \theta, \sigma) := \mathcal{J}^{(n)\circ}(f_1)/\mathcal{I}^{(n)\circ}(f_1), \tag{3.12}$$

respectively.

Clearly, $\kappa^{(n)\circ}(f_1)$ satisfies the requirement that $\kappa^{(n)\circ}(f_1) - \kappa(f_1) = \text{op}(1)$ under $\mathbb{P}_{\theta,\sigma,0;f_1}^{(n)}$. In practice, however, $\kappa^{(n)\circ}(f_1, \theta, \sigma)$ cannot be computed from the observations, and $\kappa^{(n)\circ}(f_1, \theta, \hat{\sigma}_\#)$, where $Z_i(\theta, \hat{\sigma}_\#)$ has been substituted for $Z_i(\theta, \sigma)$, is to be used instead. As in the evaluation of $\gamma^{(n)}(f_1)$ above (see (3.5)), this substitution requires mild additional regularity conditions. LeCam's third lemma then applies exactly along the same lines, implying that

$$\kappa^{(n)\circ}(f_1, \theta, \hat{\sigma}_\#) - \kappa^{(n)\circ}(f_1, \theta, \sigma) = \text{op}(1)$$

under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$ as soon as the asymptotic covariances of $\mathcal{I}^{(n)\circ}(f_1)$ and $\mathcal{J}^{(n)\circ}(f_1)$ with $\Delta_{g_1;2}^{(n)}$ are finite. A simple computation (and the strong unimodality of f_1 and g_1) shows that a sufficient conditions for this is

$$g_1 \in \mathcal{F}_{f_1}^\circ := \left\{ h_1 \in \mathcal{F}_{f_1}^* : \int_{-\infty}^{\infty} z^3 \dot{\phi}_{f_1}(z) \phi_{h_1}(z) h_1(z) dz < \infty \right. \\ \left. \text{and } \int_{-\infty}^{\infty} z \dot{\phi}_{f_1}(z) \phi_{h_1}(z) h_1(z) dz < \infty \right\} \tag{3.13}$$

(no redundancy, since $\dot{\phi}_{f_1}$ is not necessarily monotone).

Emphasize the dependence of $\Delta_{f_1;3}^{(n)}(\kappa)$ on θ and σ by writing $\Delta_{f_1;3}^{(n)}(\kappa, \theta, \sigma)$: it follows from Lemma A.5 in the Appendix that, for $f_1 \in \mathcal{F}_1^\circ$ and $g_1 \in \mathcal{F}_{f_1}^\circ$, the difference between $\Delta_{f_1;3}^{(n)}(\kappa^{(n)\circ}(f_1, \theta, \hat{\sigma}_\#), \theta, \hat{\sigma}_\#)$ and $\Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \theta, \sigma)$ is $\text{op}(1)$ under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$. Letting (still for $f_1 \in \mathcal{F}_1^\circ$)

$$T_{f_1}^{(n)\circ}(\theta, \sigma) := \frac{1}{\sqrt{n\gamma^{(n)\circ}(f_1)}} \sum_{i=1}^n \phi_{f_1}(Z_i(\theta, \sigma))(Z_i^2(\theta, \sigma) - \kappa^{(n)\circ}(f_1)), \tag{3.14}$$

where

$$\gamma^{(n)\circ}(f_1) = \gamma^{(n)\circ}(f_1, \theta, \sigma) := \mathcal{K}^{(n)}(f_1) - 2\kappa^{(n)\circ}(f_1)\mathcal{J}^{(n)}(f_1) + (\kappa^{(n)\circ}(f_1))^2\mathcal{I}^{(n)}(f_1),$$

we thus have the following result.

Proposition 3.2. *Let $f_1 \in \mathcal{F}_1^\circ$ and $g_1 \in \mathcal{F}_{f_1}^\circ$. Then:*

- (i) $T_{f_1}^{(n)\circ}(\theta, \hat{\sigma}_\#) = T_{f_1}^{(n)\circ}(\theta, \sigma) + \text{op}(1)$ is asymptotically normal, with mean zero under $\mathbb{P}_{\theta,\sigma,0;g_1}^{(n)}$, mean

$$\tau \frac{\mathcal{K}_{g_1}(f_1, g_1) - \mathcal{J}_{g_1}(f_1, g_1)\kappa_{g_1}(f_1, g_1)}{[\mathcal{K}_{g_1}(f_1) - 2\mathcal{J}_{g_1}(f_1)\kappa_{g_1}(f_1, g_1) + \mathcal{I}_{g_1}(f_1)\kappa_{g_1}^2(f_1, g_1)]^{1/2}} \tag{3.15}$$

under $\mathbb{P}_{\theta,\sigma,n^{-1/2}\tau;g_1}^{(n)}$ and variance one under both.

- (ii) The sequence of tests rejecting the null hypothesis $\mathcal{H}_\theta^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{f_1}^\circ} \mathcal{H}_{\theta;g_1}^{(n)}$ of symmetry (with specified location θ , unspecified scale σ and unspecified standardized density $g_1 \in \mathcal{F}_{f_1}^\circ$) whenever $T_{f_1}^{(n)\circ}(\theta, \hat{\sigma}_\#)$ exceeds the $(1 - \alpha)$ standard normal quantile z_α is locally asymptotically most powerful at asymptotic level α for $\mathcal{H}_\theta^{(n)}$ against $\bigcup_{\xi > 0} \bigcup_{\sigma \in \mathbb{R}_0^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_1}^{(n)}\}$.

The advantage of the test statistic (3.14) compared to (3.10) is that, irrespective of the underlying density g_1 , its behavior does not depend on the definition of the scale parameter. The case of a Gaussian reference density ($f_1 = \phi_1$), however, is slightly different due to the particular form of the score function ϕ_{f_1} ; see Section 3.3.

3.2.2. Unspecified location

We now turn to the case under which both f_1 and the location θ are unspecified. Again, θ is to be replaced with some estimator, but additional care has to be taken about the asymptotic impact of this substitution. Still, from LeCam’s third lemma, it follows that the impact, under $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$, of an estimated θ on $\Delta_{f_1;3}^{(n)}(\kappa)$ can be obtained from the asymptotic behavior of

$$\begin{pmatrix} \Delta_{f_1;3}^{(n)}(\kappa) \\ \Delta_{g_1;1}^{(n)}(\theta, \sigma, 0) \end{pmatrix} = n^{-1/2} \sum_{i=1}^n \begin{pmatrix} \phi_{f_1}(Z_i)(Z_i^2 - \kappa) \\ \sigma^{-1} \phi_{g_1}(Z_i) \end{pmatrix},$$

which is asymptotically normal with asymptotic covariance matrix

$$\begin{pmatrix} \gamma_{g_1}^\kappa(f_1) & \delta_{g_1}^\kappa(f_1, g_1) \\ \delta_{g_1}^\kappa(f_1, g_1) & \sigma^{-2} \mathcal{I}(g_1) \end{pmatrix},$$

where $\delta_{g_1}^\kappa(f_1, g_1) := \sigma^{-1}(\mathcal{J}_{g_1}(f_1, g_1) - \kappa \mathcal{I}_{g_1}(f_1, g_1))$. Clearly, this covariance $\delta_{g_1}^\kappa(f_1, g_1)$ vanishes iff $\kappa = \kappa_{g_1}(f_1, g_1)$, which, for $g_1 = f_1$, coincides with $\kappa(f_1)$.

Assuming that an estimate $\kappa^{(n)}(f_1)$ such that $\kappa^{(n)}(f_1) - \kappa_{g_1}(f_1, g_1) = o_P(1)$ under $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$ exists, $\Delta_{f_1;3}^{(n)}(\kappa^{(n)}(f_1))$ is asymptotically equivalent to $\Delta_{f_1;3}^{(n)}(\kappa(f_1))$ under $\mathbb{P}_{\theta, \sigma, 0; f_1}^{(n)}$, and asymptotically uncorrelated with $\Delta_{g_1;1}^{(n)}(\theta, \sigma, 0)$ and $\Delta_{g_1;2}^{(n)}(\theta, \sigma, 0)$ (hence, asymptotically insensitive (in probability) to root- n perturbations of both θ and σ) under $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$. It follows from Section 3.2.1 that $\kappa^{(n)\circ}(f_1, \theta, \sigma)$ defined in (3.12) is such an estimator. The same reasoning as in Section 3.2.1 implies that this still holds when substituting, in $\Delta_{f_1;3}^{(n)}(\kappa)$, any estimators $\hat{\theta}_\#$ and $\hat{\sigma}_\#$ satisfying (C1) and (C2) for θ and σ . Finally, Lemma A.5 in the Appendix ensures that $\Delta_{f_1;3}^{(n)}(\kappa^{(n)\circ}(f_1, \hat{\theta}_\#, \hat{\sigma}_\#), \hat{\theta}_\#, \hat{\sigma}_\#)$ can be substituted for $\Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \theta, \sigma)$. We thus have shown the following result.

Proposition 3.3. *Let $f_1 \in \mathcal{F}_1^\circ$ and $g_1 \in \mathcal{F}_{f_1}^\circ$. Then:*

- (i) $T_{f_1}^{(n)\circ}(\hat{\theta}_\#, \hat{\sigma}_\#) = T_{f_1}^{(n)\circ}(\theta, \hat{\sigma}_\#) + o_P(1) = T_{f_1}^{(n)\circ}(\theta, \sigma) + o_P(1)$ is asymptotically normal, with mean zero under $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$, mean (3.15) under $\mathbb{P}_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ and variance one under both.

(ii) The sequence of tests rejecting the null hypothesis $\mathcal{H}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{f_1}^\circ} \bigcup_{\theta \in \mathbb{R}} \mathcal{H}_{\theta; g_1}^{(n)}$ of symmetry (with unspecified location θ , unspecified scale σ and unspecified standardized density g_1) whenever $T_{f_1}^{(n)\circ}(\hat{\theta}_\#, \hat{\sigma}_\#)$ exceeds the $(1 - \alpha)$ standard normal quantile z_α is locally asymptotically most powerful at asymptotic level α for $\mathcal{H}^{(n)}$ against $\bigcup_{\xi > 0} \bigcup_{\theta \in \mathbb{R}} \bigcup_{\sigma \in \mathbb{R}_0^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_1}^{(n)}\}$.

This test is based on the same test statistic $T_{f_1}^{(n)\circ}$ as the specified-location test of Proposition 3.2, except that the (here unspecified) location θ is replaced by an estimator $\hat{\theta}_\#$. The local powers of the two tests coincide: asymptotically, again, there is no loss of efficiency due to the non-specification of θ .

3.3. Pseudo-Gaussian tests

Particularizing the reference density f_1 as the standard normal one ϕ_1 in the tests of Sections 3.2.1 and 3.2.2 in principle yields *pseudo-Gaussian* tests, based on the test statistics $T_{\phi_1}^{(n)\circ}(\theta)$ or $T_{\phi_1}^{(n)\circ}(\hat{\theta})$. Due to the particular form of the Gaussian score function, however, the Gaussian statistic can be given a much simpler form. Indeed, $\mathcal{I}_{g_1}(\phi_1, g_1) = \mathcal{I}(\phi_1) = a$ does not depend on g_1 , and needs not be estimated, while $\mathcal{J}_{g_1}(\phi_1, g_1) = \mathcal{J}(\phi_1) = 3a\mu_2(g_1)$, so that $\kappa_{g_1}(f_1, g_1)$ is consistently estimated by $3m_2^{(n)}(\theta)/\sigma^2$. This, after elementary computation, yields the test statistic

$$T^{(n)\dagger}(\theta) := \frac{1}{\sqrt{n\gamma^{(n)\dagger}}} \sum_{i=1}^n (X_i - \theta)((X_i - \theta)^2 - 3m_2^{(n)}(\theta)), \tag{3.16}$$

where $\gamma^{(n)\dagger} := \gamma^{(n)\dagger}(\theta) := m_6^{(n)}(\theta) - 6m_2^{(n)}(\theta)m_4^{(n)}(\theta) + 9(m_2^{(n)}(\theta))^3$. For this test statistic $T^{(n)\dagger}(\theta)$, the asymptotic shift (3.15) under $\mathbb{P}_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ now takes the form

$$\tau[5\mu_4(g_1) - 9\mu_2^2(g_1)][\mu_6(g_1) - 6\mu_2(g_1)\mu_4(g_1) + 9\mu_2^3(g_1)]^{-1/2}.$$

This shift does not depend on a anymore, and still reduces to $\tau\sqrt{6/a}$ under $\mathbb{P}_{\theta, \sigma, n^{-1/2}\tau; \phi_1}^{(n)}$ (the same value as for $\hat{T}_{\phi_1}^{(n)}(\theta, \sigma)$, which confirms that optimality under Gaussian densities has been preserved); nor does it depend on the scale.

The tests based on the asymptotically standard normal null distribution of $T^{(n)\dagger}$ are optimal under Gaussian assumptions, but remain valid when those assumptions are violated. Again, a simple Slutsky argument allows for replacing θ (if unspecified) with any consistent estimator $\hat{\theta}$ without going through discretization; moreover, (3.16) does not depend on σ . The tests based on $T^{(n)\dagger}(\theta)$ and $T^{(n)\dagger}(\bar{X}^{(n)})$ both are closely related to the traditional test of symmetry based on $b_1^{(n)}$. More precisely, under any $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$, $g_1 \in (\mathcal{F}_{\phi_1}^\circ =) \mathcal{F}_{\phi_1}$ (note that the assumption $g_1 \in \mathcal{F}_{\phi_1}^\circ = \mathcal{F}_{\phi_1}$ implies that g_1 has finite moments of order six),

$$T^{(n)\dagger}(\theta) = T^{(n)\dagger}(\bar{X}^{(n)}) + o_p(1) = S_2^{(n)} + o_p(1), \tag{3.17}$$

where $S_2^{(n)}$ is the empirically standardized form (1.3) of $b_1^{(n)}$ (see (1.3)).

Summing up, we thus have the following result.

Proposition 3.4. *Let $g_1 \in \mathcal{F}_{\phi_1}$ and $\hat{\theta} = \theta + O_P(n^{-1/2})$; recall that $\mu_k(g_1) := \int_{-\infty}^{\infty} z^k g_1(z) dz$ stands for g_1 's moment of order k . Then:*

- (i) $T^{(n)\dagger}(\hat{\theta}) = T^{(n)\dagger}(\theta) + o_P(1)$ is asymptotically normal, with mean zero under $P_{\theta, \sigma, 0; g_1}^{(n)}$, mean

$$\tau[5\mu_4(g_1) - 9\mu_2^2(g_1)]/[\mu_6(g_1) - 6\mu_2(g_1)\mu_4(g_1) + 9\mu_2^3(g_1)]^{1/2}$$

under $P_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ and variance one under both.

- (ii) The sequence of tests rejecting the null hypothesis of symmetry (with specified location θ) $\mathcal{H}_{\theta}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{\phi_1}} \mathcal{H}_{\theta; g_1}^{(n)}$ whenever $T^{(n)\dagger}(\theta)$ exceeds the $(1 - \alpha)$ standard normal quantile z_{α} is locally asymptotically most powerful at asymptotic level α against $\bigcup_{\xi > 0} \bigcup_{\sigma \in \mathbb{R}_0^+} \{P_{\theta, \sigma, \xi; \phi_1}^{(n)}\}$.
- (iii) The sequence of tests rejecting the null hypothesis of symmetry (with unspecified location) $\mathcal{H}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{\phi_1}} \bigcup_{\theta \in \mathbb{R}} \mathcal{H}_{\theta; g_1}^{(n)}$ whenever $T^{(n)\dagger}(\hat{\theta})$ exceeds the $(1 - \alpha)$ standard normal quantile z_{α} is locally asymptotically most powerful at asymptotic level α against $\bigcup_{\xi > 0} \bigcup_{\theta \in \mathbb{R}} \bigcup_{\sigma \in \mathbb{R}_0^+} \{P_{\theta, \sigma, \xi; \phi_1}^{(n)}\}$.

For the sake of completeness, we also provide (with the same notation) the following result on the asymptotic behavior of the (suboptimal) test based on $m_3^{(n)}(\theta)$. Details are left to the reader.

Proposition 3.5. *Let $g_1 \in \mathcal{F}_{\phi_1}$. Then, $S_1^{(n)} := n^{1/2}m_3^{(n)}(\theta)/(m_6^{(n)}(\theta))^{1/2}$ is asymptotically normal, with mean zero under $P_{\theta, \sigma, 0; g_1}^{(n)}$, mean $\tau[5\mu_4(g_1) - 3\kappa(g_1)\mu_2(g_1)]/\mu_6^{1/2}(g_1)$ under $P_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ and variance one under both.*

Under Gaussian densities ($g_1 = \phi_1$), the asymptotic shifts of $T_{\phi_1}^{(n)\circ}(\theta)$ (Proposition 3.3(i)) and $S_1^{(n)}$ (Proposition 3.5) are $16\tau/\sqrt{6}$ and $16\tau/\sqrt{15}$, respectively; the asymptotic relative efficiency of $T^{(n)\dagger}(\theta)$ with respect to $S_1^{(n)}$ is thus as high as 2.5 in the vicinity of Gaussian densities. This, which is not a small difference, confirms the suboptimality of $m_3^{(n)}(\theta)$ -based tests.

3.4. Laplace tests

Replacing the Gaussian reference density ϕ_1 with the double-exponential one $f_{\mathcal{L}}$, we similarly obtain the *Laplace tests*. The assumption that $f_1 \in \mathcal{F}_1^{\circ}$ unfortunately rules out $f_{\mathcal{L}}$, since $\phi_{f_{\mathcal{L}}}(z) = \text{sign}(z)/d$ is not differentiable, so that the construction of $\kappa^{(n)\circ}(f_1)$ in (3.12) does not apply for $f_1 = f_{\mathcal{L}}$. Now, a direct construction is possible: $\mathcal{I}_{g_1}(f_{\mathcal{L}}, g_1)$ indeed reduces to $2g_1(0)/d$ – which is consistently estimated by $\mathcal{I}^{(n)\circ}(f_{\mathcal{L}}) := 2\hat{g}_1(0)/d$ (where \hat{g}_1 , e.g., is some kernel estimator of g_1). Similarly, $\mathcal{J}_{g_1}(f_{\mathcal{L}}, g_1)$ reduces to $(2/d) \int_{-\infty}^{\infty} |z|g_1(z) dz$ – which is consistently estimated by $\mathcal{J}^{(n)\circ}(f_{\mathcal{L}}) := (2/nd) \sum_{i=1}^n |Z_i(\theta, \hat{\sigma}_{\#})|$; the scaling constant d is easily computed, yielding

$d = 1/(\log 2) \approx 1.44$. Then,

$$\kappa^{(n)\circ}(f_{\mathcal{L}}) := \mathcal{J}^{(n)\circ}(f_{\mathcal{L}})/\mathcal{I}^{(n)\circ}(f_{\mathcal{L}}) = \frac{1}{n\hat{g}_1(0)} \sum_{i=1}^n |Z_i(\theta, \hat{\sigma}_{\#})|$$

is such that $\kappa^{(n)\circ}(f_{\mathcal{L}}) - \kappa(f_{\mathcal{L}}) = o_P(1)$ under $\mathbb{P}_{\theta, \sigma, 0; f_{\mathcal{L}}}^{(n)}$, as required.

The Laplace tests are based on $T_{\mathcal{L}}^{(n)\circ}(\theta)$ (specified θ) or $T_{\mathcal{L}}^{(n)\circ}(\hat{\theta})$ (unspecified θ), where

$$T_{\mathcal{L}}^{(n)\circ}(\theta) := \frac{1}{\sqrt{n\gamma^{(n)\circ}(f_{\mathcal{L}})}} \sum_{i=1}^n \text{sign}(Z_i(\theta, \hat{\sigma}_{\#}))((Z_i(\theta, \hat{\sigma}_{\#}))^2 - \kappa^{(n)\circ}(f_{\mathcal{L}})) \tag{3.18}$$

with

$$\gamma^{(n)\circ}(f_{\mathcal{L}}) = \frac{m_4^{(n)}}{\hat{\sigma}_{\#}^4} - 2 \frac{m_2^{(n)}}{n\hat{\sigma}_{\#}^2\hat{g}_1(0)} \sum_{i=1}^n |Z_i(\theta, \hat{\sigma}_{\#})| + \left(\frac{1}{n\hat{g}_1(0)} \sum_{i=1}^n |Z_i(\theta, \hat{\sigma}_{\#})| \right)^2.$$

These tests share with the Gaussian Fechner test (see [3]) the use of the score function $z \mapsto \text{sign}(z)z^2$. The orthogonalization, however, differs, since the Fechner and Edgeworth families the tests were built on are different. The following proposition summarizes their properties; details are left to the reader.

Proposition 3.6. *Let $g_1 \in \mathcal{F}_{f_{\mathcal{L}}}$, $\hat{\theta} = \theta + O_P(n^{-1/2})$ and denote by $\mu_{|k|}(g_1) := \int_{-\infty}^{\infty} |z|^k g_1(z) dz$ the absolute moment of order k of g_1 . Then,*

- (i) $T_{\mathcal{L}}^{(n)\circ}(\hat{\theta}) = T_{\mathcal{L}}^{(n)\circ}(\theta) + o_P(1)$ is asymptotically normal, with mean zero under $\mathbb{P}_{\theta, \sigma, 0; g_1}^{(n)}$, mean

$$\frac{\tau[4\mu_{|3|}(g_1) - 2\mu_{|1|}^2(g_1)/g_1(0)]}{[\mu_4(g_1) - 2\mu_2(g_1)\mu_{|1|}(g_1)/g_1(0) + \mu_{|1|}^2(g_1)/(g_1(0))^2]^{1/2}}$$

under $\mathbb{P}_{\theta, \sigma, n^{-1/2}\tau; g_1}^{(n)}$ and variance one under both.

- (ii) The sequence of tests rejecting the null hypothesis of symmetry (with specified location θ) $\mathcal{H}_{\theta}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{f_{\mathcal{L}}}} \mathcal{H}_{\theta; g_1}^{(n)}$ whenever $T_{\mathcal{L}}^{(n)\circ}(\theta)$ exceeds the $(1 - \alpha)$ standard normal quantile z_{α} is locally asymptotically most powerful at asymptotic level α against $\bigcup_{\xi > 0} \bigcup_{\sigma \in \mathbb{R}_0^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_{\mathcal{L}}}^{(n)}\}$.
- (iii) The sequence of tests rejecting the null hypothesis of symmetry (with unspecified location) $\mathcal{H}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_{f_{\mathcal{L}}}} \bigcup_{\theta \in \mathbb{R}} \mathcal{H}_{\theta; g_1}^{(n)}$ whenever $T_{\mathcal{L}}^{(n)\circ}(\hat{\theta})$ exceeds the $(1 - \alpha)$ standard normal quantile z_{α} is locally asymptotically most powerful at asymptotic level α against $\bigcup_{\xi > 0} \bigcup_{\theta \in \mathbb{R}} \bigcup_{\sigma \in \mathbb{R}_0^+} \{\mathbb{P}_{\theta, \sigma, \xi; f_{\mathcal{L}}}^{(n)}\}$.

Comparing the asymptotic shifts of the pseudo-Gaussian tests and the Laplace ones yields asymptotic relative efficiency values; the asymptotic efficiency of tests based on $T^{(n)\dagger}$ with respect

to those based on $T_{\mathcal{L}}^{(n)\circ}$ is 1.76 in the vicinity of Gaussian densities and 0.7 in the vicinity of double exponential ones. Finally, note that the empirical median $X_{1/2}^{(n)}$ here provides a much more sensible estimator of θ than the empirical mean $\bar{X}^{(n)}$; it has been used for $\hat{\theta}$ in the simulations of $T_{\mathcal{L}}^{(n)\circ}(\hat{\theta})$ in Section 4.

4. Finite sample performances

We performed a first simulation study on the basis of $N = 5000$ independent samples of size $n = 100$ from (2.2), with normal and double-exponential densities f_1 and skewness parameter values $\xi = 0.1$ and $\xi = 0.2$. Each of those samples was subjected, at asymptotic level $\alpha = 5\%$, to the classical specified location test of skewness based on $m_3^{(n)}(\theta)$ (i.e., on (1.1)), the (optimal) pseudo-Gaussian tests based on $b_1^{(n)}$ (i.e., on (1.3)) and the corresponding Laplace and logistic tests. For the sake of completeness, the two *triples tests* proposed by Randles et al. [13], which are based on the signs of $X_i + X_j - 2X_k$, $1 \leq i < j < k \leq n$, are also included in this simulation study. Those tests, which are location-invariant, do not follow from any argument of group invariance and are not distribution-free.

Rejection frequencies are reported in Table 1.

Note that all tests considered here, except for Randles', are extremely conservative, and in most cases hardly reach the nominal 5% rejection frequency under the null. Randles' tests, on the other hand, significantly over-reject, which does not facilitate comparisons. Despite that,

Table 1. Rejection frequencies (out of $N = 5000$ replications), under various symmetric and skewed normal and double-exponential distributions from the Edgeworth families (2.2), with $\xi = 0, 0.1, 0.2$, of the classical tests of skewness, based on $m_3^{(n)}(\theta)$ and $b_1^{(n)}$, the Gaussian, Laplace and logistic tests, and the *triples tests* $T_{R1}^{(n)}$ and $T_{R2}^{(n)}$ of [13]

Test	$\mathcal{SN}(\xi)$			$\mathcal{SL}(\xi)$		
	ξ			ξ		
	0	0.1	0.2	0	0.1	0.2
$m_3^{(n)}(\theta)$	0.0372	0.1136	0.0996	0.0306	0.6938	0.8722
$T^{(n)\dagger}(\theta)$	0.0434	0.7276	0.9958	0.0252	0.4596	0.6774
$b_1^{(n)}$	0.0416	0.6986	0.9746	0.0444	0.7458	0.8930
$T_{\mathcal{L}}^{(n)\circ}(\theta)$	0.0520	0.5424	0.9474	0.0406	0.9090	0.9998
$T_{\mathcal{L}}^{(n)\circ}(X_{1/2}^{(n)})$	0.0280	0.4440	0.8360	0.0284	0.8838	0.9960
$T_{\text{Log}}^{(n)\circ}(\theta)$	0.0492	0.7336	0.9954	0.0378	0.8516	0.9894
$T_{\text{Log}}^{(n)\circ}(\bar{X}^{(n)})$	0.0362	0.6626	0.9716	0.0384	0.8516	0.9880
$T_{R1}^{(n)}$	0.0518	0.6786	0.9606	0.0576	0.9276	0.9986
$T_{R2}^{(n)}$	0.0608	0.6992	0.9640	0.0650	0.9350	0.9988

the tests based on $b_1^{(n)}$, $T_{\mathcal{L}}^{(n)\circ}(X_{1/2}^{(n)})$ and $T_{\text{Log}}^{(n)\circ}(\bar{X}^{(n)})$ exhibit excellent performances, and largely outperform those based on $m_3^{(n)}(\theta)$ (although the latter requires θ to be known).

The Edgeworth families considered throughout this paper, however, served as a theoretical guideline in the construction of our Edgeworth testing procedures, and never were meant as an actual data generating process. One could argue that analyzing performances under alternatives of the Edgeworth type creates an unfair bias in favor of our methods. Therefore, we also generated $N = 5000$ independent samples of size $n = 100$ from the skew-normal $\mathcal{SN}(\lambda)$ and skew- t $St(\nu, \lambda)$ densities (with $\nu = 2$, $\nu = 4$ and $\nu = 8$ degrees of freedom) defined by Azzalini and Capitanio [1], for various values of their skewness coefficient λ ($\lambda = 0$ implying symmetry); since the sign of λ is not directly related to that of ξ , we only performed two-sided tests. That class of skewed densities was chosen in view of its increasing popularity among practitioners.

None of the tests considered in this simulation example are optimal in this Azzalini and Capitanio context. Inspection of Table 2 nevertheless reveals that the classical tests of skewness based on $m_3^{(n)}(\theta)$ and $b_1^{(n)}$ collapse under t_2 and t_4 , which have infinite sixth-order moments, and under the related $St(2, \lambda)$ and $St(4, \lambda)$ densities. The same tests fail to achieve the 5% nominal level under the Student distribution with 8 degrees of freedom (despite finite sixth-order moments) and show weak performance under the $St(8, \lambda)$ density. Remark that the suboptimality of the test based on $m_3^{(n)}(\theta)$, which, as a consequence of Proposition 3.1, may be considered as an artificial consequence of the choice of skewed families of the Edgeworth type, nevertheless also very neatly appears here. The triples tests behave uniformly well; note, however, their tendency to over-rejection, in particular under Student densities.

5. Conclusions and perspectives

We have derived the optimal tests for testing the hypothesis of symmetry within families of skewed densities mimicking, in the Gaussian case, the type of local asymmetry observed in a central limit behaviour. The resulting tests were obtained for specified and unspecified locations, under specified or unspecified densities (satisfying appropriate assumptions). Local powers and asymptotic relative efficiencies are computed and finite-sample performances are investigated in the context of classical skew-normal and skew- t distributions.

Establishing the optimality properties of the traditional test based on the Pearson–Fisher coefficient $b_1^{(n)}$ was one of the objectives of this work, and we show that this test is indeed optimal in the vicinity of Gaussian symmetry. Interestingly, its optimality holds for the specified-location as well as for the unspecified-location hypothesis of symmetry and is preserved if centering, in the computation of the test statistic, is based on robust root- n -consistent estimators of location rather than on the sample mean $\bar{X}^{(n)}$.

Table 3 provides a summary of the various tests described throughout Section 3, along with their validity and optimality properties.

These tests naturally extend into nonparametric rank-based ones. The hypothesis of symmetry indeed enjoys strong group invariance features. The null hypothesis $\mathcal{H}_\theta^{(n)}$ of symmetry with respect to θ is generated by the group $\mathcal{G}_\theta^{(n)}$, \circ of all transformations \mathcal{G}_h of \mathbb{R}^n such that

Table 2. Rejection frequencies (out of $N = 5000$ replications), under various symmetric and related skew-normal and skew- t distributions [1] $\mathcal{SN}(\lambda)$ and $\mathcal{St}(\nu, \lambda)$ ($\nu = 2, 4, 8$ and various λ) of the classical tests of skewness, based on $m_3^{(n)}(\theta)$ and $b_1^{(n)}$, the Gaussian, Laplace and logistic tests, and the *triples tests* $T_{R1}^{(n)}$ and $T_{R2}^{(n)}$ of [13]

Test	λ															
	$\mathcal{SN}(\lambda)$				$\mathcal{St}(2, \lambda)$				$\mathcal{St}(4, \lambda)$				$\mathcal{St}(8, \lambda)$			
	0	1	2	3	0	2	4	6	0	2	4	6	0	2	4	6
$m_3^{(n)}(\theta)$	0.0476	0.0482	0.0952	0.1936	0.0072	0.0106	0.0126	0.0144	0.0192	0.0190	0.0298	0.0436	0.0316	0.0608	0.1302	0.1754
$T^{(n)\dagger}(\theta)$	0.0374	0.0634	0.2988	0.5942	0.0046	0.0118	0.0182	0.0268	0.0144	0.0184	0.0586	0.1134	0.0260	0.1422	0.4078	0.5428
$b_1^{(n)}$	0.0418	0.0616	0.3066	0.6130	0.0172	0.0232	0.0308	0.0396	0.0252	0.0302	0.0754	0.1298	0.0322	0.1604	0.4186	0.5484
$T_{\mathcal{L}}^{(n)\circ}(\theta)$	0.0460	0.0690	0.2682	0.5406	0.0180	0.0414	0.0740	0.0988	0.0332	0.0406	0.1304	0.2514	0.0456	0.2054	0.5640	0.6906
$T_{\mathcal{L}}^{(n)\circ}(X_{1/2}^{(n)})$	0.0334	0.0472	0.2022	0.4736	0.0168	0.0288	0.0520	0.0678	0.0206	0.0302	0.0950	0.1798	0.0304	0.1518	0.4834	0.6260
$T_{\text{Log}}^{(n)\circ}(\theta)$	0.0468	0.0742	0.3542	0.7010	0.0154	0.0286	0.0496	0.0666	0.0236	0.0318	0.1082	0.2086	0.0342	0.2104	0.5846	0.7508
$T_{\text{Log}}^{(n)\circ}(\bar{X}^{(n)})$	0.0354	0.0568	0.2988	0.6426	0.0144	0.0256	0.0408	0.0492	0.0228	0.0276	0.0882	0.1694	0.0288	0.1712	0.5046	0.6696
$T_{R1}^{(n)}$	0.0540	0.0778	0.3602	0.7082	0.0618	0.1032	0.1798	0.2312	0.0508	0.0636	0.1842	0.3444	0.0530	0.2592	0.6766	0.8336
$T_{R2}^{(n)}$	0.0598	0.0886	0.3812	0.7258	0.0656	0.1098	0.1882	0.2424	0.0556	0.0688	0.1938	0.3582	0.0598	0.2740	0.6940	0.8422

Table 3. A summary of the various test statistics considered throughout the paper, with reference to their definitions, the testing problem (specified or unspecified location – scale throughout remains unspecified) addressed, the standardized densities g_1 for which they are valid, the densities at which they are optimal and the proposition (lemma) where their asymptotic properties are described

Test statistic	Reference	Location θ	Validity (comments)	Optimality (comments)	Asymptotic properties
$S_1^{(n)}(\theta)$ or $m_3^{(n)}(\theta)$ (classical)	(1.1)	Specified	Any g_1 with finite 6th-order moments	Suboptimal at the Gaussian	Proposition 3.5
$S_2^{(n)}$ or $b_1^{(n)}$ (classical)	(1.3)	Unspecified	Any g_1 with finite 6th-order moments	At the Gaussian, uniformly in θ, σ (= $T^{(n)\dagger}(\bar{X}^{(n)}) + o_p(1)$)	Proposition 3.4 (3.17)
$T_{f_1}^{(n)}(\theta, \hat{\sigma}_\#)$	(3.1)	Specified	Specified $f_1 \in \mathcal{F}_1$	At f_σ , uniformly in σ	Proposition 3.1
$T_{f_1}^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#)$	(3.1)	Unspecified	Specified $f_1 \in \mathcal{F}_1$	At f_σ , uniformly in θ, σ	Proposition 3.1
$\hat{T}_{f_1}^{(n)}(\theta, \hat{\sigma}_\#)$	(3.10)	Specified	Any $g_1 \in \mathcal{F}_{f_1}^*$ (3.9)	At f_σ , uniformly in σ (bad behavior under local alternatives and $g_1 \neq f_1$)	Lemma 3.1
$T_{f_1}^{(n)\circ}(\theta, \hat{\sigma}_\#)$	(3.14)	Specified	Any $g_1 \in \mathcal{F}_1^\circ$ (3.13)	At f_σ , uniformly in σ (same as $\hat{T}_{f_1}^{(n)}$, except for its bad behavior under local alternatives and $g_1 \neq f_1$)	Proposition 3.2
$T_{f_1}^{(n)\circ}(\hat{\theta}_\#, \hat{\sigma}_\#)$	(3.14)	Unspecified	Any $g_1 \in \mathcal{F}_1^\circ$ (3.13)	At f_σ , uniformly in θ, σ	Proposition 3.3
$T^{(n)\dagger}(\theta)$ (pseudo-Gaussian)	(3.16)	Specified	Any $g_1 \in \mathcal{F}_{\phi_1}$ (implies finite 6-order moments)	At the Gaussian, uniformly in σ	Proposition 3.4
$T^{(n)\dagger}(\hat{\theta})$ (pseudo-Gaussian)	(3.16)	Unspecified	Any $g_1 \in \mathcal{F}_{\phi_1}$ (implies finite 6-order moments)	At the Gaussian, uniformly in θ, σ	Proposition 3.5
$T_{\mathcal{L}}^{(n)\circ}(\theta)$ (Laplace)	(3.18)	Specified	Any $g_1 \in \mathcal{F}_{f_{\mathcal{L}}}$ (implies finite 4-order moments)	At the double exponential, uniformly in σ	Proposition 3.6
$T_{\mathcal{L}}^{(n)\circ}(\hat{\theta})$ (Laplace)	(3.18)	Unspecified	Any $g_1 \in \mathcal{F}_{f_{\mathcal{L}}}$ (implies finite 4-order moments)	At the double exponential, uniformly in θ, σ	Proposition 3.6

$g_h(x_1, \dots, x_n) := (h(x_1), \dots, h(x_n))$, where $\lim_{x \rightarrow \pm\infty} h(x) = \pm\infty$, and $x \mapsto h(x)$ is continuous, monotone increasing and skew-symmetric with respect to θ (that is, satisfies $h(\theta - z) - \theta = -(h(\theta + z) - \theta)$). A maximal invariant for that group is known to be the vector $(s_1(\theta), \dots, s_n(\theta))$, along with the vector $(R_{+,1}^{(n)}(\theta), \dots, R_{+,n}^{(n)}(\theta))$, where $s_i(\theta)$ is the sign of $X_i - \theta$ and $R_{+,i}^{(n)}(\theta)$ the rank of $|X_i - \theta|$ among $|X_1 - \theta|, \dots, |X_n - \theta|$. General results on semi-parametric efficiency [7] indicate that, in such a context, the expectation of the central sequence $\mathbf{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta})$ conditional on those *signed ranks* yields a version of the semi-parametrically efficient (at f_1 and $\boldsymbol{\vartheta}$) central sequence.

That rank-based approach is adopted in a companion paper [4]. For instance, the rank-based counterpart of the specified- θ test statistic of Proposition 3.6(ii) is the (strictly distribution-free, irrespective of any moment assumptions) van der Waerden test based on

$$T_{\text{vdW}}^{(n)}(\theta) := \frac{1}{\sqrt{n\gamma^{(n)}(\phi_1)}} \sum_{i=1}^n s_i(\theta) \Phi^{-1}\left(\frac{n+1+R_{+,i}^{(n)}(\theta)}{2(n+1)}\right) \left(\left(\Phi^{-1}\left(\frac{n+1+R_{+,i}^{(n)}(\theta)}{2(n+1)}\right) \right)^2 - 3 \right),$$

where $\gamma^{(n)}(\phi_1) := n^{-1} \sum_{r=1}^n \Phi^{-1}\left(\frac{n+1+r}{2(n+1)}\right) \left(\left(\Phi^{-1}\left(\frac{n+1+r}{2(n+1)}\right) \right)^2 - 3 \right)^2$ and Φ stands for the standard normal distribution function. The unspecified- θ case under such approach, however, is considerably more delicate.

Appendix

A.1. Proof of Proposition 2.1

The proof relies on [16], Lemma 1, which involves a set of six jointly sufficient conditions. Most of them readily follow from the form of local likelihoods and are left to the reader. The most delicate one is the quadratic mean differentiability of $(\theta, \sigma, \xi) \mapsto g_{\theta, \sigma, \xi; f_1}^{1/2}(x)$, which we establish in the following lemma, where $g_{\theta, \sigma, \xi; f_1}(x)$ is the density defined in (2.2).

Lemma A.1. *Let $f_1 \in \mathcal{F}_1$, $\theta \in \mathbb{R}$, $\sigma \in \mathbb{R}_0^+$ and $\xi \in \mathbb{R}$. Define*

$$\begin{aligned} g_{\theta, \sigma, \xi; f_1}(x) &:= \sigma^{-1} f_1\left(\frac{x-\theta}{\sigma}\right) - \frac{\xi}{\sigma} f_1\left(\frac{x-\theta}{\sigma}\right) \left(\left(\frac{x-\theta}{\sigma}\right)^2 - \kappa(f_1) \right) I[|x-\theta| \leq \sigma |z^*|] \\ &\quad - \text{sign}(\xi) \sigma^{-1} f_1\left(\frac{x-\theta}{\sigma}\right) \{ I[x-\theta > \text{sign}(-\xi)\sigma |z^*|] \\ &\quad \quad - I[x-\theta < \text{sign}(\xi)\sigma |z^*|] \}, \end{aligned}$$

$$D_\theta g_{\theta, \sigma, 0; f_1}^{1/2}(x) := \frac{1}{2} \sigma^{-3/2} f_1^{1/2}\left(\frac{x-\theta}{\sigma}\right) \phi_{f_1}\left(\frac{x-\theta}{\sigma}\right),$$

$$D_\sigma g_{\theta, \sigma, 0; f_1}^{1/2}(x) := \frac{1}{2} \sigma^{-3/2} f_1^{1/2}\left(\frac{x-\theta}{\sigma}\right) \left(\left(\frac{x-\theta}{\sigma}\right) \phi_{f_1}\left(\frac{x-\theta}{\sigma}\right) - 1 \right)$$

and

$$D_\xi g_{\theta,\sigma,\xi;f_1}^{1/2}(x)|_{\xi=0} := \frac{1}{2}\sigma^{-1/2}f_1^{1/2}\left(\frac{x-\theta}{\sigma}\right)\phi_{f_1}\left(\frac{x-\theta}{\sigma}\right)\left(\left(\frac{x-\theta}{\sigma}\right)^2 - \kappa(f_1)\right).$$

Then, as r, s and $t \rightarrow 0$,

- (i) $\int \{g_{\theta+t,\sigma+s,r;f_1}^{1/2}(x) - g_{\theta+t,\sigma+s,0;f_1}^{1/2}(x) - r D_\xi g_{\theta+t,\sigma+s,\xi;f_1}^{1/2}(x)|_{\xi=0}\}^2 dx = o(r^2).$
- (ii) $\int \left\{ g_{\theta+t,\sigma+s,0;f_1}^{1/2}(x) - g_{\theta,\sigma,0;f_1}^{1/2}(x) - \begin{pmatrix} t \\ s \end{pmatrix}' \begin{pmatrix} D_\theta g_{\theta,\sigma,0;f_1}^{1/2}(x) \\ D_\sigma g_{\theta,\sigma,0;f_1}^{1/2}(x) \end{pmatrix} \right\}^2 dx = o\left(\left\| \begin{pmatrix} t \\ s \end{pmatrix} \right\|^2\right).$
- (iii) $\int \{(D_\xi g_{\theta+t,\sigma+s,\xi;f_1}^{1/2}(x)|_{\xi=0} - D_\xi g_{\theta,\sigma,\xi;f_1}^{1/2}(x)|_{\xi=0})\}^2 dx = o(1).$
- (iv) $\int \left\{ g_{\theta+t,\sigma+s,r;f_1}^{1/2}(x) - g_{\theta,\sigma,0;f_1}^{1/2}(x) - \begin{pmatrix} t \\ s \\ r \end{pmatrix}' \begin{pmatrix} D_\theta g_{\theta,\sigma,0;f_1}^{1/2}(x) \\ D_\sigma g_{\theta,\sigma,0;f_1}^{1/2}(x) \\ D_\xi g_{\theta,\sigma,\xi;f_1}^{1/2}(x)|_{\xi=0} \end{pmatrix} \right\}^2 dx = o\left(\left\| \begin{pmatrix} t \\ s \\ r \end{pmatrix} \right\|^2\right).$

Proof. (i) Decompose $\int \{g_{\theta+t,\sigma+s,r;f_1}^{1/2}(x) - g_{\theta+t,\sigma+s,0;f_1}^{1/2}(x) - r D_\xi g_{\theta+t,\sigma+s,\xi;f_1}^{1/2}(x)|_{\xi=0}\}^2 dx$ into $a_1 + 2a_2$, where

$$a_1 = \int_{|u|<|z^*|} \left\{ \left(\frac{1}{\sigma+s}f_1(u)\right)^{1/2} [1+r\phi_{f_1}(u)(u^2-\kappa(f_1))]^{1/2} - \left(\frac{1}{\sigma+s}f_1(u)\right)^{1/2} - \frac{r}{2}(\sigma+s)^{-1/2} \frac{\dot{f}_1(u)}{f_1^{1/2}(u)}(u^2-\kappa(f_1)) \right\}^2 (\sigma+s) du$$

and

$$a_2 = \int_{u>|z^*|} \left\{ ((\sigma+s)^{-1}f_1(u))^{1/2} - \frac{r}{2}(\sigma+s)^{-1/2} \frac{\dot{f}_1(u)}{f_1^{1/2}(u)}(u^2-\kappa(f_1)) \right\}^2 (\sigma+s) du.$$

Since, for $|x| < 1$, $(1+x)^{1/2} = 1 + \frac{x}{2}(1+\lambda x)^{-1/2}$ for some $\lambda \in (0, 1)$, one easily obtains that

$$a_1 = \frac{r^2}{4} \int_{|u|<|z^*|} \left\{ (\sigma+s)^{-1/2} \frac{\dot{f}_1(u)}{f_1^{1/2}(u)}(u^2-\kappa(f_1)) \times ((1+\lambda r\phi_{f_1}(u)(u^2-\kappa(f_1)))^{-1/2} - 1) \right\}^2 (\sigma+s) du.$$

For $|u| < 1$, one has $(1 - (1 + \lambda u)^{-1/2})^2 \leq 2 \frac{2-\lambda}{1-\lambda}$, and the integrand is bounded by

$$2 \frac{2-\lambda}{1-\lambda} (u^2 - \kappa(f_1))^2 \left(\frac{\dot{f}_1^{1/2}(u)}{f_1^{1/2}(u)} \right)^2,$$

which is square-integrable; Lebesgue’s dominated convergence theorem thus implies that a_1 is $o(r^2)$. Turning to a_2 , we have that $a_2 \leq C((\sigma + s)^{-1} a_{21} + a_{22})$, where

$$a_{21} := \int_{u>|z^*|} f_1(u) \, du \quad \text{and} \quad a_{22} := \frac{r^2}{4} \int_{u>|z^*|} \left(\frac{\dot{f}_1(u)}{f_1^{1/2}(u)} \right)^2 (u^2 - \kappa(f_1))^2 \, du.$$

The definition of \mathcal{F}_1 implies that $a_{21} = O((z^*)^{-\beta})$, hence that $a_{21} = o(r^2)$ if $r(z^*)^{\beta/2} \rightarrow \infty$ as $r \rightarrow 0$. This latter condition holds, since $\phi_{f_1}(z) = o(z^{\beta/2-2})$ and since the definition of z^* entails that

$$-1 = r(z^*)^{\beta/2} \frac{\phi_{f_1}(z^*)}{z^{*(\beta/2-2)}} \frac{(z^{*2} - \kappa(f_1))}{z^{*2}}.$$

An application of Lebesgue’s dominated convergence theorem again yields $a_{22} = o(r^2)$.

- (ii) This is a particular case of Lemma A.1 in [6] (here in a simpler univariate context).
- (iii) The fact that $D_\xi g_{\theta, \sigma, \xi; f_1}^{1/2}(x)|_{\xi=0}$ is square-integrable implies that

$$\|D_\xi g_{\theta+t, \sigma, \xi; f_1}^{1/2}(x)|_{\xi=0} - D_\xi g_{\theta, \sigma, \xi; f_1}^{1/2}(x)|_{\xi=0}\|_{L^2} = o(1)$$

as t tends to zero. Define $f_{1;\text{exp}}(x) := f_1(e^x)$ and $(f_{1;\text{exp}}^{1/2}(x))' := \frac{1}{2} f_1^{-1/2}(e^x) \dot{f}_1(e^x) e^x$. For the perturbation of σ , we have

$$\begin{aligned} & \int \left\{ (D_\xi g_{\theta, \sigma+s, \xi; f_1}^{1/2}(x)|_{\xi=0} - D_\xi g_{\theta, \sigma, \xi; f_1}^{1/2}(x)|_{\xi=0}) \right\}^2 \, dx \\ &= 2\sigma \int_0^\infty \left| \sigma^{-1/2} z \left(\left(1 + \frac{s}{\sigma}\right)^{-3/2} (f_{1;\text{exp}}^{1/2})' \left(\ln(z) - \ln\left(1 + \frac{s}{\sigma}\right) \right) - (f_{1;\text{exp}}^{1/2})'(\ln(z)) \right) \right. \\ & \quad \left. - \sigma^{-1/2} z^{-1} \kappa(f_1) \right. \\ & \quad \left. \times \left(\left(1 + \frac{s}{\sigma}\right)^{1/2} (f_{1;\text{exp}}^{1/2})' \left(\ln(z) - \ln\left(1 + \frac{s}{\sigma}\right) \right) + (f_{1;\text{exp}}^{1/2})'(\ln(z)) \right) \right|^2 \, dz \\ & \leq C(c_1 + c_2), \end{aligned}$$

where

$$c_1 = \int_{-\infty}^\infty \left(e^{(3/2)(u-\ln(1+s/\sigma))} (f_{1;\text{exp}}^{1/2})' \left(u - \ln\left(1 + \frac{s}{\sigma}\right) \right) - e^{(3/2)u} (f_{1;\text{exp}}^{1/2})'(u) \right)^2 \, du$$

and

$$c_2 = \int_{-\infty}^{\infty} \left(e^{-(1/2)u - \ln(1+s/\sigma)} (f_{1;\text{exp}}^{1/2})' \left(u - \ln \left(1 + \frac{s}{\sigma} \right) \right) - e^{-(1/2)u} (f_{1;\text{exp}}^{1/2})'(u) \right)^2 du.$$

Now, both $e^{-(1/2)u} (f_{1;\text{exp}}^{1/2})'(u)$ and $e^{(3/2)u} (f_{1;\text{exp}}^{1/2})'(u)$ are square-integrable since $f_1 \in \mathcal{F}_1$. Therefore, quadratic mean continuity implies that c_1 and c_2 are $o(1)$ as $s \rightarrow 0$.

(iv) The left-hand side in (iv) is bounded by $C(b_1 + b_2 + b_3)$, where

$$b_1 = \int \{g_{\theta+t, \sigma+s, r; f_1}^{1/2}(x) - g_{\theta+t, \sigma+s, 0; f_1}^{1/2}(x) - r D_{\xi} g_{\theta+t, \sigma+s, \xi; f_1}^{1/2}(x)\}^2 dx,$$

$$b_2 = \int \left\{ g_{\theta+t, \sigma+s, 0; f_1}^{1/2}(x) - g_{\theta, \sigma, 0; f_1}^{1/2}(x) - (t, s) \begin{pmatrix} D_{\theta} g_{\theta, \sigma, 0; f_1}^{1/2}(x) \\ D_{\sigma} g_{\theta, \sigma, 0; f_1}^{1/2}(x) \end{pmatrix} \right\}^2 dx$$

and

$$b_3 = \int \{r(D_{\xi} g_{\theta+t, \sigma+s, \xi; f_1}^{1/2}(x) - D_{\xi} g_{\theta, \sigma, \xi; f_1}^{1/2}(x))\}^2 dx.$$

The result then follows from (i)–(iii). □

A.2. Asymptotic linearity

A.2.1. Asymptotic linearity of $\Delta_{f_1;3}^{(n)}$

The asymptotic linearity of $\Delta_{f_1;3}^{(n)}$ is required in the construction of the optimal parametric test of Section 3.2.2. Note that the proof below needs uniform local asymptotic normality in θ and σ only.

Proposition A.1. *Let $f_1 \in \mathcal{F}_1$ and $g_1 \in \mathcal{F}_{f_1}$. Then, under $P_{\theta, \sigma, 0; g_1}^{(n)}$, as $n \rightarrow \infty$,*

- (i) $\Delta_{f_1;3}^{(n)}(\theta + n^{-1/2}t, \sigma, 0) = \mathbf{\Delta}_{f_1;3}^{(n)}(\theta, \sigma, 0) - t\sigma^{-1}(\mathcal{J}_{g_1}(f_1, g_1) - \kappa(f_1)\mathcal{I}_{g_1}(f_1, g_1)) + o_P(1)$ for all $t \in \mathbb{R}$.
- (ii) $\mathbf{\Delta}_{f_1;3}^{(n)}(\theta, \sigma + n^{-1/2}s, 0) = \mathbf{\Delta}_{f_1;3}^{(n)}(\theta, \sigma, 0) + o_P(1)$ for all $s \in \mathbb{R}$.

Proof. Define

$$D_{f_1;3}^{(n)}(\theta, \sigma) := n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_i^{(n)}(\theta, \sigma))(Z_i^{(n)}(\theta, \sigma))^2.$$

Letting $K_{f_1}(u) := \phi_{f_1}(G_{1+}^{-1}(u)) (G_{1+}^{-1}(u))^2$ with $G_{1+}(z) := (2G_1(z) - 1)I[z \geq 0]$, note that

$$D_{f_1;3}^{(n)}(\theta, \sigma, 0) = n^{-1/2} \sum_{i=1}^n s_i(\theta) K_{f_1}(G_{1+}(|Z_i^{(n)}(\theta, \sigma)|)).$$

Writing $\theta^{(n)}$ for $\theta + n^{-1/2}t$, $Z_i^{(n)}$ for $Z_i^{(n)}(\theta, \sigma)$, $s_i^{(n)}$ for $\text{sign}(Z_i^{(n)})$, $Z_{i;n}^{(n)}$ for $Z_i^{(n)}(\theta^{(n)}, \sigma)$ and $s_{i;n}^{(n)}$ for $\text{sign}(Z_{i;n}^{(n)})$, let us show that

$$n^{-1/2} \sum_{i=1}^n s_{i;n}^{(n)} K_{f_1}(G_{1+}(|Z_{i;n}^{(n)}|)) - n^{-1/2} \sum_{i=1}^n s_i^{(n)} K_{f_1}(G_{1+}(|Z_i^{(n)}|)) + t\sigma^{-1} \mathcal{J}_{g_1}(f_1, g_1)$$

is $o_p(1)$; the proofs of (ii) and that of

$$n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_{i;n}^{(n)}) = n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_i^{(n)}) - t\sigma^{-1} \mathcal{I}_{g_1}(f_1, g_1) + o_p(1)$$

follow along the same lines and are therefore left to the reader. Let

$$K_{f_1}^{(l)}(u) := K_{f_1}\left(\frac{2}{l}\right) l \left(u - \frac{1}{l}\right) I\left[\frac{1}{l} < u \leq \frac{2}{l}\right] + K_{f_1}(u) I\left[\frac{2}{l} < u \leq 1 - \frac{2}{l}\right] \\ + K_{f_1}\left(1 - \frac{2}{l}\right) l \left(\left(1 - \frac{1}{l}\right) - u\right) I\left[1 - \frac{2}{l} < u \leq 1 - \frac{1}{l}\right].$$

Continuity of $u \mapsto K_{f_1}(u)$ implies continuity of $u \mapsto K_{f_1}^{(l)}(u)$ on the interval $]0, 1[$. Moreover, since this function is compactly supported, it is bounded, for any (sufficiently large) $l \in \mathbb{N}_0$, by the monotone increasing function $u \mapsto K_{f_1}(u)$. Let E_0 denote expectation under $P_{\theta, \sigma, 0; g_1}^{(n)}$. One easily shows that $D_{f_1;3}^{(n)}(\theta, \sigma)$ decomposes into $D_1^{(n,m)} + D_2^{(n,m)} - R_1^{(n,m)} + R_2^{(n,m)} + R_3^{(m)}$ where, defining $\mathcal{J}_{g_1}^{(l)}(f_1, g_1) := \int_0^1 K_{f_1}^{(l)}(u) \phi_{g_1}(G_1^{-1}(u)) du$,

$$D_1^{(n,l)} = n^{-1/2} \sum_{i=1}^n s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - n^{-1/2} \sum_{i=1}^n s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)) \\ - n^{1/2} E_0[s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|))], \\ D_2^{(n,l)} = n^{1/2} E_0[s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|))] + t\sigma^{-1} \mathcal{J}_{g_1}^{(l)}(f_1, g_1), \\ R_1^{(n,l)} = n^{-1/2} \sum_{i=1}^n s_i (K_{f_1}(G_{1+}(|Z_i^{(n)}|)) - K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))), \\ R_2^{(n,l)} = n^{-1/2} \sum_{i=1}^n s_{i;n} (K_{f_1}(G_{1+}(|Z_{i;n}^{(n)}|)) - K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)))$$

and

$$R_3^{(l)} = t\sigma^{-1} (\mathcal{J}_{g_1}(f_1, g_1) - \mathcal{J}_{g_1}^{(l)}(f_1, g_1)).$$

In order to conclude, we prove that $D_1^{(n,l)}$ and $D_2^{(n,l)}$ are $o_p(1)$ under $P_{\theta, \sigma, 0; g_1}^{(n)}$, as $n \rightarrow \infty$, for fixed l , and that $R_1^{(n,l)}$, $R_2^{(n,l)}$ and $R_3^{(l)}$ are $o_p(1)$ under the same sequence of hypotheses, as

$l \rightarrow \infty$, uniformly in n . For the sake of convenience, those three results are treated separately (Lemmas A.2–A.4).

Lemma A.2. For any fixed l , $D_1^{(n,l)} = o_P(1)$ as $n \rightarrow \infty$, under $P_{\theta,\sigma,0;g_1}^{(n)}$.

Lemma A.3. For any fixed l , $D_2^{(n,l)} = o_P(1)$ as $n \rightarrow \infty$, under $P_{\theta,\sigma,0;g_1}^{(n)}$.

Lemma A.4. (i) Under $P_{\theta,\sigma,0;g_1}^{(n)}$, $R_1^{(n,l)} = o_P(1)$ as $l \rightarrow \infty$, uniformly in n .

(ii) $R_2^{(n,l)} = o_P(1)$ as $l \rightarrow \infty$, under $P_{\theta,\sigma,0;g_1}^{(n)}$ (for n sufficiently large), uniformly in n .

(iii) $R_3^{(l)}$ is $o(1)$ as $l \rightarrow \infty$.

Proof of Lemma A.2. Consider the i.i.d. variables

$$T_i^{(n,l)} := s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)).$$

One easily verifies that $D_1^{(n,l)} = n^{-1/2} \sum_{i=1}^n (T_i^{(n,l)} - E_0[T_i^{(n,l)}])$. Writing Var_0 for variances under $P_{\theta,\sigma,0;g_1}^{(n)}$, we have that

$$\begin{aligned} E_0(D_1^{(n,l)}) &\leq n^{-1} E_0 \left[\left(\sum_{i=1}^n (T_i^{(n,l)} - E_0[T_i^{(n,l)}]) \right)^2 \right] \\ &\leq n^{-1} \text{Var}_0 \left[\sum_{i=1}^n (T_i^{(n,l)} - E_0[T_i^{(n,l)}]) \right] = \text{Var}_0[T_i^{(n,l)}] \leq E_0[(T_i^{(n,l)})^2], \end{aligned}$$

and it only remains to show that

$$E_0[(T_i^{(n,l)})^2] = E_0[(s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)))^2] = o(1)$$

as $n \rightarrow \infty$. Now,

$$\begin{aligned} &(s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)))^2 \\ &= (s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)) + s_{i;n} K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)) \\ &\quad - s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)))^2 \\ &\leq 2(K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)))^2 + 2(K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|)))^2 (s_{i;n} - s_i)^2. \end{aligned}$$

Because $u \mapsto K_{f_1}^{(l)}(u)$ is continuous and $|Z_{i;n}^{(n)} - Z_i^{(n)}|$ is $o_P(1)$, $K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|)) - K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))$ also is $o_P(1)$. Moreover, since $K_{f_1}^{(l)}$ is bounded, this convergence to zero also holds in quadratic mean. Similarly, $K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))(s_{i;n} - s_i) = o_P(1)$ since $K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))$

is bounded and $|s_{i;n} - s_i|$ is $o_p(1)$. Finally, both $s_{i;n}$ and s_i are bounded, implying that this convergence to zero also holds in quadratic mean. \square

Proof of Lemma A.3. Let $B_1^{(n,l)} := n^{-1/2} \sum_{i=1}^n s_i K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))$. As $n \rightarrow \infty$, under $P_{\theta, \sigma, 0; g_1}^{(n)}$, $B_1^{(n,l)}$ is asymptotically $\mathcal{N}(0, E[(K_{f_1}^{(l)}(U))^2])$, where U stands for a random variable uniformly distributed over the unit interval. Also, letting $B_2^{(n,l)} := n^{-1/2} \sum_{i=1}^n s_{i;n} \times K_{f_1}^{(l)}(G_{1+}(|Z_{i;n}^{(n)}|))$, it follows from ULAN that $B_2^{(n,l)} - t\sigma^{-1} \mathcal{J}_{g_1}^{(l)}(f_1, g_1)$ is asymptotically $\mathcal{N}(0, E[(K_{f_1}^{(l)}(U))^2])$ as $n \rightarrow \infty$, under $P_{\theta, \sigma, 0; g_1}^{(n)}$. Since $D_1^{(n,l)} = B_2^{(n,l)} - B_1^{(n,l)} - E_0[B_2^{(n,l)}] = o_p(1)$, we have that $B_2^{(n,l)} - E_0[B_2^{(n,l)}]$ is asymptotically $\mathcal{N}(0, E[(K_{f_1}^{(l)}(U))^2])$ as $n \rightarrow \infty$, under $P_{\theta, \sigma, 0; g_1}^{(n)}$. Therefore, still as $n \rightarrow \infty$, $D_2^{(n,l)} = E_0[B_2^{(n,l)}] - t\sigma^{-1} \mathcal{J}_{g_1}^{(l)}(f_1, g_1) = o(1)$. \square

Proof of Lemma A.4. (i) We have that

$$\begin{aligned} E_0[(R_1^{(n,l)})^2] &\leq CE_0[(K_{f_1}(G_{1+}(|Z_i^{(n)}|)) - K_{f_1}^{(l)}(G_{1+}(|Z_i^{(n)}|))]^2] \\ &= C \int_{-\infty}^{\infty} (K_{f_1}(u) - K_{f_1}^{(l)}(u))^2 du; \end{aligned}$$

for any $u \in]0, 1[$, $K_{f_1}^{(l)}(u)$ converges to $K_{f_1}(u)$ and the integrand is bounded (uniformly in l) by $4(K_{f_1}(u))^2$, which is integrable on $]0, 1[$. Lebesgue’s dominated convergence theorem thus implies that $E_0[(R_1^{(n,l)})^2] = o(1)$ as $l \rightarrow \infty$, uniformly in n .

(ii) The claim here is the same as in (i), with $Z_{i;n}^{(n)}$ replacing $Z_i^{(n)}$. Accordingly, (ii) holds under $P_{\theta^{(n)}, \sigma, 0; g_1}^{(n)}$. That it also holds under $P_{\theta, \sigma, 0; g_1}^{(n)}$ follows from Lemma 3.5 in [9].

(iii) Note that

$$\begin{aligned} |\mathcal{J}_{g_1}(f_1, g_1) - \mathcal{J}_{g_1}^{(l)}(f_1, g_1)|^2 &= \left| \int_0^1 \phi_{g_1}(G_1^{-1}(u))(K_{f_1}^{(l)}(u) - K_{f_1}(u)) du \right|^2 \\ &\leq \mathcal{I}(g_1) \int_0^1 ((K_{f_1}^{(l)}(u) - K_{f_1}(u))^2) du, \end{aligned}$$

where the integrand is bounded by $4(K_{f_1}(u))^2$, which is square-integrable. Pointwise convergence of $K_{f_1}^{(l)}(u)$ to $K_{f_1}(u)$ implies that $\mathcal{J}_{g_1}(f_1, g_1) - \mathcal{J}_{g_1}^{(l)}(f_1, g_1) = o(1)$ as $l \rightarrow \infty$. The result follows. \square

A.2.2. Substitution of $\Delta_{f_1;3}^{(n)}(\kappa^{(n)\circ}(f_1, \hat{\theta}_{\#}, \hat{\sigma}_{\#}), \hat{\theta}_{\#}, \hat{\sigma}_{\#})$ for $\Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \theta, \sigma)$.

Lemma A.5. Let $f_1 \in \mathcal{F}_1^{\circ}$ and $g_1 \in \mathcal{F}_{f_1}^{\circ}$. Then, under $P_{\theta, \sigma, 0; g_1}^{(n)}$:

- (i) $\Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \hat{\theta}_{\#}, \hat{\sigma}_{\#}) - \Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \theta, \sigma) = o_p(1)$.
- (ii) $\Delta_{f_1;3}^{(n)}(\kappa^{(n)\circ}(f_1, \hat{\theta}_{\#}, \hat{\sigma}_{\#}), \hat{\theta}_{\#}, \hat{\sigma}_{\#}) - \Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \hat{\theta}_{\#}, \hat{\sigma}_{\#}) = o_p(1)$.

$$(iii) \Delta_{f_1;3}^{(n)}(\kappa^{(n)\circ}(f_1, \hat{\theta}_\#, \hat{\sigma}_\#), \hat{\theta}_\#, \hat{\sigma}_\#) - \Delta_{f_1;3}^{(n)}(\kappa_{g_1}(f_1, g_1), \theta, \sigma) = o_P(1).$$

Proof. Part (i) is a direct consequence of Proposition A.1. The left-hand side in (ii) can be written as

$$T_1^{(n)} \times T_2^{(n)} := (\kappa^{(n)\circ}(f_1, \hat{\theta}_\#, \hat{\sigma}_\#) - \kappa_{g_1}(f_1, g_1)) \times n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_i^{(n)}(\hat{\theta}_\#, \hat{\sigma}_\#)) \quad (A.1)$$

where $T_1^{(n)}$ is $o_P(1)$. Now, ULAN implies that

$$T_2^{(n)} = n^{-1/2} \sum_{i=1}^n \phi_{f_1}(Z_i^{(n)}(\theta, \sigma)) + (\sigma^{-1} \mathcal{I}_{g_1}(f_1, g_1), 0) n^{1/2} \left(\begin{pmatrix} \hat{\theta}_\# \\ \hat{\sigma}_\# \end{pmatrix} - \begin{pmatrix} \theta \\ \sigma \end{pmatrix} \right) + o_P(1), \quad (A.2)$$

as $n \rightarrow \infty$ under $P_{\theta, \sigma, 0; g_1}^{(n)}$. Hence, the central limit theorem and the root- n -consistency of $\hat{\theta}_\#$ and $\hat{\sigma}_\#$ entail that (A.2) is $O_P(1)$; the result follows. As for (iii), it is a direct consequence of (i) and (ii). \square

Acknowledgements

Marc Hallin's research supported by the Sonderforschungsbereich "Statistical modelling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (Deutsche Forschungsgemeinschaft) and a Discovery Grant of the Australian Research Council. This work was partly completed while visiting ORFE and the Bendheim Center at Princeton University. Davy Paindaveine's research supported by a Mandat d'Impulsion Scientifique of the Fonds National de la Recherche Scientifique, Communauté française de Belgique.

References

- [1] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *J. Roy. Statist. Soc. Ser. B* **65** 367–389. [MR1983753](#)
- [2] Beran, R. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* **2** 63–74. [MR0345295](#)
- [3] Cassart, D., Hallin, M. and Paindaveine, D. (2008). Optimal detection of Fechner asymmetry. *J. Statist. Plann. Inference* **138** 2499–2525. [MR2414260](#)
- [4] Cassart, D., Hallin, M. and Paindaveine, D. (2010). A class of optimal rank-based tests of symmetry. Submitted.
- [5] Choi, S., Hall, W.J. and Schick, A. (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.* **24** 841–861. [MR1394992](#)
- [6] Hallin, M. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756. [MR2329465](#)
- [7] Hallin, M. and Werker, B.J.M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9** 137–165. [MR1963675](#)

- [8] Hollander, M. (1988). Testing for symmetry. In *Encyclopedia of Statistical Sciences* (N.L. Johnson and S. Kotz, eds.) **9** 211–216. New York: Wiley.
- [9] Jurečková, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.
- [10] Kreiss, J.-P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133. [MR0885727](#)
- [11] Le Cam, L.M. (1986). *Asymptotic Methods in Statistical Decision Theory*. New York: Springer. [MR0856411](#)
- [12] Le Cam, L.M. and Yang, G.L. (2000). *Asymptotics in Statistics*, 2nd ed. New York: Springer. [MR1784901](#)
- [13] Randles, R.H., Fligner, M.A., Policello, G.E. and Wolfe, D.A. (1980). An asymptotically distribution-free test for symmetry versus asymmetry. *J. Amer. Statist. Assoc.* **75** 168–172. [MR0568588](#)
- [14] Stein, C. (1956). Efficient nonparametric testing and estimation. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* **1** 187–196. Berkeley, CA: Univ. California Press. [MR0084921](#)
- [15] Stone, C. (1975). Adaptive maximum likelihood estimation of a location parameter. *Ann. Statist.* **3** 267–284. [MR0362669](#)
- [16] Swensen, A.R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *J. Multivariate Anal.* **16** 54–70. [MR0778489](#)
- [17] van der Vaart, A.W. (1988). *Asymptotic Statistics*. Cambridge, UK: Cambridge Univ. Press.

Received March 2007 and revised February 2010