Journal of Multivariate Analysis 193 (2023) 105118

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

A new distance based measure of asymmetry

Jonas Baillien [a], Irène Gijbels [a,*], Anneleen Verhasselt [b]

[a] KU Leuven, Department of Mathematics, Celestijnenlaan 200B, Box 2400, B-3001 Leuven (Heverlee), Belgium
[b] Center for Statistics, Data Science Institute, Hasselt University, Agoralaan-building D, B-3590 Diepenbeek, Belgium

A R T I C L E I N F O

Article history:
Received 29 April 2022
Received in revised form 6 October 2022
Accepted 6 October 2022
Available online 15 October 2022

AMS 2010 subject classifications:
primary 62H05
secondary 62Fxx
62GXX

Keywords:
Asymmetry measure
Mode
Nonparametric estimation
Parametric estimation
Reflective symmetry
Uniform consistency

A B S T R A C T

In a univariate setting there is a near unanimous agreement on the notion of skewness. Nevertheless, many more skewness measures, or also called asymmetry measures (or indices) exist, each with their benefits. Extending the concept of skewness or asymmetry to a multivariate setting is a much harder problem. Attempts have been made, but the unanimity of the univariate setting is no longer present. Most asymmetry indices are scalar or vector based measures, but this can lead to a loss of information concerning asymmetry. To this end, we propose a novel functional asymmetry index which is based on the natural idea of reflective symmetry around the mode. The proposed index is also extended to the multivariate setting and a summarizing scalar (or vector based index in multivariate setting) is derived from it.

1. Introduction

Quantifying asymmetry (or skewness) is a well known problem. As early as the late 19th century, great mathematicians as Pearson came up with nowadays widely known measures to capture asymmetry. Multivariate extensions came later with Mardia's skewness [34] and subsequently many more measures have been proposed. The need for asymmetry measures has grown with the rise of econometrics for risk calculation, but it also finds its use in normality testing and all fields which rely on this. Although the wording asymmetry 'measure' is of standard use for a quantity that indicates the level of asymmetry of a distribution, we also utilize the word 'index' as the former might misleadingly be linked to a probabilistic measure.

Denote with X a univariate random variable taking values in $B \subseteq \mathbb{R}$, with density function $f_X(x)$, cumulative distribution function $F_X(x)$ and quantile function $Q_X(p) = \inf_{x \in B}\{F_X(x) \geq p\}$. The earliest measures of skewness for a random variable X were based on measures of location for X. These mainly consisted in comparing the mean ($\mu_X = E[X]$), median ($M_X = Q_X(0.5)$) and mode ($\mathcal{M}_X = \arg\max_{x \in B} f_X(x)$). It has been agreed upon by several authors (see e.g., [4]) that a measure of asymmetry $\gamma(X)$ should posses certain desirable properties. These are

(P1) $\gamma(X)$ is location and scale (for positive scaling factors) invariant, i.e., $\forall a > 0$ and $b \in \mathbb{R}$: $\gamma(aX + b) = \gamma(X)$;
(P2) $\gamma(X) = 0$ for symmetric X;
(P3) $\gamma(-X) = -\gamma(X)$;

* Corresponding author.
E-mail address: irene.gijbels@kuleuven.be (I. Gijbels).

https://doi.org/10.1016/j.jmva.2022.105118

(P4) Let Y have cumulative distribution function $G(y)$, if $G^{-1}(F_X(x))$ is a convex function, then $\gamma(Y) \geq \gamma(X)$.

Properties (P1) to (P3) together can be expressed as $\gamma(aX+b) = \text{sign(a)}\gamma(X)$, for any $a, b \in \mathbb{R}$. The reason these properties are separated is from a historical point of view. For the earliest notions of skewness, it was already a base criterion that a measure of skewness should be unaffected by scaling or location shifts. Also the measure equaling zero under symmetry was unanimously agreed upon. So what then with negative scaling? The most logical line of thought starting from Properties (P1) and (P2) would be Property (P3) as a reflection of the distribution makes a left skewed distribution right skewed and vice versa, without changing anything else. So the sign of the skewness changes, but not the magnitude. Property (P4) saw life in [45] and is used to put an ordering based on a skewness measure, i.e., when the condition holds for F_X and G, then $\gamma(F_X) \leq \gamma(G)$ always. Over the course of time, many other conditions have been stated, depending on the skewness measure. A nice review is given in [32]. The condition stated in Property (P4) is among the strongest, but is arguably not a true necessity as other ad hoc ordering schemes can be devised. Examples of such can be found in [5,15]. In essence, this property can thus be replaced with any other condition which, if it holds for a class of distributions, puts an ordering on the entire class based on its skewness.

One could argue that for an easier interpretation, $\gamma(X) \in [-1, 1]$, which ensures boundedness of the measure. However, as most basic measures of skewness do not satisfy this, we do not impose this boundedness, at a first stage.

The remainder of the paper is structured as follows. In Section 2 a small review on existing univariate and multivariate asymmetry measures is given. In Section 3 we propose a new measure of asymmetry as well as a multivariate extension of it. Section 4 calculates the newly proposed asymmetry measure for several examples. Section 5 studies methods for estimating the measure and establishes some asymptotic properties of the estimators. In Section 6 we compare, for some illustrative examples, our measure with the classical measure of asymmetry, and finally in Section 7 we provide some concluding remarks. The paper comes with some Supplementary Material, containing additional information in the form of tables and figures to accompany Sections 4 and 6.

2. Review on existing asymmetry measures

2.1. Univariate measures of asymmetry

Based on the three classical measures of location (mean, median and mode), at the end of the 19th century Pearson proposed a few measures to express asymmetry. These were, as proposed in respectively [38,39],

$$\frac{\mu_X - \mathcal{M}_X}{\left\{E[(X-\mu_X)^2]\right\}^{1/2}}, \qquad \frac{3(\mu_X - M_X)}{\left\{E[(X-\mu_X)^2]\right\}^{1/2}}. \tag{1}$$

Later on, as more knowledge and insights in statistics and distributional theory became available, other measures of asymmetry have been introduced. These extend the above two by taking into account extra information of the distribution. The primary examples are the following ones.

- An asymmetry measure based on first and second order differences between quantiles (see [12,48]):
$$\frac{Q_X(0.25) - 2Q_X(0.5) + Q_X(0.75)}{Q_X(0.75) - Q_X(0.25)}. \tag{2}$$
This is known as the Bowley coefficient of skewness.
- Classical skewness (also known as Pearson's moment coefficient of skewness):
$$\text{SK}(X) = \frac{E[(X-\mu_X)^3]}{\left\{E[(X-\mu_X)^2]\right\}^{3/2}}. \tag{3}$$
- A skewness measure based on the difference between mean and median (see [21]):
$$\frac{\mu_X - M_X}{E[|X - M_X|]}. \tag{4}$$

For a symmetric distribution, all three measures of location coincide and all of the above measures equal zero. But in deciding which one is best, there is some debate. The most commonly used one however, is the classical skewness (3).

All measures meet the four desirable properties of a good skewness measure. However, they also have their downsides. For measures (1), (3) and (4) the problem is that they are heavily influenced by outliers as they depend on the mean μ_X. Measure (2) does not share this property, but it only takes into account the 50% most central data, thereby neglecting the tails which might also contain useful information on skewness. A last issue with all these measures is interpretability, which from an intrinsic idea viewpoint is clear, but given a value from the measure, it is not immediately clear what this reflects.

Some other measures of skewness have been introduced, but these are more situational or do not fulfill most of the desirable properties mentioned above. The idea of comparing certain quantiles of the distribution however, is a common concept that can be applied to obtain new measures of asymmetry. For more general measures of asymmetry, some options have been proposed. First there are extensions on previous measures of asymmetry. A nice example of this is the measure introduced by [16], see also in [22] among others, building upon the idea behind (2), i.e., considering

$$\frac{Q_X(1-p) - 2Q_X(0.5) + Q_X(p)}{Q_X(1-p) - Q_X(p)}, \qquad 0 < p < 1. \tag{5}$$

This circumvents to some extent the loss of information in the tails, but gives rise to the question which p one should take. For $p = 0.25$, (2) is obtained, which is also dubbed the quartile skewness. For $p = 0.125$ one obtains the octile skewness etc. A more detailed treatment on related skewness measures is given in [3]. For robustness reasons, considering asymmetry around the mean is not an option.

A quantile function fully characterizes a distribution function. In that sense one could replace quantiles in (5) also by expectiles, that also fully characterize a distribution function, and exist when X has a finite first moment. For given p, the pth order expectile, denoted $\mathrm{e}_X(p)$, satisfies $p = E\left[\,|X - \mathrm{e}_X(p)| I\{X \le \mathrm{e}_X(p)\}\,\right]\{E\left[\,|X - \mathrm{e}_X(p)|\,\right]\}^{-1}$. In contrast to quantiles, expectiles are sensitive to information in the tails (since expectation based) and in that sense circumvent that drawback of a quantiles based measure of asymmetry. In [18] the authors introduce and study the expectile-based measure of skewness obtained by replacing in (5) quantiles by expectiles. Multivariate extensions of (5), as well as of its expectile-based alternative are however not straightforward, requiring defining quantiles and expectiles in a multivariate setting. For a detailed study of this expectile-based measure of skewness in a univariate setting see [18].

In [37] the measure

$$\eta(X) = \begin{cases} -\mathrm{Corr}(F_X(X), f_X(X)) & \text{if } 0 < \mathrm{Var}(f_X(X)) < \infty, \\ 0 & \text{if } \mathrm{Var}(f_X(X)) = 0, \end{cases} \tag{6}$$

is used as a tool to quantify asymmetry. The measure in expression (6) fulfills the first three requirements for a good asymmetry measure, but not the fourth one. The advantage of this measure is that it also makes sense for a monotonically increasing/decreasing density function. In [17] critique on this measure is posed as a tool to quantify asymmetry. The main comment being that asymmetry for monotonical functions is very subjective with respect to the observer's view and that the asymmetry measure (6) is actually just a test for exponentiality, i.e., it yields 0 for the exponential distribution only.

Throughout the rest of the paper we assume densities, univariate or multivariate, to be unimodal. When the distribution is known to be skewed but unimodal, intuitively it makes more sense to compare both sides of the density function around its mode as we want to compare both halves of a distribution with respect to its peak. There are also more practical reasons why the mode might be a better location measure in this context, as mentioned in [23]. A possible measure can thus be the total probability mass on each side of the mode, but also distances between the mode and a pre-specified point. In [11] the "odds-asymmetry" function is proposed. Assume that X is a continuous random variable with a unimodal density f_X which is rooted (i.e., a density with limiting values zero at both endpoints of its support). Then for each $p \in (0, 1)$ there exist two points $x_\ell(p)$ and $x_r(p)$, one to the left and one to the right of the mode, for which $f_X(x_\ell(p)) = f_X(x_r(p)) = pf_X(\mathcal{M}_X)$. The odds-asymmetry function is then constructed by comparing the distances of $x_\ell(p)$ and $x_r(p)$ to the mode. If we denote these with $d_\ell(p) = \mathcal{M}_X - x_\ell(p)$ and $d_r(p) = x_r(p) - \mathcal{M}_X$, then the odds-asymmetry function is given by

$$\rho : (0, 1) \to (0, \infty) : p \to \rho(p) = \frac{d_r(p)}{d_\ell(p)}. \tag{7}$$

If (7) equals one for each $p \in (0, 1)$, then X is symmetric. This idea is further improved in [15] where the signed proportionate difference function is defined as

$$\gamma : (0, 1) \to (-1, 1) : p \to \gamma(p) = \frac{d_r(p) - d_\ell(p)}{d_r(p) + d_\ell(p)} = \frac{x_r(p) - 2\mathcal{M}_X + x_\ell(p)}{x_r(p) - x_\ell(p)}. \tag{8}$$

This provides an asymmetry function which is constant 0 in case of symmetry, negative for values of p for which the function is left skewed and positive otherwise. The interesting part about this measure is that it allows for different asymmetry on different parts of the density, which gives more detailed information compared to a single measure for the entirety of X. If one still wants a single scalar asymmetry measure, an option is to integrate (8) over p which results in the overall (integrated) measure $\gamma^* = 1 - 2F_X(\mathcal{M}_X)$, which is the measure of asymmetry proposed in [4] (see also [15] for details on this). The condition that f_X needs to be rooted can also be dropped, but this limits the range of values p can take.

In Fig. 1 we illustrate the idea and construction of the measures (7) and (8) for $p = 0.55$, on the plotted density. The red horizontal line indicates the value of $pf_X(\mathcal{M}_X)$.

2.2. Multivariate measures of asymmetry

Consider now a d-dimensional random variable $\mathbf{X} \in \mathbb{R}^d$ with mean $\boldsymbol{\mu}_\mathbf{X}$ and covariance matrix $\boldsymbol{\Sigma}$. Throughout the paper all vectors are denoted as column vectors, with $\mathbf{X}^\top$ denoting the transpose of a vector $\mathbf{X}$ (or matrix). Further, vectors and matrices will be denoted with boldface notations. Some multivariate measures of skewness for $\mathbf{X}$ have been proposed. The first notable attempt was made in [34], in which the author proposes a rather straightforward multivariate extension

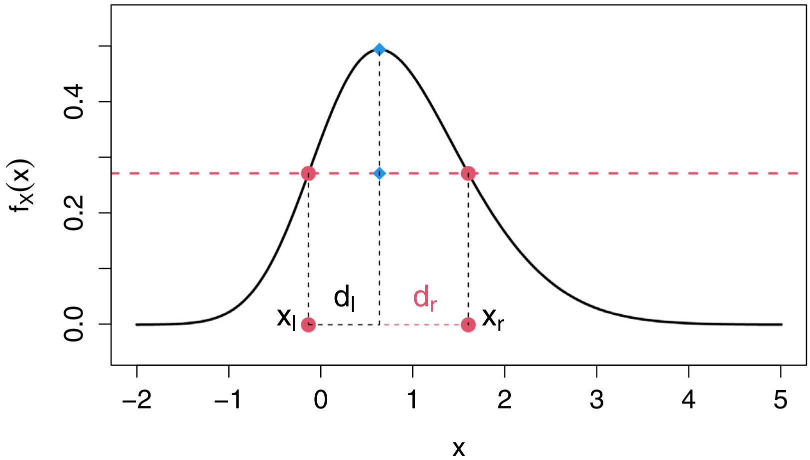


Fig. 1. Graphical illustration of ideas behind (7) and (8) for the plotted density and $p = 0.55$.

to the univariate skewness measure (3). With $\mathbf{X_1}$ and $\mathbf{X_2}$ being independent copies of $\mathbf{X}$, Mardia's skewness measure is defined as

$$\beta_d(\mathbf{X}) = E\left\{\left[(\mathbf{X_1} - \boldsymbol{\mu}_{\mathbf{X}})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X_2} - \boldsymbol{\mu}_{\mathbf{X}})\right]^3\right\}. \tag{9}$$

For standardized variables $\mathbf{Y} = \boldsymbol{\Sigma}^{-1/2}\,(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \in \mathbb{R}^d$, this can be reduced to

$$\beta_d(\mathbf{Y}) = \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{\ell=1}^{d}\left[E(Y_jY_kY_\ell)\right]^2.$$

This measure is by far the most well known and popular way to express multivariate skewness. Appealing properties are that β_d is affine invariant and easy to compute numerically. However it has also received some criticism as no negative values are allowed. With $\mathbf{Y}_1$ and $\mathbf{Y}_2$ two independent copies of $\mathbf{Y}$, the above measure equals $E\left[\left(\mathbf{Y}_1^\top\mathbf{Y}_2\right)^3\right]$, as was noticed by [25]. As stated in, for example, [25,44], $\beta_d(\mathbf{Y})$ is just the norm of the vector containing all third order cumulants of $\mathbf{Y}$. Let $\boldsymbol{\kappa}_3$ denote this vector of cumulants.

Some extensions of Mardia's skewness measure have subsequently been proposed. The first one is for the standardized variables $\mathbf{Y}$ proposed in [35], which is a vector valued measure expressing skewness for the different variables

$$\mathbf{s}(\mathbf{Y}) = E\left[\left(\sum_{j=1}^{d} Y_j^2\right)\mathbf{Y}\right].$$

This measure has then been further adapted in [29] to fill the gap created by not including all interactions between the components of $\mathbf{Y}$. In that paper Kollo's measure of skewness is proposed, defined as

$$\mathbf{b}(\mathbf{Y}) = E\left[\left(\sum_{j=1}^{d}\sum_{k=1}^{d} Y_jY_k\right)\mathbf{Y}\right]. \tag{10}$$

This measure of multivariate skewness has recently gained popularity in the financial context to model skewness in e.g., asset returns. See further for example Section 6.1 in [44] for more discussion on these and other multivariate measures of skewness.

The measure in (10) can be zero, even for asymmetric distributions, as mentioned in for example [25]. In that paper it is further remarked that the measures $\mathbf{s}(\mathbf{Y})$ and $\mathbf{b}(\mathbf{Y})$ (among others) can also be expressed in terms of $\boldsymbol{\kappa}_3$, the vector of third order cumulants of $\mathbf{Y}$. The authors in [25] then propose further improvements of the above measures of multivariate skewness by considering only distinct elements of cumulant vectors (of third and other orders).

A second flow of multivariate skewness measures is obtained by projecting the data on a one-dimensional hyperplane and calculating the univariate skewness (mostly (3), but in general any other univariate skewness measure can be used) of the projected data. The first such type of skewness measure was proposed in [33], and is defined as the maximal skewness that can be obtained by projecting the data on a vector $\mathbf{u}$ on the unit sphere $\mathcal{S}^d = \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\| = 1\}$

$$\beta_d^*(\mathbf{Y}) = \sup_{\mathbf{u}\in\mathcal{S}^d}\left\{E\left[\left(\mathbf{u}^\top\mathbf{Y}\right)^3\right]\right\}^2. \tag{11}$$

A similar measure to this is obtained by taking for $\mathbf{u}$ the eigenvectors of the covariance matrix of $\mathbf{X}$, scaled by the corresponding eigenvalue. Denote by $\boldsymbol{\Lambda}$ the diagonal matrix of eigenvalues, and by $\boldsymbol{\Gamma}$ the orthogonal matrix consisting of columns that are the associated eigenvectors of $\boldsymbol{\Sigma}$, such that $\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^\top = \boldsymbol{\Lambda}$. The measure proposed in [43] is then given by

$$\beta_d^2(\mathbf{X}) = \frac{1}{d}\sum_{j=1}^{d}\left(\frac{E\left\{[\boldsymbol{\Gamma}_{\cdot j}^\top(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})]^3\right\}}{\boldsymbol{\Lambda}_{jj}^{3/2}}\right)^2, \tag{12}$$

with $\boldsymbol{\Gamma}_{\cdot j}$ denoting the jth column of the matrix $\boldsymbol{\Gamma}$, and $\boldsymbol{\Lambda}_{jj}$ the jth diagonal element of $\boldsymbol{\Lambda}$. The interpretation of (12) is the averaged squared skewness along the principal components of the covariance matrix, which is not affine invariant in contrast to (11). See for example [44].

A third group of measures proposed in the literature are often situation specifically viable or generalizations of other univariate skewness measures. We in particularly mention Isogai's measure of skewness (see [24]). This is a multivariate extension of the measures in (1). The measure is given by

$$S = (\boldsymbol{\mu}_\mathbf{X} - \mathcal{M}_\mathbf{X})^\top g^{-1}(\boldsymbol{\Sigma})(\boldsymbol{\mu}_\mathbf{X} - \mathcal{M}_\mathbf{X}),$$

with $g(.)$ an appropriate function, which is more often than not taken to be the identity function. Another measure is proposed in [42] and is based on Renyi's entropy of a specific order. Formally stated in [9], the measure is given by

$$S(f) = \mathrm{var}(\log(f_\mathbf{X}(\mathbf{X}))) \approx G'(\boldsymbol{\mu}_\mathbf{X})^\top \boldsymbol{\Sigma} G'(\boldsymbol{\mu}_\mathbf{X}),$$

with $G'(\cdot)$ the first derivative of $\log(f_\mathbf{X}(\cdot))$.

Of all these criteria, only those of Mardia and Kollo (provided in expressions (9) and (10) respectively) are frequently used in practice and one can create new criteria by combining third order cumulants relevant to the application at hand.

To conclude with, a multivariate extension of (8) is proposed in [1]. Let $\mathbf{x}_0 \in \mathcal{S}^d$ be given. Take the d-variate random vector $\mathbf{X}$ with unimodal density function $f_\mathbf{X}(\mathbf{x})$ and mode $\mathcal{M}_\mathbf{X}$. Denote with $a_+(p, \mathbf{x}_0)$ and $a_-(p, \mathbf{x}_0)$ the positive, respectively negative solution of $f_\mathbf{X}(\mathcal{M}_X + a\mathbf{x}_0) = pf_\mathbf{X}(\mathcal{M}_\mathbf{X})$. The multivariate skewness measure proposed in [1] is then given by

$$\gamma(p, \mathbf{x}_0) = \frac{a_+(p, \mathbf{x}_0) + a_-(p, \mathbf{x}_0)}{a_+(p, \mathbf{x}_0) - a_-(p, \mathbf{x}_0)}. \tag{13}$$

This measure depends on the choice of $\mathbf{x}_0$, which can be seen as a direction in which the skewness function is to be evaluated. A global measure for skewness, depending only on $\mathbf{x}_0$ can then again be taken by averaging or integrating (13) over p.

3. Newly proposed asymmetry index

3.1. A new distance based index of asymmetry in the univariate setting

The idea of using the mode as measure of location in deriving a univariate asymmetry measure is intriguing, although comes with the burden of having to find the mode, which might not be an easy problem. Especially since for empirical measures, the median is much easier to find. When looking at a multivariate situation, utilizing the mode might gain favor again as quantiles are not uniquely defined in a multivariate setting whereas (under the assumption of unimodality) the mode is still uniquely defined, albeit harder to find. The measure we propose is related to (8), i.e., we opt for an asymmetry function instead of a single scalar (or vector in $\mathbb{R}^d$).

Denote with $X \in \mathbb{R}$ a univariate random variable with density function $f_X(x)$, cumulative distribution function $F_X(x)$ and mode $\mathcal{M}_X$. We propose the measure

$$\gamma_X(s) = \frac{f_X(\mathcal{M}_X + s) - f_X(\mathcal{M}_X - s)}{f_X(\mathcal{M}_X)}, \tag{14}$$

with $s \in \mathcal{S}$, $\mathcal{S} \subset \mathbb{R}$ compact and containing 0. Note that $f_X(\mathcal{M}_X - s) = f_{2\mathcal{M}_X - X}(\mathcal{M}_X + s)$, where $f_{2\mathcal{M}_X - X}$ is the reflection of f_X around its mode $\mathcal{M}_X$.

The interpretation of (14) is best explained graphically; see therefore Fig. 2. In the top panel we depict the univariate skew-normal density (in black solid line), defined as

$$f_X(x; \xi, \omega, \alpha) = 2\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\frac{x-\xi}{\omega}\right), \tag{15}$$

with ϕ and Φ respectively the standard normal density and cumulative distribution function, and where $\xi \in \mathbb{R}$ is a location parameter, $\omega > 0$ is a scale parameter, and $\alpha \in \mathbb{R}$ is a skewing parameter. Presented in the top panel of Fig. 2 is also the reflection $f_{2\mathcal{M}_X - X}(x) = f_X(2\mathcal{M}_X - x)$ of f_X around its mode (in red dashed line). If we take the difference between $f_X(x+s)$ and $f_{2\mathcal{M}_X - X}(x+s)$ for each $s > 0$ and scale this by the maximal attainable difference, i.e., $f_X(\mathcal{M}_X)$, we obtain the bottom panel.

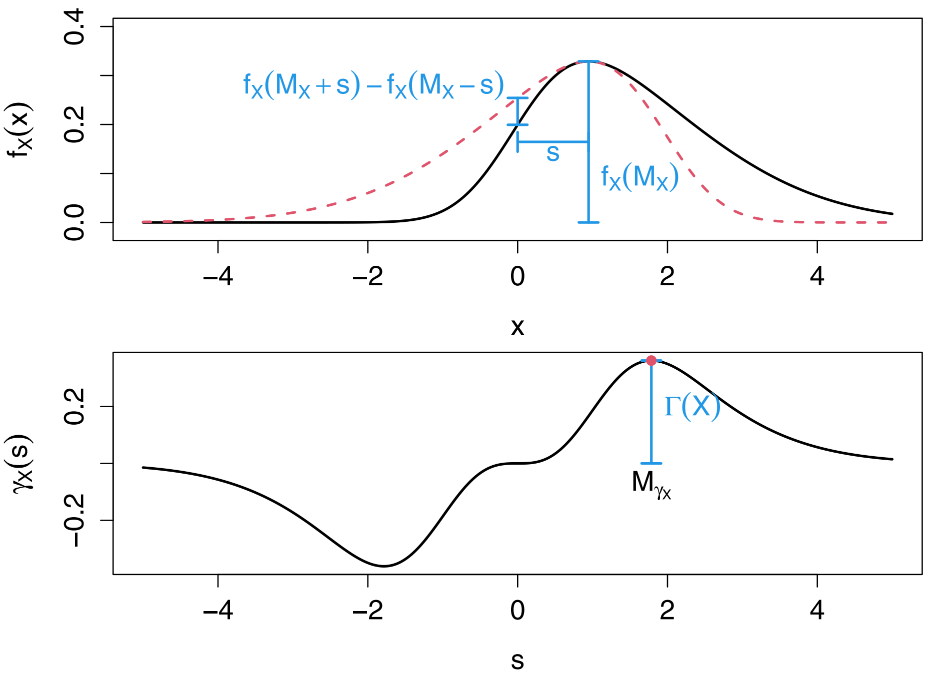


Fig. 2. Illustration of (14) applied to the skew-normal density in (15) with parameters $\xi = 0$, $\omega = 2$ and $\alpha = 3$. The top panel shows f_X (black solid line) and $f_{2\mathcal{M}_X - X}$ (in red dashed line). The bottom panel shows $\gamma_X(s)$.

It is obvious from (14) that $\gamma_X(s)$ is an odd function, hence we get the value of $\gamma_X(s)$ for $s < 0$ for free from these for $s > 0$. In the bottom panel of Fig. 2 two more quantities are indicated. These are

$$\mathcal{M}_{\gamma_X} = \arg\max_{s\in\mathcal{S}}(\gamma_X(s)),$$

i.e., the mode of $\gamma_X(\cdot)$, and

$$\Gamma(X) = \max_{s\in\mathcal{S}} \frac{f_X(\mathcal{M}_X + s) - f_X(\mathcal{M}_X - s)}{f_X(\mathcal{M}_X)}.$$

Under symmetry, $\gamma_X(s) = 0$ everywhere, while for any asymmetry present there exists a $s \neq 0$ where $\gamma_X(s) \neq 0$. The interpretation of (14) is thus the difference between the density and its reflected counterpart around the point of symmetry (in this case the mode) relative to the maximal attainable difference (the density in the mode).

As this is a skewness function (in s), Property (P4) is somewhat difficult to obtain as we are no longer dealing with a single scalar, but with a function that changes as s changes. With this we mean that for s_1 the distribution might be right skewed, but for some $s_2 \neq s_1$ it might just as well be left skewed. But we can use an ordering as proposed in [15].

Theorem 1. *The skewness index as defined in (14) satisfies Properties (P2) and (P3), but not Property (P1). Moreover, a logical ordering is that "X is more skewed to the right than Y i.f.f. $\gamma_X(s) > \gamma_Y(s)$, $\forall s > 0$".*

Proof. Take $Y = aX + b$, with $a \in \mathbb{R} \setminus \{0\}$ and $b \in \mathbb{R}$. Then Y has density function $f_Y(y) = \text{sign}(a)a^{-1}f_X((y - b)/a)$ and mode $\mathcal{M}_Y = a\mathcal{M}_X + b$. Hence we obtain that

$$\begin{aligned}\gamma_Y(s) = \frac{f_Y(\mathcal{M}_Y + s) - f_Y(\mathcal{M}_Y - s)}{f_Y(\mathcal{M}_Y)} &= \frac{\frac{1}{a}f_X(\frac{a\mathcal{M}_X+b+s-b}{a}) - \frac{1}{a}f_X(\frac{a\mathcal{M}_X+b-s-b}{a})}{\frac{1}{a}f_X(\frac{a\mathcal{M}_X+b-b}{a})} \\ &= \frac{f_X(\mathcal{M}_X + \frac{s}{a}) - f_X(\mathcal{M}_X - \frac{s}{a})}{f_X(\mathcal{M}_X)} = \gamma_X\left(\frac{s}{a}\right).\end{aligned}$$

From this, it immediately follows that $\gamma_Y(s)$ is a rescaled version of $\gamma_X(s)$ and hence location invariant, but not scale invariant. Property (P1) is thus only partially satisfied. By taking $a = -1$ we also immediately obtain that $\gamma_{-X}(s) = \gamma_X(-s) = -\gamma_X(s)$, as $\gamma_X(s)$ is an odd function in s. This makes that Property (P3) is satisfied. For Property (P2), note that for symmetric X, $f_X(\mathcal{M}_X - s) = f_X(\mathcal{M}_X + s)$, for all s. Since the mode is the point of symmetry for unimodal distributions, from (14) it follows that $\gamma_X(s) = 0$ for symmetric X. □

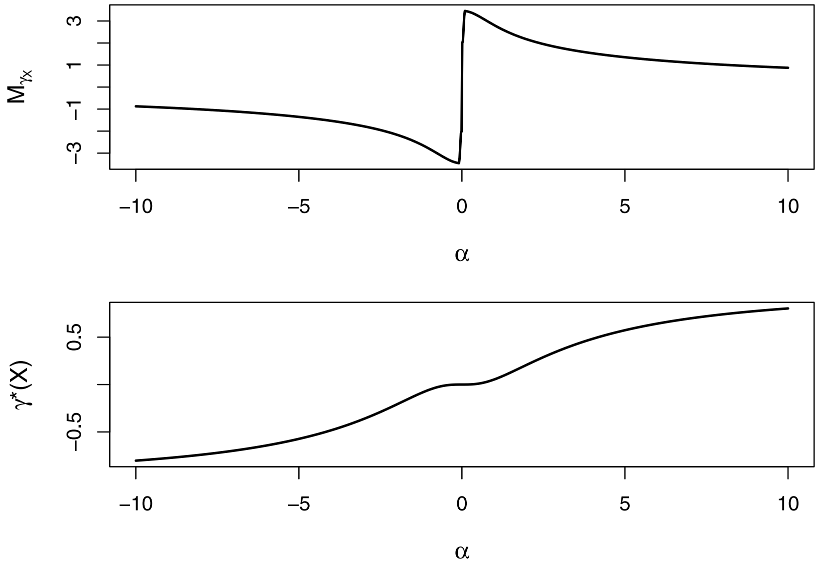


Fig. 3. Illustration of (16) applied to a skew-normal distribution (15) with parameters $\xi = 0$, $\omega = 2$. The top panel shows $\mathcal{M}_{\gamma_X}$, the bottom panel shows $\gamma^*(X)$ as function of α.

Furthermore the asymmetry function $\gamma_X(s)$ satisfies some basic properties, which are summarized in Corollary 1.

Corollary 1. *For the asymmetry function $\gamma_X(s)$ the following holds.*

(i) *The function $\gamma_X(s)$ is an odd function.*
Consequently $\mathcal{M}_{\gamma_X}$ may alternatively be obtained from $\arg\max_{s \geq 0} |\gamma_X(s)|$ multiplied with the sign of $\gamma_X(\cdot)$ at this argument.

(ii) *Suppose that X is rooted at the two tails, i.e., with $a < b$ the lower and upper endpoints of the support of X, it holds that $\lim_{x\to a} f_X(x) = 0 = \lim_{x\to b} f_X(x)$. Under this assumption it holds that $\lim_{s\to\pm\infty} \gamma_X(s) = 0$.*

(iii) *The function $\gamma_X(s)$ is bounded: $-1 \leq \gamma_X(s) \leq 1$.*

(iv) *A scale invariant version can be obtained by, for example, defining $s_X := \frac{s}{\sqrt{var(X)}}$ and considering $\gamma_X(s_X)$.*

Proof. Statements (i) and (ii) are immediate from the definition of $\gamma_X(s)$.

For statement (iii) it suffices to recall that f_X is unimodal with unique mode $\mathcal{M}_X$ and hence it holds that for any s,

$$0 \leq f_X(\mathcal{M}_X - s) \leq f_X(\mathcal{M}_X), \qquad 0 \leq f_X(\mathcal{M}_X + s) \leq f_X(\mathcal{M}_X),$$

which implies that

$$-f_X(\mathcal{M}_X) \leq f_X(\mathcal{M}_X + s) - f_X(\mathcal{M}_X) \leq f_X(\mathcal{M}_X + s) - f_X(\mathcal{M}_X - s) \leq f_X(\mathcal{M}_X),$$

from which the statement follows.

Statement (iv) follows immediately from the proof of Theorem 1. □

As for a scalar summarizing measure of (14), we take into account both the location of $\gamma_X(s)$ and its value. This is with a multivariate extension in mind as it gives the direction of asymmetry and the magnitude. The scalar (or vector in higher dimensions) summarizing measure of asymmetry we propose is

$$\gamma^*(X) = \begin{cases} \Gamma(X)\,\text{sign}\left(\mathcal{M}_{\gamma_X}\right) & \text{if } \mathcal{M}_{\gamma_X} \neq 0, \\ 0 & \text{if } \mathcal{M}_{\gamma_X} = 0 \text{ or } [\gamma_X(s) = 0,\ \forall s], \\ 1 & \text{if } X \text{ is a half-type distribution with decreasing density}, \\ -1 & \text{if } X \text{ is a half-type distribution with increasing density}. \end{cases} \tag{16}$$

This measure applied to the same example as above yields Fig. 3, in which $\mathcal{M}_{\gamma_X}$ and $\gamma^*(X)$ are given as a function of α (the skewing parameter) of a skew-normal distribution (15) with $\xi = 0$ and $\omega = 2$. As we would expect, the skewness is gradually increasing as α increases, passing through zero for $\alpha = 0$. This is in line with the intrinsic purpose of the parameter. For $\alpha \to \infty$, $\gamma^*(X)$ approaches one and the distribution turns into a half-normal distribution (defined on the positive halfline), with decreasing density.

Proposition 1. *The measure $\gamma^*(X)$ proposed in (16) satisfies the desirable Properties (P1)–(P3) mentioned above. For an ordering based on (16) we use that X is more skewed to the right than Y i.f.f. $\gamma^*(X) > \gamma^*(Y)$.*

Proof. Location invariance still holds because both $\Gamma(X)$ and $\mathcal{M}_X$ are location invariant. However, scale invariance also holds since for $Y = aX$, with $a > 0$

$$\Gamma(Y) = \max_{s\in\mathcal{S}}(\gamma_Y(s)) = \max_{s\in\mathcal{S}}(\gamma_Y(as)) = \max_{s\in\mathcal{S}}(\gamma_X(s)) = \Gamma(X),$$

$\frac{\mathcal{M}_{\gamma_X}}{|\mathcal{M}_{\gamma_X}|}$ always has unit length (or is zero) and the sign of $\Gamma(X)$ does not change for $a > 0$. Hence Property (P1) is now satisfied. If X is symmetric, it follows that $\Gamma(X) = 0$ and thus $\gamma^*(X) = 0$ which shows Property (P2) also holds. For Property (P3), it can easily be seen that $\Gamma(X) = \Gamma(-X)$ and $\mathcal{M}_{\gamma_X} = -\mathcal{M}_{\gamma_{\tilde{X}}}$, where we denoted $\tilde{X} = 2\mathcal{M}_X - X$, because $\gamma_X(s)$ is an odd function. □

3.2. Multivariate extension

As we are reflecting a random variable around its mode, a natural extension to the multivariate setting is to apply the same technique as used in (14). Consider a d-variate random vector $\mathbf{X} \in \mathbb{R}^d$, with unimodal density and mode $\mathcal{M}_{\mathbf{X}}$. The measure proposed in (14) then becomes

$$\gamma_{\mathbf{X}}(\mathbf{s}) = \frac{f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s})}{f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}})}, \tag{17}$$

where $\mathbf{s} \in \mathcal{S}$ ($\subset \mathbb{R}^d$ compact and containing $\mathbf{0}$) is a vector with origin in $\mathbf{0}$. Denoting the mode of this function by

$$\mathcal{M}_{\gamma_{\mathbf{X}}} = \arg\max_{\mathbf{s}\in\mathcal{S}} \gamma_{\mathbf{X}}(\mathbf{s}),$$

the multivariate extension of $\mathcal{M}_{\gamma_X}$. The multivariate analogue of (16) then becomes

$$\boldsymbol{\gamma}^*(\mathbf{X}) = \begin{cases} \frac{\Gamma(\mathbf{X})\mathcal{M}_{\gamma_{\mathbf{X}}}}{\|\mathcal{M}_{\gamma_{\mathbf{X}}}\|} & \text{if } \mathcal{M}_{\gamma_{\mathbf{X}}} \neq \mathbf{0}, \\ \mathbf{0} & \text{if } \mathcal{M}_{\gamma_{\mathbf{X}}} = \mathbf{0} \text{ or } [\gamma_{\mathbf{X}}(\mathbf{s}) = 0, \ \forall \mathbf{s}]. \end{cases} \tag{18}$$

Proposition 2. *Both $\gamma_{\mathbf{X}}(\mathbf{s})$ and $\boldsymbol{\gamma}^*(\mathbf{X})$ are translation invariant, but not scale invariant. $\Gamma(\mathbf{X})$ however, is affine invariant.*

Proof. Let $\mathbf{Y} = \mathbf{AX} + \mathbf{B}$ with $\mathbf{A} \in \mathbb{R}^{d\times d}$ non-singular and $\mathbf{B} \in \mathbb{R}^d$. Denote by $\gamma_{\mathbf{X}}(\mathbf{s})$ and $\gamma_{\mathbf{Y}}(\mathbf{s})$ the asymmetry function for respectively $\mathbf{X}$ and $\mathbf{Y}$. Also denote $\mathcal{M}_{\gamma_{\mathbf{X}}} = \arg\max_{\mathbf{s}\in\mathbf{S}}(\gamma_{\mathbf{X}}(\mathbf{s}))$, $\mathcal{M}_{\gamma_{\mathbf{Y}}} = \arg\max_{\mathbf{s}\in\mathbf{S}}(\gamma_{\mathbf{Y}}(\mathbf{s}))$, $\Gamma(\mathbf{X}) = \max_{\mathbf{s}\in\mathbf{S}}(\gamma_{\mathbf{X}}(\mathbf{s}))$ and $\Gamma(\mathbf{Y}) = \max_{\mathbf{s}\in\mathbf{S}}(\gamma_{\mathbf{Y}}(\mathbf{s}))$. We have that the density of $\mathbf{Y}$ is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{B})\right).$$

Consequently, it holds that $\mathcal{M}_{\mathbf{Y}} = \mathbf{A}\mathcal{M}_{\mathbf{X}} + \mathbf{B}$. Therefore we get

$$\begin{aligned} \gamma_{\mathbf{Y}}(\mathbf{s}) &= \frac{\frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathcal{M}_{\mathbf{Y}} - \mathbf{B} + \mathbf{s})\right) - \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathcal{M}_{\mathbf{Y}} - \mathbf{B} - \mathbf{s})\right)}{\frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathcal{M}_{\mathbf{Y}} - \mathbf{B})\right)} \\ &= \frac{f_{\mathbf{X}}\left(\mathcal{M}_{\mathbf{X}} + \mathbf{A}^{-1}\mathbf{s}\right) - f_{\mathbf{X}}\left(\mathcal{M}_{\mathbf{X}} - \mathbf{A}^{-1}\mathbf{s}\right)}{f_{\mathbf{X}}\left(\mathcal{M}_{\mathbf{X}}\right)} = \gamma_{\mathbf{X}}(\mathbf{A}^{-1}\mathbf{s}). \end{aligned}$$

From this the translation invariance of $\gamma_{\mathbf{X}}(\mathbf{s})$ follows, and also its scale non-invariance. Moreover we obtain that $\mathcal{M}_{\gamma_{\mathbf{Y}}} = \mathbf{A}\mathcal{M}_{\gamma_{\mathbf{X}}}$, and $\Gamma(\mathbf{Y}) = \Gamma(\mathbf{X})$. Consequently it holds that $\boldsymbol{\gamma}^*(\mathbf{Y}) = \frac{\Gamma(\mathbf{Y})\mathbf{A}\mathcal{M}_{\gamma_{\mathbf{X}}}}{\|\mathbf{A}\mathcal{M}_{\gamma_{\mathbf{X}}}\|}$, and hence $\boldsymbol{\gamma}^*(\mathbf{X})$ is thus location, but not scale invariant. □

That $\boldsymbol{\gamma}^*(\mathbf{Y})$ is not scale invariant, unlike in the univariate case, is to be expected. In terms of scale it remains the same (i.e., in norm), but because the measure also takes into account the direction of asymmetry, it is to be expected that any sort of rotation induces a change in the main direction of asymmetry.

Similar properties as these stated in Corollary 1 continue to hold for the multivariate extension $\gamma_{\mathbf{X}}$. Statement (i) (first part), and statements (ii) and (iii) are straightforward to see. Concerning the generalization of statement (iv), note from the proof of Proposition 2 that for $\mathbf{Y} = \mathbf{AX}$, with $\mathbf{A} \in \mathbb{R}^{d\times d}$ a non-singular matrix, we have that $\gamma_{\mathbf{Y}}(\mathbf{s}) = \gamma_{\mathbf{X}}\left(\mathbf{A}^{-1}\mathbf{s}\right)$. Therefore by defining, for $\mathbf{X}$ with variance–covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$, the vector $\mathbf{s}_{\mathbf{X}} = \boldsymbol{\Sigma}_{\mathbf{X}}^{-1/2}\mathbf{s}$ and considering $\gamma_{\mathbf{X}}(\mathbf{s}_{\mathbf{X}})$ we obtain a scale-free or scale invariant version of $\gamma_{\mathbf{X}}(\cdot)$.

Our proposed asymmetry measure has as a benefit over (13) that there is no dependency on a chosen direction $\mathbf{x}_0$ and gives a full overview of the asymmetry of the entire distribution instead of just one cross section. The downside however is that when no closed form expression for the density is available it requires evaluation of the density on a grid. In lower dimensional settings however, this is still a feasible task.

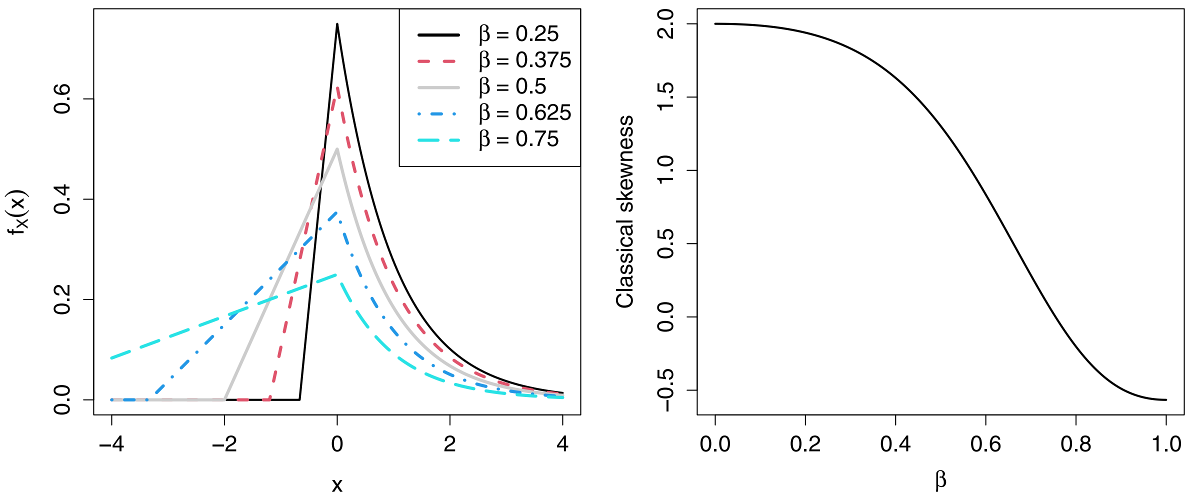


Fig. 4. Left: Plots of the density in (19) for various values of β. Right: The classical skewness measure $\mathrm{SK}(X)$, defined in (3), as a function of β.

4. Examples

4.1. Simple illustrative example

We start with a simple example, originating from [46], which is often used as a counterexample on the rule of thumb for checking right-skewness which states: "For a right-skewed distribution $\mathcal{M}_X < M_X < \mu_X$". Consider the density function

$$f_X(x;\beta) = \begin{cases} 0 & \text{if } x < -\frac{2\beta}{1-\beta}, \\ (1-\beta)(1+\frac{1-\beta}{2\beta}x) & \text{if } -\frac{2\beta}{1-\beta} \le x < 0, \\ (1-\beta)e^{-x} & \text{if } x \ge 0. \end{cases} \tag{19}$$

The parameter $\beta \in (0, 1)$ determines the amount of mass in the triangular region. Fig. 4 (left panel) depicts the density for several values of β. Note that the density is asymmetric, for any value of β. The classical skewness measure $\mathrm{SK}(X)$ in (3) can be calculated, noting that the kth moment of the density in (19), for $k \in \mathbb{N}$ exists and equals

$$E(X^k) = (1-\beta)c_\beta^{k+1}\left[\frac{1}{k+1}(-1)^{k+2} + \frac{1}{k+2}(-1)^{k+3}\right] + (1-\beta)k!, \tag{20}$$

where the constant $c_\beta = 2\beta/(1-\beta)$. Since $\mathrm{SK}(X) = \left[E(X^3) - 3\mu_X \mathrm{Var}(X) - \mu_X^3\right] [\mathrm{Var}(X)]^{-3/2}$, an expression for the classical skewness measure is easily obtained using (20) for $k = 1, 2$ and 3, and is given by

$$\mathrm{SK}(X) = \frac{6 - \frac{1}{20}c_\beta^4 - 3(1-\beta)\left[1 - \frac{c_\beta^2}{6}\right]\left[2 + \frac{c_\beta^3}{12}\right] + 2(1-\beta)^2\left[1 - \frac{c_\beta^2}{6}\right]^3}{(1-\beta)^{1/2}\left\{\left[2 + \frac{c_\beta^3}{12}\right] - (1-\beta)\left[1 - \frac{c_\beta^2}{6}\right]^2\right\}^{3/2}}. \tag{21}$$

This classical skewness measure, as a function of the parameter β, is depicted in the right panel of Fig. 4.

From (21) it is easily seen that when β tends to zero, then $\mathrm{SK}(X)$ tends to the value 2, whereas for β tending to one, $\mathrm{SK}(X)$ tends to $-2^3/5 \approx -0.565$. For $\beta = 0.75$, the measure $\mathrm{SK}(X)$ is approximately 0.023, which would be interpreted as a density that is slightly skewed to the right. However in this example, skewness is hard to express via a single scalar. For values of β between, say, 0.5 and 0.75, it seems that the density is left skewed. The finite left and infinite right tail however counteract this in the classical skewness, resulting in a positive skewness.

By considering our functional measure, a more detailed picture can be obtained. Obviously the mode of the density (19) is zero and $f_X(0;\beta) = 1-\beta$. It is further easy to calculate that

$$\gamma_X(s) = \begin{cases} -e^s & \text{if} \quad s < -\frac{2\beta}{1-\beta}, \\ -e^s + 1 + \frac{1-\beta}{2\beta}s & \text{if} \quad -\frac{2\beta}{1-\beta} \le s < 0, \\ e^{-s} - 1 + \frac{1-\beta}{2\beta}s & \text{if} \quad 0 \le s \le \frac{2\beta}{1-\beta}, \\ e^{-s} & \text{if} \quad s > \frac{2\beta}{1-\beta}, \end{cases} \qquad \mathcal{M}_{\gamma_X} = \begin{cases} \frac{2\beta}{1-\beta} & \text{if} \quad 0 < \beta \le 1/3, \\ \frac{2\beta}{1-\beta} & \text{if} \quad 1/3 < \beta < 1,\ [h_1(\beta) \ge h_2(\beta)], \\ \ln\left(\frac{2\beta}{1-\beta}\right) & \text{if} \quad 1/3 < \beta < 1,\ [h_1(\beta) < h_2(\beta)], \end{cases}$$

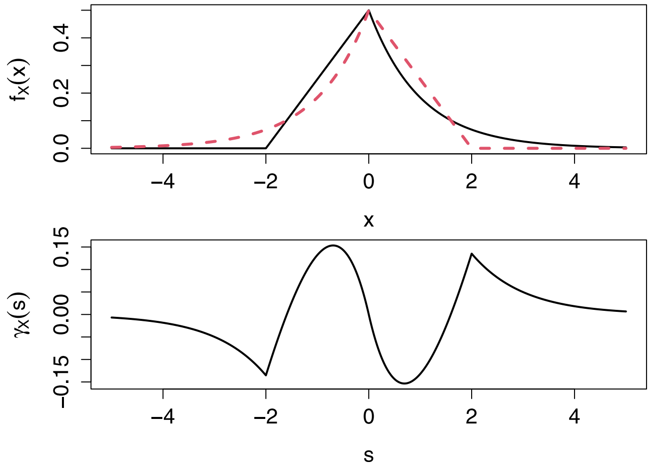


Fig. 5. Density (19) with $\beta = 0.50$, together with the density for $2\mathcal{M}_X - X$ (dashed red line) and $\gamma_X(s)$.

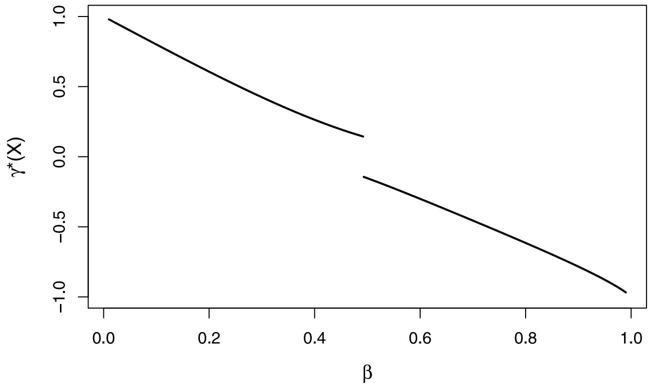


Fig. 6. The function $\gamma^*(X)$ for the density in Eq. (19).

where $h_1(\beta) = \exp\left\{-\frac{2\beta}{1-\beta}\right\}$ and $h_2(\beta) = -\frac{1-\beta}{2\beta} + 1 + \frac{1-\beta}{2\beta}\ln\left(\frac{1-\beta}{2\beta}\right)$. Noting that the two functions h_1 and h_2 cross at the value $\beta \approx 0.4927$, and $h_1(\beta) \geq h_2(\beta)$ (respectively $h_1(\beta) < h_2(\beta)$) for values of $\beta \leq 0.4927$ (respectively $\beta > 0.4927$) we then get

$$\mathcal{M}_{\gamma_X} = \begin{cases} \frac{2\beta}{1-\beta} & \text{if} \quad 0 < \beta \leq 0.4927, \\ \ln\left(\frac{2\beta}{1-\beta}\right) & \text{if} \quad 0.4927 < \beta < 1, \end{cases} \qquad \Gamma(X) = \gamma_X\left(\mathcal{M}_{\gamma_X}\right) = \begin{cases} h_1(\beta) & \text{if} \quad 0 < \beta \leq 0.4927, \\ h_2(\beta) & \text{if} \quad 0.4927 < \beta < 1, \end{cases}$$

and $\gamma^*(X) = \text{sign}(\mathcal{M}_{\gamma_X})\Gamma(X)$.

The function $\gamma_X(s)$ is presented in the bottom panel of Fig. 5, for $\beta = 0.50$. The mode $\mathcal{M}_{\gamma_X}$ is positive for $\beta \leq 0.4927$ and negative for $\beta > 0.4927$. The same holds then of course for the overall measure $\gamma^*(X)$, which is plotted in Fig. 6 as a function of β. It can be seen that the measure $\gamma^*(X)$ can reach maximal asymmetry (i.e., half distributions). Indeed this happens for β tending to zero: density (19) gives in this limit the standard exponential density. The overall measure is never equal to zero. This is in a way natural as this type of distribution can never be symmetric. In that light there is a clear advantage over e.g., moment based skewness as this can give a skewness of 0, indicating symmetry.

4.2. Elliptical distributions

The general formulation of a multivariate elliptical distribution based on a univariate density generator $\widetilde{f}$ is (according to [7])

$$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}\det(\boldsymbol{\Sigma})^{\frac{1}{2}} k_d}\widetilde{f}\left((\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right). \tag{22}$$

In this, $\boldsymbol{\mu} \in \mathbb{R}^d$ is a location parameter vector, $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ a positive definite matrix, and $\widetilde{f}$ such that it has mean zero, unit standard deviation and

$$k_d = \int_0^\infty t^{d-1}\widetilde{f}(t^2)dt < \infty.$$

For almost all popular densities, $\widetilde{f}$ is a unimodal symmetric density function. Location and scale shifts are introduced by respectively $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. We then have the following proposition.

Proposition 3. *Let $\mathbf{X} \in \mathbb{R}^d$ with density function $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ be an elliptical distribution as in (22), generated by a symmetric, unimodal, standardized density generator $\widetilde{f}$. Then it holds that $\forall \mathbf{s} \in \mathcal{S}$, $\gamma_{\mathbf{X}}(\mathbf{s}) = \mathbf{0}$ and subsequently $\boldsymbol{\gamma}^*(\mathbf{X}) = \mathbf{0}$.*

Proof. The proof is simple. Since $\widetilde{f}$ is symmetric, unimodal with mode 0, and standardized, we have from (22) that $\mathcal{M}_{\mathbf{X}} = \boldsymbol{\mu}$. Hence

$$\begin{aligned} f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}\det(\boldsymbol{\Sigma})^{\frac{1}{2}}k_d}\widetilde{f}\left((\mathcal{M}_{\mathbf{X}} - \mathbf{s} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathcal{M}_{\mathbf{X}} - \mathbf{s} - \boldsymbol{\mu})\right) \\ &= \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}\det(\boldsymbol{\Sigma})^{\frac{1}{2}}k_d}\widetilde{f}\left((-\mathbf{s})^\top \boldsymbol{\Sigma}^{-1}(-\mathbf{s})\right) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}\det(\boldsymbol{\Sigma})^{\frac{1}{2}}k_d}\widetilde{f}\left(\mathbf{s}^\top \boldsymbol{\Sigma}^{-1}\mathbf{s}\right) \\ &= \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}\det(\boldsymbol{\Sigma})^{\frac{1}{2}}k_d}\widetilde{f}\left((\mathcal{M}_{\mathbf{X}} + \mathbf{s} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathcal{M}_{\mathbf{X}} + \mathbf{s} - \boldsymbol{\mu})\right) = f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Sigma}). \end{aligned}$$

Thus $\gamma_{\mathbf{X}}(\mathbf{s})$ is zero, $\forall \mathbf{s} \in \mathcal{S}$, and so is $\boldsymbol{\gamma}^*(\mathbf{X})$. □

This implies that for any multivariate unimodal elliptical distribution (22), the measures $\gamma_{\mathbf{X}}(\mathbf{s})$ and $\boldsymbol{\gamma}^*(\mathbf{X})$ are zero. This is what one would expect.

4.3. Skew-normal and skew-elliptical distributions

We next consider the skew-normal distribution. Proposed in [6], the density function of a univariate skew-normal random variable is given by

$$\phi(x; \alpha) = 2\phi(x)\Phi(\alpha x) \qquad x \in \mathbb{R},$$

with $\alpha \in \mathbb{R}$ a skewing parameter. Introducing a location parameter $\xi \in \mathbb{R}$ and a scale parameter $\omega > 0$ leads to the density in (15).

A multivariate extension to the skew-normal, and in fact a whole family of skew-symmetric distributions, has emerged the years afters. These are dubbed skew-elliptical distributions and have as density function

$$h_{\mathbf{X}}(\mathbf{x}) = 2f_d(\mathbf{x})G(w(\mathbf{x})), \qquad \mathbf{x} \in \mathbb{R}^d, \tag{23}$$

with f_d an elliptical density as in (22) of a random variable $\mathbf{Y}$, G the cumulative distribution function of an absolutely continuous random variable T, symmetric around 0 and independent of $\mathbf{Y}$, and w a real-valued function antisymmetric around zero (i.e., $w(-\mathbf{x}) = -w(\mathbf{x})$, for all $\mathbf{x}$), called the skewing function. A particular choice of odd function is $w(\mathbf{x}) = \boldsymbol{\alpha}^\top\mathbf{x}$, with $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^\top \in \mathbb{R}^d$, the vector of skewing parameters.

Here we focus primarily on the skew-normal case, obtained from (23) by taking f_d a multivariate normal density with zero mean vector and with variance–covariance matrix $\bar{\boldsymbol{\Omega}}$, and taking $w(\mathbf{x}) = \boldsymbol{\alpha}^\top\mathbf{x}$. Denote by $\mathbf{X}_0$ a random vector with this skew-normal density. Location and scale parameters, denoted by respectively $\boldsymbol{\xi}$ and a diagonal matrix $\boldsymbol{\omega} = \mathrm{diag}(\omega_1, \dots, \omega_d) \in \mathbb{R}^{d\times d}$, are then introduced by considering the random vector $\mathbf{X} = \boldsymbol{\omega}\mathbf{X}_0 + \boldsymbol{\xi}$ (see [8]). Note that the mean vector of the random vector $\mathbf{X}$ equals $\boldsymbol{\xi}$ and its variance–covariance matrix is $\boldsymbol{\Omega} = \boldsymbol{\omega}^\top\bar{\boldsymbol{\Omega}}\boldsymbol{\omega}$, the rescaled version of the original variance–covariance matrix of $\mathbf{X}_0$. In this specific setting, (23) becomes

$$f_{\mathbf{X}}(\mathbf{x}) = 2\phi_d(\mathbf{x} - \boldsymbol{\xi}; \boldsymbol{\Omega})\Phi(\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}(\mathbf{x} - \boldsymbol{\xi})), \qquad \mathbf{x} \in \mathbb{R}^d, \tag{24}$$

a d-variate skew-normal distribution with mean vector $\boldsymbol{\xi}$ and variance–covariance matrix $\boldsymbol{\Omega}$, referred to as $\mathrm{SN}_d(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$. In [7], Chapter 5, it is stated that the $\mathrm{SN}_d(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ has a unique mode at

$$\mathcal{M}_{\mathbf{X}} = \boldsymbol{\xi} + \frac{m_0{}^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}, \tag{25}$$

with $\alpha^* = (\boldsymbol{\alpha}^\top\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha})^{1/2}$ and m_0^* is the mode of a univariate $\mathrm{SN}_1(0, 1, \alpha^*)$ distribution for which unfortunately no explicit expression is available. It simplifies however the problem of finding a multivariate mode to that of finding a univariate mode.

For obtaining $\gamma_{\mathbf{X}}(\mathbf{s})$ and in extension $\boldsymbol{\gamma}^*(\mathbf{X})$, we need to find $f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}+\mathbf{s})-f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}-\mathbf{s})$ and $f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}})$. Using (24) and (25) we find

$$
\begin{aligned}
f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}) &= 2c\exp\left\{-\frac{1}{2}(\frac{m_0{}^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha})^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha})\right\}\Phi(\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha})\\
&= 2c\exp\left\{-\frac{1}{2}(\frac{m_0^*}{\alpha^*})^2\boldsymbol{\alpha}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}\right\}\Phi(\alpha^*m_0^*) = 2c\exp\left\{-\frac{1}{2}(m_0^*)^2\right\}\Phi(\alpha^*m_0^*),
\end{aligned}
$$

where c is the normalizing constant of the normal density. Further we get

$$
\begin{aligned}
f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}+\mathbf{s})-f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}-\mathbf{s}) &= 2c\exp\left\{-\frac{1}{2}(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}+\mathbf{s})^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}+\mathbf{s})\right\}\Phi(\boldsymbol{\alpha}^\top\frac{m_0^*}{\alpha^*}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}+\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&\quad -2c\exp\left\{-\frac{1}{2}(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}-\mathbf{s})^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}-\mathbf{s})\right\}\Phi(\boldsymbol{\alpha}^\top\frac{m_0^*}{\alpha^*}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}-\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&= 2c\exp\left\{-\frac{1}{2}(\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}+\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1})(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}+\mathbf{s})\right\}\Phi(m_0^*\alpha^*+\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&\quad -2c\exp\left\{-\frac{1}{2}(\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}-\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1})(\frac{m_0^*}{\alpha^*}\boldsymbol{\omega}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}-\mathbf{s})\right\}\Phi(m_0^*\alpha^*-\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&= 2c\exp\left\{-\frac{1}{2}[(\frac{m_0^*}{\alpha^*})^2\boldsymbol{\alpha}^\top\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}+2\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}+\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}\mathbf{s}]\right\}\Phi(m_0^*\alpha^*+\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&\quad -2c\exp\left\{-\frac{1}{2}[(\frac{m_0^*}{\alpha^*})^2\boldsymbol{\alpha}^\top\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}-2\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}+\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}\mathbf{s}]\right\}\Phi(m_0^*\alpha^*-\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\\
&= 2c\exp\left\{-\frac{1}{2}(m_0^*)^2-\frac{1}{2}\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\left[\exp\left\{-\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\Phi(m_0^*\alpha^*+\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\right.\\
&\qquad\qquad\left.-\exp\left\{\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\Phi(m_0^*\alpha^*-\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})\right].
\end{aligned}
$$

Combining these two results, we obtain

$$
\gamma_{\mathbf{X}}(\mathbf{s}) = \frac{\exp\left\{-\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\Phi(m_0^*\alpha^*+\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})-\exp\left\{\frac{m_0^*}{\alpha^*}\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\Phi(m_0^*\alpha^*-\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s})}{\exp\left\{\frac{1}{2}\mathbf{s}^\top\boldsymbol{\omega}^{-1}\bar{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}^{-1}\mathbf{s}\right\}\Phi(m_0^*\alpha^*)}. \tag{26}
$$

From (26), a few interesting results can be derived. First of all, when $\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}\mathbf{s}=0$, $\gamma_{\mathbf{X}}(\mathbf{s})=0$. Since this is linear in $\mathbf{s}$, the zero-level contour of (26) is thus a straight line. Related to this, when $\boldsymbol{\alpha}=\mathbf{0}$, $m_0^*=0$ and consequently (26) will be zero $\forall\mathbf{s}\in\mathcal{S}$.

A general formula for skew-elliptical distributions (23), with skewing function $w(\mathbf{x})=\boldsymbol{\alpha}^\top\mathbf{x}$, is given by

$$
\gamma_{\mathbf{X}}(\mathbf{s}) = \frac{f_d(\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}+\mathbf{s}-\boldsymbol{\xi}))G(\boldsymbol{\alpha}^T\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}+\mathbf{s}-\boldsymbol{\xi}))-f_d(\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}+\mathbf{s}-\boldsymbol{\xi}))G(\boldsymbol{\alpha}^T\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}-\mathbf{s}-\boldsymbol{\xi}))}{f_d(\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}-\boldsymbol{\xi}))G(\boldsymbol{\alpha}^T\boldsymbol{\omega}^{-1}(\mathcal{M}_{\mathbf{X}}-\boldsymbol{\xi}))}.
$$

Explicit formulae for the mode of skew-elliptical distributions are only available in selective cases, e.g., for skew-normal and skew-t distributions. Even in those cases, the modes depend on an equation that needs numerical solving. Also finding an explicit maximizer of this equation is not analytically possible, hence no additional extra information can be given on $\Gamma(\mathbf{X})$ and $\boldsymbol{\gamma}^*(\mathbf{X})$. Fig. 7 contains the contourplot of a bivariate skew-normal distribution with parameters $\boldsymbol{\xi}=(1.25,-2.6)^\top$, $\boldsymbol{\Omega}=\begin{bmatrix}3.5 & -0.9\\ -0.9 & 5.8\end{bmatrix}$ and $\boldsymbol{\alpha}=(6,-6)^\top$ together with its reflected version. In Fig. 8 the measure (17) of this distribution is shown. Note that the contours of the asymmetry measure are always odd functions in $\mathbf{s}$ with respect to the origin (in the $\mathbf{s}$ plane) or the mode (in the data-plane). The proof of this is easily derived from (17). For this model $\mathcal{M}_{\gamma_{\mathbf{X}}}=(0.6653;-0.8554)^\top$, $\Gamma(\mathbf{X})=0.7837$ and $\boldsymbol{\gamma}^*(\mathbf{X})=(0.4812;-0.6186)^\top$.

4.4. Two piece asymmetric distributions

We first introduce a specific family of univariate two piece asymmetric distributions, called quantile based asymmetric distributions (QBA-distributions) in the literature. The density of a QBA-distributed random variable X is given by

$$
f_X(x;\boldsymbol{\eta}) = \frac{2\alpha(1-\alpha)}{\phi}\begin{cases} f\left(-(1-\alpha)\frac{x-\mu}{\phi};\boldsymbol{\kappa}\right) & \text{if } x\le\mu,\\ f\left(\alpha\frac{x-\mu}{\phi};\boldsymbol{\kappa}\right) & \text{if } x>\mu.\end{cases} \tag{27}
$$

In this f is the symmetric reference density, which is assumed to be unimodal and standardized, and possibly coming with some parameter(s), collected in the parameter vector $\boldsymbol{\kappa}$. The two piece asymmetric distribution in (27) depends on

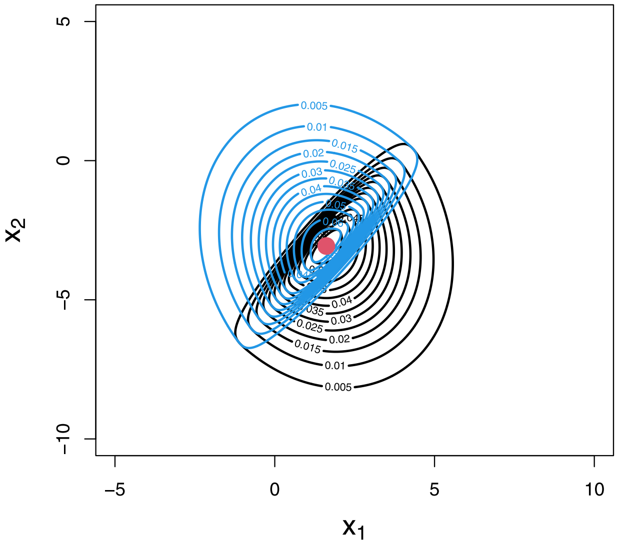


Fig. 7. Contour plots of the skew-normal distribution (black) with its mode (red) and reflected contours (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

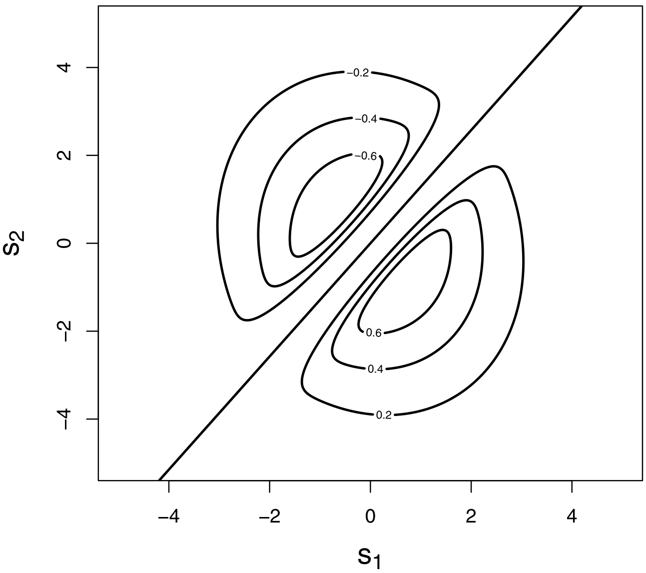


Fig. 8. Proposed asymmetry measure Eq. (17) applied to the skew-normal distribution.

the parameter vector $\boldsymbol{\eta} = (\alpha, \mu, \phi, \boldsymbol{\kappa}^\top)^\top$. Skewness is governed by the parameter $\alpha \in (0, 1)$, and $\mu \in \mathbb{R}$ and $\phi \in (0, \infty)$ are respectively a location and scale parameter. More information on the QBA-family, and statistical inference for it can be found in [20]. We make a distinction between $\alpha \leq 0.5$ (right skewness) and $\alpha > 0.5$ (left skewness). The mode of (27) is $\mathcal{M}_X = \mu$ (the location parameter), and it easily seen that

$$\gamma_X(s) = \begin{cases} 2\left[f\left((1-\alpha)\frac{s}{\phi}\right) - f\left(\alpha\frac{s}{\phi}\right)\right] & \text{if } s \leq 0, \\ 2\left[f\left(\alpha\frac{s}{\phi}\right) - f\left((1-\alpha)\frac{s}{\phi}\right)\right] & \text{if } s > 0. \end{cases}$$

The equation we need to solve in order to find $\mathcal{M}_{\gamma_X}$ is given by

$$\frac{f'\left(\frac{\alpha s}{\phi}\right)}{f'\left(\frac{(1-\alpha)s}{\phi}\right)} = \frac{1-\alpha}{\alpha}. \tag{28}$$

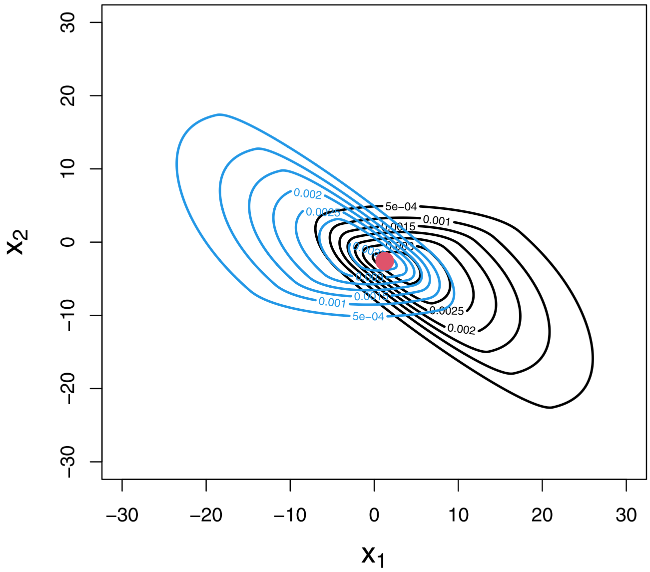


Fig. 9. Contourplots of the linear combination distribution (black) with its mode (red) and reflected contours (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

If $\alpha > 0.5$ (respectively $\alpha \leq 0.5$) then $(1-\alpha)/\alpha < 1$ (respectively more than or equal to 1), and hence in the former case $\mathcal{M}_{\gamma_X} < 0$ whereas in the latter case $\mathcal{M}_{\gamma_X} \geq 0$.

The solution to (28) for the QBA-normal, -Laplace, -logistic and -Student's t distribution can be found in the third column of Table S1 in the Supplementary Material. The last column of Table S1 provides information on $\Gamma(X)$ for these four univariate QBA-distributions.

For the multivariate setting we consider a linear combination of QBA-distributed random variables. Let $\mathbf{X} = \mathbf{A}^\top \mathbf{Z} + \boldsymbol{\mu}_a$ ($\mathbf{A} \in \mathbb{R}^{d\times d}$ and $\boldsymbol{\mu}_a \in \mathbb{R}^d$), in which the random vector $\mathbf{Z}$ consists of independent components Z_j, for $j = 1, \ldots, d$, with the density of Z_j denoted by f_{Z_j}. The random vector $\mathbf{X}$ then has density function

$$f_{\mathbf{X}}(\mathbf{x}; \mathbf{A}, \boldsymbol{\mu}_a, \boldsymbol{\eta}) = |\det(\mathbf{A})|^{-1} \prod_{j=1}^{d} f_{Z_j}((\mathbf{x} - \boldsymbol{\mu}_a)^T \left(\mathbf{A}^{-1}\right)_{.,j}; \boldsymbol{\eta}_j),$$

where $\left(\mathbf{A}^{-1}\right)_{.,j}$ denotes the jth column of the matrix $\mathbf{A}^{-1}$, and in which $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^\top, \ldots, \boldsymbol{\eta}_d^\top)^\top$. We assume that each of the Z_j follows a QBA-distribution given by (27). For identifiability reason we assume $\mu_j = 0$ and $\phi_j = 1$. It then immediately follows that the parameter $\boldsymbol{\mu}_a$ represents the location of the mode. From this, we have that $f_X(\mathcal{M}_{\mathbf{X}}) = |\det(\mathbf{A})|^{-1} \prod_{j=1}^{d} f_{Z_j}(0)$, where for notational simplicity we avoid to write the dependence on the parameters. In order to find $\gamma_{\mathbf{X}}(\mathbf{s})$ (and subsequently $\gamma^*(\mathbf{X})$), an expression for $f_{\mathbf{X}}(\boldsymbol{\mu}_a+\mathbf{s}; \mathbf{A}, \boldsymbol{\mu}_a, \boldsymbol{\eta}) - f_{\mathbf{X}}(\boldsymbol{\mu}_a-\mathbf{s}; \mathbf{A}, \boldsymbol{\mu}_a, \boldsymbol{\eta})$ is still required. One obtains

$$\gamma_{\mathbf{X}}(\mathbf{s}) = \frac{\prod_{j=1}^{d} f_{Z_j}(\mathbf{s}^\top \left(\mathbf{A}^{-1}\right)_{.,j}) - \prod_{j=1}^{d} f_{Z_j}(-\mathbf{s}^\top \left(\mathbf{A}^{-1}\right)_{.,j})}{\prod_{j=1}^{d} f_{Z_j}(0)},$$

and

$$\boldsymbol{\gamma}^*(\mathbf{X}) = \frac{\left(\prod_{j=1}^{d} f_{Z_j}(\mathcal{M}_{\gamma_{\mathbf{X}}}^\top \left(\mathbf{A}^{-1}\right)_{.,j}) - \prod_{j=1}^{d} f_{Z_j}(-\mathcal{M}_{\gamma_{\mathbf{X}}}^\top \left(\mathbf{A}^{-1}\right)_{.,j})\right) \mathcal{M}_{\gamma_{\mathbf{X}}}}{\left\|\mathcal{M}_{\gamma_{\mathbf{X}}}\right\| \prod_{j=1}^{d} f_{Z_j}(0)},$$

with $\mathcal{M}_{\gamma_{\mathbf{X}}} = \arg\max_{\mathbf{s}\in\mathcal{S}} \gamma_{\mathbf{X}}(\mathbf{s})$. No closed form expression exist for $\Gamma(\mathbf{X})$ and $\boldsymbol{\gamma}^*(\mathbf{X})$, hence numerical methods are required to arrive at the desired quantities. By the elegance and computational simplicity of the proposed asymmetry measure, this is easily achieved.

In Fig. 9 a linear combination model of QBA-distributions with parameters $\boldsymbol{\alpha} = (0.25, 0.75)^\top$, $\boldsymbol{\mu}_a = (1.25, -2.6)^\top$ and $A = \begin{bmatrix} 2 & 0.5 \\ -1.5 & 1.5 \end{bmatrix}$ is shown. The univariate components are (in order) a QBA-normal and QBA-logistic distribution. Fig. 10 represents the proposed asymmetry measure applied to this distribution. In this case, the zero-level contour of (17) no longer is a straight line, but is curved. Also the contour lines are more irregularly shaped compared to these in Fig. 8. Here we have $\mathcal{M}_{\gamma_{\mathbf{X}}} = (5.9580; 1.4557)^\top$, $\Gamma(\mathbf{X}) = 0.6754$, and $\boldsymbol{\gamma}^*(\mathbf{X}) = (0.6561; 0.1603)^\top$.

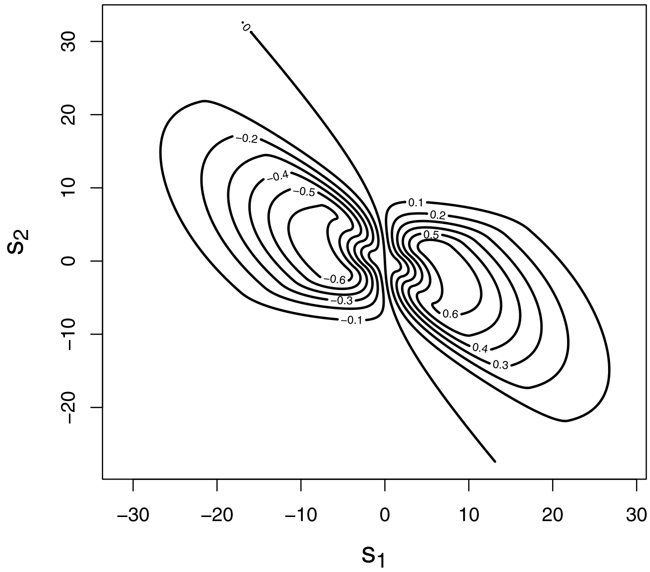


Fig. 10. Proposed asymmetry measure Eq. (17) applied to the linear combination distribution.

4.5. Transformation of scale distributions

In [1] another popular skewing mechanism is used to illustrate their asymmetry index. The distributions they consider are Transformation of Scale (ToS) distributions, which are described in, among others, [19,26,27,31]. ToS distributions introduce asymmetry by starting from a symmetric unimodal reference density g (with mode at zero) and rescaling its argument by a function $r(x)$. Usually the scaling function depends on a certain parameter λ to control the asymmetry introduced. The density of a ToS distributed r.v. is then given by $f(x;\lambda) = g(r(x;\lambda))$.

If $\frac{\partial}{\partial x} r(x;\lambda) > 0$ then also $f(\cdot)$ is unimodal with mode at zero. We assume this holds from hereon. For such ToS distributions, the asymmetry index $\gamma_X(s)$ equals

$$\gamma_X(s) = \frac{g(r(s;\lambda)) - g(r(-s;\lambda))}{g(r(0;\lambda))}.$$

Finding $\mathcal{M}_{\gamma_X}$, the mode of this function, is not straightforward, since it involves evaluating the derivative of both the functions $g(\cdot)$ and $r(\cdot;\lambda)$.

In [19] some other assumptions for ToS distributions are put in place. Denote by $r^{-1}(y;\lambda)$, the inverse function of $r(\cdot;\lambda)$. If $\frac{\partial}{\partial y} r^{-1}(y;\lambda) + \frac{\partial}{\partial y} r^{-1}(-y;\lambda) = 2$ and a function $H(y;\lambda)$ is taken such that $H(0,\lambda) = 0$ and $r^{-1}(y;\lambda) = y + H(y;\lambda)$, then mode invariance is obtained for any combination of $g(\cdot)$ and $r(\cdot;\lambda)$.

The specific example we use to illustrate our asymmetry index is a specific ToS distribution from [19], which is mode-invariant for any $f(\cdot)$ and $H(\cdot)$. Therein, $r^{-1}(y;\lambda)$ is chosen as

$$r^{-1}(y;\lambda) = y + a_\lambda \frac{\sqrt{1+\lambda^2 y^2} - 1}{\lambda}, \tag{29}$$

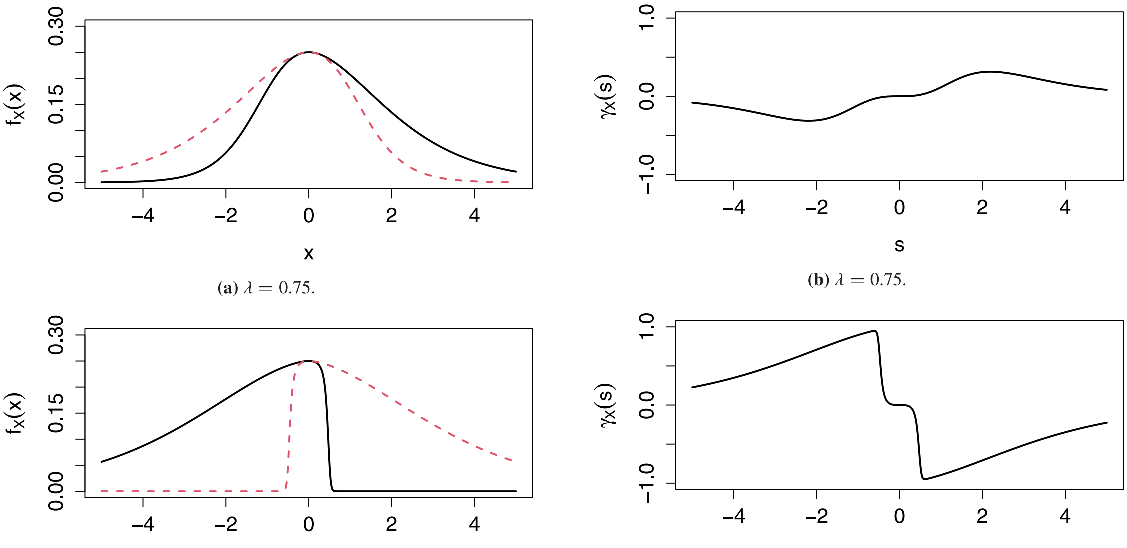
with $a_\lambda = 1 - exp(-\lambda^2)$. From (29) one is able to find an explicit expression for $r(x;\lambda)$, given by

$$r(x;\lambda) = \begin{cases} \frac{\lambda x + a_\lambda - a_\lambda\sqrt{(\lambda x + a_\lambda)^2 + 1 - a_\lambda^2}}{\lambda(1-a_\lambda^2)} & \text{if } \lambda \neq 0, \\ x & \text{if } \lambda = 0. \end{cases} \tag{30}$$

For $g(\cdot)$ we take the standard logistic distribution. In Fig. 11, the density and $\gamma_X(s)$-function are plotted for $\lambda = 0.75$ and $\lambda = -2$.

In analogy with the QBA-distributions, a way to extend these univariate ToS distributions to the multivariate setting is by taking affine combinations as is also done in [1]. For a random vector $\mathbf{Z} = (Z_1,\ldots,Z_d)^\top$, denote the density of the jth component by $f_j(z_j;\lambda_j) = g_j(r_j(z_j;\lambda_j))$, for $j \in \{1,\ldots,d\}$. The density of $\mathbf{X} = \mathbf{A}^\top\mathbf{Z} + \boldsymbol{\mu}_a$ is then given by

$$f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\lambda}) = |\det(\mathbf{A})|^{-1} \prod_{j=1}^{d} g_j\left(r_j\left((\mathbf{x}-\boldsymbol{\mu}_a)^\top(\mathbf{A}^{-1})_{.,j};\lambda_j\right)\right)$$



(a) $\lambda = 0.75$. (b) $\lambda = 0.75$.

(c) $\lambda = -2$. (d) $\lambda = -2$.

Fig. 11. Left column: Density $f_X(x;\lambda) = g(r(x;\lambda))$ of a mode-invariant ToS-distribution with $g(\cdot)$ a standard logistic density and $r(x;\lambda)$ as in (30), for $\lambda = 0.75$ (top panel left), and $\lambda = -2$ (bottom panel left). The reflected density, i.e., $f_{2\mathcal{M}_X - X}(x;\lambda)$ is plotted as a dashed line. Right column: corresponding asymmetry function $\gamma_X(s)$.

and subsequently

$$\gamma_{\mathbf{X}}(\mathbf{s}) = \frac{\prod_{j=1}^{d} g_j\left(r_j(\mathbf{s}^\top\left(\mathbf{A}^{-1}\right)_{\cdot,j};\lambda_j)\right) - \prod_{j=1}^{d} g_j\left(r_j(-\mathbf{s}^\top\left(\mathbf{A}^{-1}\right)_{\cdot,j};\lambda_j)\right)}{\prod_{j=1}^{d} g_j\left(r_j(0;\lambda_j)\right)}.$$

From the latter expression, the summarizing indices can be derived.

As a bivariate example, we apply the same affine combination as in Section 4.4 to the context of ToS distributions, by taking g_1 a standard normal and g_2 a logistic distribution and both r_1 and r_2 as in (30) with $\lambda_1 = 0.75$ and $\lambda_2 = -2$. For this example Figs. 12 and 13 are obtained.

5. Estimation of the asymmetry measure

Given an i.i.d. sample $\widetilde{\mathbf{X}}_n = (\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(n)})^\top$ from $\mathbf{X}$, we need to estimate the asymmetry measure (14) or (17). When an estimator of the density function of $\mathbf{X}$ is available, its mode can be determined so the problem of estimating the asymmetry function can be converted into a problem of estimating a multivariate density and its mode. We discuss parametric and nonparametric estimation methods.

5.1. Parametric estimation

Assume that the random vector $\mathbf{X}$ has a unimodal density function $f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta})$, depending on a parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^q$, with the parameter space $\boldsymbol{\Theta}$ a compact subset of $\mathbb{R}^q$. The mode of the density, denoted by $\mathcal{M}_{\mathbf{X}}(\boldsymbol{\theta})$, depends on the unknown parameter vector. Denote the true parameter vector by $\boldsymbol{\theta}_0$. Based on the i.i.d. sample $\widetilde{\mathbf{X}}_n$ we obtain an estimator $\widehat{\boldsymbol{\theta}}_n$ for the parameter vector $\boldsymbol{\theta}$. The true density $f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta}_0)$ is then approximated by the fitted density $f_{\mathbf{X}}(\mathbf{x};\widehat{\boldsymbol{\theta}}_n)$, and an estimator for the mode $\mathcal{M}_{\mathbf{X}}(\boldsymbol{\theta}_0)$ is then the mode of the estimated density. We introduce the following notations:

$$\widehat{\mathcal{M}}_{\mathbf{X},n} = \arg\max_{\mathbf{x}\in\mathbb{R}^d} f_{\mathbf{X}}(\mathbf{x};\widehat{\boldsymbol{\theta}}_n), \qquad \widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}) = \frac{f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}+\mathbf{s};\widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}-\mathbf{s};\widehat{\boldsymbol{\theta}}_n)}{f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n};\widehat{\boldsymbol{\theta}}_n)},$$
$$\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n} = \arg\max_{\mathbf{s}\in\mathcal{S}} \widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}), \qquad \widehat{\Gamma}_n(\mathbf{X}) = \max_{\mathbf{s}\in\mathcal{S}} \widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}), \qquad \widehat{\boldsymbol{\gamma}}_n^*(\mathbf{X}) = \widehat{\Gamma}_n(\mathbf{X})\frac{\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n}}{\left\|\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n}\right\|}, \tag{31}$$

where $\mathbf{s}$ takes values in a certain compact set $\mathcal{S} \subset \mathbb{R}^d$ containing $\mathbf{0}$.

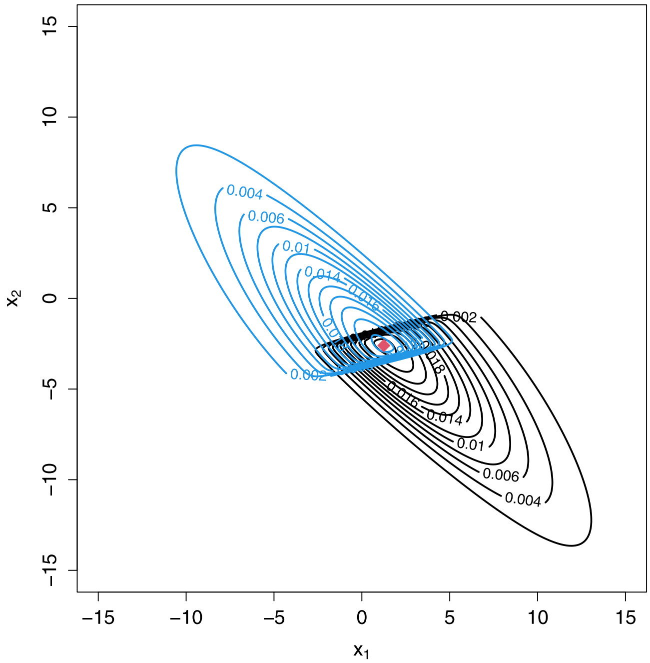


Fig. 12. Contourplots of the linear combination of ToS distributions (black) with its mode (red) and reflected contours (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

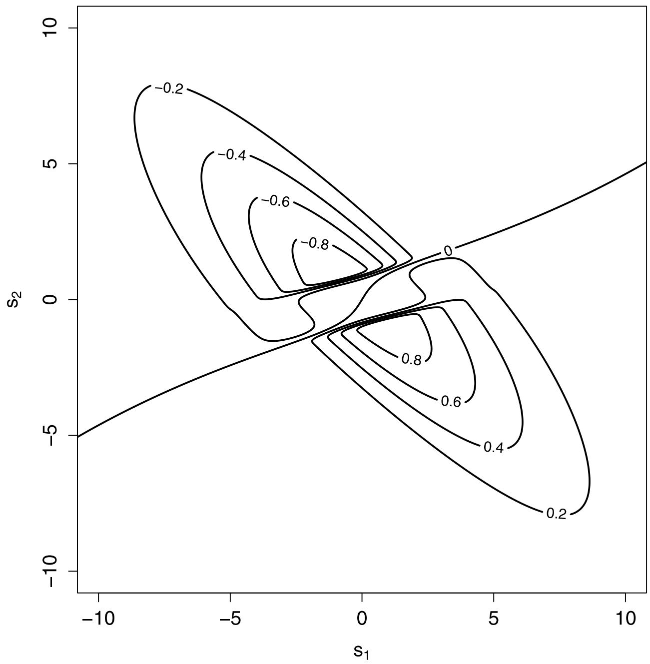


Fig. 13. Proposed asymmetry measure Eq. (17) applied to the linear combination of ToS distributions.

To show consistency of $\widehat{\gamma}_n^*(\mathbf{X})$, we need the concept of uniform consistency in probability. A random quantity $g_n(x)$, considered as a function of x, is uniformly consistent in probability to $g(x)$ if

$$\text{for all } \epsilon > 0 \quad \lim_{n\to\infty} \Pr\left(\sup_{x\in\mathbb{R}} |g_n(x) - g(x)| > \epsilon\right) = 0.$$

We denote this uniform in probability consistency as $g_n(x) \overset{\text{unif. P}}{\to} g(x)$, as $n \to \infty$.

The following theorem states that under some mild conditions on the density and the parameter vector estimator, the estimator $\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s})$ of $\gamma_{\mathbf{X}}(\mathbf{s})$ is consistent in probability, uniformly; and that the estimators $\widehat{\Gamma}_n(\mathbf{X})$, $\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n}$ and $\widehat{\boldsymbol{\gamma}}_n^*(\mathbf{X})$ are consistent estimators of respectively $\Gamma(\mathbf{X})$, $\mathcal{M}_{\gamma_{\mathbf{X}}}$ and $\boldsymbol{\gamma}^*(\mathbf{X})$. The proof of the theorem is provided in Appendix A.

Theorem 2. *Let $\widetilde{\mathbf{X}}_n = (\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)})^\top$ be an i.i.d. sample from a random vector $\mathbf{X} \in \mathbb{R}^d$ with continuous unimodal density function $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}_0)$. Also let $\widehat{\boldsymbol{\theta}}_n$ be a consistent estimator of $\boldsymbol{\theta}_0$, i.e., $\widehat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0$, as $n \to \infty$. If $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ is uniformly continuous in both $\mathbf{x}$ and $\boldsymbol{\theta}$, as well as uniformly bounded $\forall \mathbf{x}$, and $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, and $\widehat{\mathcal{M}}_{\mathbf{X},n}(\boldsymbol{\theta})$ is uniformly continuous in $\boldsymbol{\theta}$, then*

(i) *$\forall \varepsilon > 0$ it holds that $\lim_{n\to\infty} \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}) - \gamma_{\mathbf{X}}(\mathbf{s})\right| > \varepsilon\right) = 0$, i.e., $\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}) \overset{unif.\ P}{\to} \gamma_{\mathbf{X}}(\mathbf{s})$, as $n \to \infty$;*
(ii) *$\widehat{\Gamma}_n(\mathbf{X}) \xrightarrow{P} \Gamma(\mathbf{X})$, as $n \to \infty$;*
(iii) *$\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n} \xrightarrow{P} \mathcal{M}_{\gamma_{\mathbf{X}}}$, as $n \to \infty$;*
(iv) *$\widehat{\boldsymbol{\gamma}}_n^*(\mathbf{X}) \xrightarrow{P} \boldsymbol{\gamma}^*(\mathbf{X})$, as $n \to \infty$.*

5.2. Nonparametric estimation

There are various kinds of nonparametric density estimators available. We focus here on kernel density estimation (see among others [14]). Denote with $K(\mathbf{x})$ a d-variate density function called the kernel. Following [41], we will require that the kernel satisfies the following 3 conditions

(K1) $\int_{\mathbb{R}^d} K(\mathbf{u})d\mathbf{u} = 1$;
(K2) $\int_{\mathbb{R}^d} \mathbf{u}K(\mathbf{u})d\mathbf{x} = 0$ element-wise;
(K3) $\int_{\mathbb{R}^d} \mathbf{u}\mathbf{u}^\top K(\mathbf{u})d\mathbf{u} = m_2(K)\mathbb{I}_d < \infty$, with $m_2(K) = \int_{\mathbb{R}} u_j^2 K(\mathbf{u})du_j$ (for all $j \in \{\ldots, d\}$).

Further consider $\mathbf{H}$ a $d \times d$ symmetric positive definite matrix, and denote its (positive) determinant by $|\mathbf{H}|$ (shorthand notation for the notation $\det(\mathbf{H})$ used before). The matrix $\mathbf{H}$ is called the bandwidth matrix. The multivariate kernel density estimator for $f_{\mathbf{X}}(\mathbf{x})$ is then given by

$$\widehat{f}_{\mathbf{X}}(\mathbf{x}; \mathbf{H}) = \frac{1}{n\,|\mathbf{H}|^{1/2}} \sum_{i=1}^{n} K(\mathbf{H}^{-1/2}(\mathbf{x} - \mathbf{X}^{(i)})) = \frac{1}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{X}^{(i)}), \tag{32}$$

where we denoted the rescaled kernel $K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K\left(\mathbf{H}^{-1/2}\mathbf{x}\right)$. Expression (32) involves two important ingredients. The first is the kernel function. Examples of univariate kernel function include the normal kernel, Epanechnikov kernel and uniform kernel. In a multivariate setting, one often uses the standard multivariate normal kernel, i.e.,

$$K(\mathbf{x}) = \phi_d(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}\mathbf{x}^\top\mathbf{x}}, \quad \text{with rescaled version} \quad K_{\mathbf{H}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\,|\mathbf{H}|^{1/2}} e^{-\frac{1}{2}\mathbf{x}^\top\mathbf{H}^{-1}\mathbf{x}}.$$

We will show a consistency result similar to that in Section 5.1 for the univariate case. Extending this to the multivariate case is straightforward, and therefore not elaborated on in detail. For estimation of $\gamma_X(s)$ both the density estimate and the estimated mode are used. Given $X_1, \ldots, X_n$ is an i.i.d. sample from X, the univariate kernel density estimator is

$$\widehat{f}_n(x; h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \tag{33}$$

where $h > 0$ in a bandwidth parameter. The mode estimation is performed on the estimated density. Some notations are then:

$$\widehat{\mathcal{M}}_{X,n} = \arg\max_{x\in\mathbb{R}} \widehat{f}_n(x; h), \qquad \widehat{\gamma}_{X,n}(s) = \frac{\widehat{f}_n(\widehat{\mathcal{M}}_{X,n} + s; h) - \widehat{f}_n(\widehat{\mathcal{M}}_{X,n} - s; h)}{\widehat{f}_n(\widehat{\mathcal{M}}_{X,n}; h)},$$

and subsequently, based on these, the estimators as in (31).

Proposition 4 states the conditions under which a kernel density estimator, in the univariate case ($d = 1$), converges in probability (and uniformly), to the true density, and under which the mode estimator $\widehat{\mathcal{M}}_{X,n}$ consistently estimates the mode $\mathcal{M}_X$. For a proof see Theorem 3 A in [36].

Proposition 4. *Suppose $X_1, \ldots, X_n$ is an i.i.d. sample from X with density $f_X(x)$ (on $\mathbb{R}$). Let $\widehat{f}_n(x; h)$ be the kernel density estimator in (33), with $h = h_n$ a bandwidth sequence, and the kernel $K(u)$ is uniformly continuous. Assume also that the following conditions hold*

(C1) *$f_X(x)$ is uniform continuous and uniformly bounded on $\mathbb{R}$;*
(C2) *$f_X(x)$ has a unique mode $\mathcal{M}_X$;*
(C3) *$h \to 0$ and $nh^2 \to \infty$, as $n \to \infty$.*

Then we have that

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \Pr\left(\sup_{\mathbf{x}\in\mathbb{R}} \left|\widehat{f}_n(x; h) - f_X(x)\right| > \varepsilon\right) = 0, \qquad \widehat{\mathcal{M}}_{X,n} \xrightarrow{P} \mathcal{M}_X, \quad \text{as } n \to \infty. \tag{34}$$

The above results allow to establish uniform in probability results for the estimator $\widehat{\gamma}_{X,n}(s)$ of the proposed asymmetry measure $\gamma_X(s)$. This is established in Theorem 3, the proof of which is deferred to Appendix A.

Theorem 3. *Under (K1)-(K3) for a uniformly bounded kernel function, (C1)-(C3) and with $s \in \mathcal{S} \subset \mathbb{R}$ compact, it holds that, as $n \to \infty$,*

$$f_n(\widehat{\mathcal{M}}_{X,n} + s; h) \xrightarrow{P} f_X(\mathcal{M}_X + s) \text{ uniformly in } s, \tag{35}$$

and

$$\widehat{\gamma}_{X,n}(s) \xrightarrow{P} \gamma_X(s) \text{ uniformly in } s.$$

From Theorem 3 it immediately follows that $\widehat{\Gamma}_n(X) \xrightarrow{P} \Gamma(X)$, as $n \to \infty$, since

$$\left|\widehat{\Gamma}_n(X) - \Gamma(X)\right| = \left|\sup_{s\in\mathcal{S}} \widehat{\gamma}_{X,n}(s) - \sup_{s\in\mathcal{S}} \gamma_X(s)\right| \le \sup_{s\in\mathcal{S}} \left|\widehat{\gamma}_{X,n}(s) - \gamma_X(s)\right| \xrightarrow{P} 0, \text{ as } n \to \infty.$$

Combining this with the fact that $\gamma_X(s)$ is uniformly continuous in s as this is only appearing in the numerator of its definition and $f_X(x)$ is uniformly continuous, a similar proof as that of Theorem 3 A in [36] provides the result that

$$\widehat{\mathcal{M}}_{\gamma_X,n} \xrightarrow{P} \mathcal{M}_{\gamma_X}, \quad \text{as } n \to \infty.$$

Hence we also have

$$\widehat{\gamma}_n^*(X) \xrightarrow{P} \gamma^*(X), \quad \text{as } n \to \infty.$$

Multivariate extensions of the results obtained in [36] (and recalled in Proposition 4), are given in, among others, [13, 30,40,47], and [14]. For a matrix $\mathbf{A}$ denote by Vec($\mathbf{A}$) the column vector obtained by stacking all columns of $\mathbf{A}$ into one long column vector. The conditions necessary for Proposition 4 become

(MC1) $f_{\mathbf{X}}(\mathbf{x})$ is uniform continuous on $\mathbb{R}^d$, and uniformly bounded;
(MC2) $f_{\mathbf{X}}(\mathbf{x})$ has a unique mode $\mathcal{M}_{\mathbf{X}}$, i.e., $\forall \mathbf{s} \in \mathbb{R}\backslash\{0\} : f_X(\mathcal{M}_{\mathbf{X}}) > f_X(\mathcal{M}_{\mathbf{X}} + \mathbf{s})$;
(MC3) The kernel function $\mathbf{K}$ is bounded, Hölder continuous and $\lim_{\|\mathbf{x}\|\to\infty} \|\mathbf{x}\|\, \mathbf{K}(\mathbf{x}) = 0$;
(MC4) $\text{Vec}(\mathbf{H}) \to \mathbf{0}$, and $n\,|\det \mathbf{H}| \to \infty$, and $n \to \infty$.

6. Comparison to classical skewness in the univariate case

We want to investigate how our measure of skewness (16) behaves in comparison to (3), the classical measure of skewness. For this, we consider three distributions:

- the Gamma distribution
$$\Gamma(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad \alpha, \beta, x > 0,$$
- the beta distribution
$$B(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad \alpha, \beta > 0; x \in [0, 1],$$
- the univariate quantile-based asymmetric normal distribution (see (27), with f the standard normal density), for $\mu \in \mathbb{R}$, $\alpha \in (0, 1)$, $\phi > 0$,
$$\text{AND}(x; \alpha, \mu, \phi) = \alpha(1-\alpha)\sqrt{\frac{2}{\pi\phi^2}} \begin{cases} e^{-\frac{\alpha^2}{2}(\frac{x-\mu}{\phi})^2} & \text{for } x > \mu, \\ e^{-\frac{(1-\alpha)^2}{2}(\frac{x-\mu}{\phi})^2} & \text{for } x \le \mu. \end{cases}$$

In Table S2 of the Supplementary Material we list for these three distributions the following quantities: the mode $\mathcal{M}_X$, the classical measure of skewness SK(X), and the discussed quantities $\mathcal{M}_{\gamma_X}$ and $\gamma^*(X)$ (the new summarizing measure of asymmetry).

As can be seen from Table S2, for the $\Gamma(\alpha, \beta)$ distribution, classical skewness only depends on the parameter α. For the $B(\alpha, \beta)$ distribution this is both α and β and for the AND(α, μ, ϕ) this is again only α. To compare the asymmetry measure SK(X) in (3) with the measure $\gamma^*(X)$ defined in (16), visual tools are used. We refer to Figures S1–S5 in the Supplementary Material. Both measures, SK(X) and $\gamma^*(X)$, as a function of the parameter α are given in Figures S1 to S3 in the Supplementary Material. Our proposed measure behaves the same as the classical skewness in all three situations with the sole difference that for the Gamma- and beta-distribution it has a flat area corresponding to parameter values where the density is unimodal with mode in 0. Hence, by definition, $\gamma^*(X)$ equals 1 as the distribution can be regarded as being of the half-distribution type. These two distributions also highlight the positive aspect that $\gamma^*(X)$ is bounded, whereas classical skewness approaches infinity as $\alpha \to 0$. For the AND(α, μ, ϕ) distribution, both SK(X) and $\gamma^*(X)$ are very similar, with only minor differences in value. The trained eye might also notice some noise in the curves of $\gamma^*(X)$. This happens mostly for more extreme levels of skewness and is completely due to numerical instabilities of the optimization algorithm used in finding $\mathcal{M}_{\gamma_X}$ for the distributions in which no closed form formula is available.

As no explicit expression of (16) is available for the Gamma-distribution, there might be dependence on β. In order to check this, a range of values for β is considered in Fig. S4 in the Supplementary Material. As can be seen, all curves coincide so (16) only depends on α as is also the case for classical skewness. There is again some numerical instability going on for higher levels of skewness. Fig. S2, on the beta-distribution, shows only the behavior in α for fixed β. This behavior is completely symmetric in both parameters, but for completeness, Fig. S5 in the Supplementary Material is added. This figure shows the contour lines of both the classical skewness and our proposed asymmetry measure for both α and β larger than one. Both resemble each other except the aforementioned boundedness of our proposed measure.

7. Discussion

In this paper we proposed a novel measure of asymmetry, which in contrast to most common measures, provides a function of asymmetry. From this function, a global scalar measure can be derived. A multivariate extension is also introduced with the main advantage that any deviations from symmetry with respect to the mode can be detected. By doing this, a more detailed picture of asymmetry is provided and regions of interest can be better investigated. The latter multivariate asymmetry function can also be summarized in a vectorial measure which gives an indication of the magnitude of the asymmetry as well as its direction.

We also presented methods for estimating the asymmetry function and show consistency of both a parametric and nonparametric estimator under some mild conditions. The downside of the asymmetry measure is that it is computationally heavy to obtain over the entire domain of the function, even in moderately high dimensions. This is entirely due to it being a function over the domain, which requires evaluation over a grid to obtain the full picture. This downside is however inherent to any kind of asymmetry measure that looks in detail into asymmetry and goes beyond moment expressions.

CRediT authorship contribution statement

Jonas Baillien: Conceptualization, Investigation, Methodology, Software, Writing – original draft, Reviewing. **Irène Gijbels:** Conceptualization, Methodology, Reviewing and editing, Supervision, Funding acquisition. **Anneleen Verhasselt:** Conceptualization, Methodology, Reviewing and editing, Supervision, Funding acquisition.

Acknowledgments

The authors are grateful to an Associate Editor and two reviewers for their valuable comments which led to an improvement of the manuscript. The first and second author gratefully acknowledge support from the Research Fund KU Leuven [C16/20/002 project]. The third author was supported by Special Research Fund (Bijzonder Onderzoeksfonds) of Hasselt University [BOF14NI06].

Appendix A. Proofs

Proof of Theorem 2. First, we will show that

$$f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) \xrightarrow{\mathrm{P}} f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\theta}_0) \text{ uniformly over } \mathbf{s}, \text{ as } n \to \infty. \tag{A.1}$$

In showing this, the uniform continuous mapping theorem (UCMT, [10,28]) is used. By uniform continuity in $\mathbf{x}$ of $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$, combined with the UCMT and

$$\lim_{n\to\infty} \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|\widehat{\mathcal{M}}_{\mathbf{X},n} + s - (\mathcal{M}_{\mathbf{X}} + s)\right| > \varepsilon\right) = 0,$$

statement (A.1) follows.

Proof of statement (i). As we are dealing with bounded density functions, there exists some $U < \infty$ such that $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}) < U$ $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$. It is also safe to assume that $\exists 0 < L < U < \infty : L < f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)$. Consider now $\mathcal{B}_\delta = \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \delta\}$. By consistency of $\widehat{\boldsymbol{\theta}}_n$ to $\boldsymbol{\theta}_0$, uniform continuity of $\widehat{\mathcal{M}}_{\mathbf{X},n}(\boldsymbol{\theta})$ in $\boldsymbol{\theta}$, and uniform continuity of $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ in both $\mathbf{x}$ and $\boldsymbol{\theta}$ we also have that for any $\widehat{\boldsymbol{\theta}}_n \in \mathcal{B}_\delta$ it holds that, if we take δ appropriately small, $\left|f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)\right| < \frac{L}{2}$. In this way, we have when $\widehat{\boldsymbol{\theta}}_n \in \mathcal{B}_\delta$ and for $\varepsilon > 0$

$$\Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s}) - \gamma_{\mathbf{X}}(\mathbf{s})\right| > \varepsilon\right) \tag{A.2}$$

$$= \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|\frac{f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n)}{f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n)} - \frac{f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\theta}_0) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s}; \boldsymbol{\theta}_0)}{f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)}\right| > \varepsilon\right)$$

$$\leq \Pr\Bigg(\sup_{\mathbf{s}\in\mathcal{S}} \Big| f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)(f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n))$$

$$- f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n)(f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\theta}_0) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s}; \boldsymbol{\theta}_0))\Big| > \frac{\varepsilon L^2}{2}\Bigg)$$

$$= \Pr\Bigg(\sup_{\mathbf{s}\in\mathcal{S}} \Big| f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n)$$

$$+ f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\theta}_0)\Big| > \frac{\varepsilon L^2}{4}\Bigg)$$

$$+ \Pr\Bigg(\sup_{\mathbf{s}\in\mathcal{S}} \Big| f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n)$$

$$+ f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s}; \boldsymbol{\theta}_0)\Big| > \frac{\varepsilon L^2}{4}\Bigg)$$

$$\leq \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)\right| > \frac{\varepsilon L^2}{8U}\right)$$

$$+ \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} + \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} + \mathbf{s}; \boldsymbol{\theta}_0)\right| > \frac{\varepsilon L^2}{8U}\right)$$

$$+ \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}}; \boldsymbol{\theta}_0)\right| > \frac{\varepsilon L^2}{8U}\right)$$

$$+ \Pr\left(\sup_{\mathbf{s}\in\mathcal{S}} \left|f_{\mathbf{X}}(\widehat{\mathcal{M}}_{\mathbf{X},n} - \mathbf{s}; \widehat{\boldsymbol{\theta}}_n) - f_{\mathbf{X}}(\mathcal{M}_{\mathbf{X}} - \mathbf{s}; \boldsymbol{\theta}_0)\right| > \frac{\varepsilon L^2}{8U}\right). \tag{A.3}$$

Denote the event in (A.2) by E, and note that

$$P(E) = P\left(E \cap \left\{\widehat{\boldsymbol{\theta}}_n \in \mathcal{B}_\delta\right\}\right) + P\left(E \cap \left\{\widehat{\boldsymbol{\theta}}_n \in \mathcal{B}_\delta^C\right\}\right).$$

The first term herein tends to zero by (A.3), and the last term is bounded by $P\left(\widehat{\boldsymbol{\theta}}_n \in \mathcal{B}_\delta^C\right)$ which tends to zero as n tends to infinity. This concludes the proof of statement (i).

Proofs of statements (ii)–(iv). In showing consistency of $\widehat{\mathcal{M}}_{\gamma_{\mathbf{X}},n}$ and $\widehat{\Gamma}_n(\mathbf{X})$, we rely on Theorem 4.1.1 of [2] that states conditions under which the maximum and, in case the objective function is uniquely maximized, also the maximizer is consistent. These conditions are

(M1) $\mathbf{s} \in \mathcal{S}$ compact;
(M2) $\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s})$ is a continuous, measurable function of $\mathbf{s}$ for all $\mathbf{s} \in \mathcal{S}$;
(M3) $\widehat{\gamma}_{\mathbf{X},n}(\mathbf{s})$ is a uniformly consistent (in probability) estimator of $\gamma_{\mathbf{X}}(\mathbf{s})$.

The former two conditions are (easily) met. The third condition is (i) of this theorem. Applying Theorem 4.1.1 of [2] concludes parts (ii) and (iii). Part (iv) then automatically follows from (ii) and (iii), completing the proof. □

Proof of Theorem 3. We have

$$\sup_{s\in\mathcal{S}} \left|f_n(\widehat{\mathcal{M}}_{X,n} + s; h) - f_X(\mathcal{M}_X + s)\right| \leq \sup_{s\in\mathcal{S}} \left|f_n(\widehat{\mathcal{M}}_{X,n} + s; h) - f_n(\mathcal{M}_X + s; h)\right| + \sup_{s\in\mathcal{S}} \left|f_n(\mathcal{M}_X + s; h) - f_X(\mathcal{M}_X + s)\right|.$$

By Proposition 4, for all $\varepsilon > 0$ the second term is smaller than $\frac{\varepsilon}{2}$ when $n \to \infty$. For the first term, we use (34) together with Assumption (C1) to fall back on the UCMT applied to $f_n(.; h)$ to obtain that $\sup_{s\in\mathcal{S}} \left|f_n(\widehat{\mathcal{M}}_{X,n} + s; h) - f_n(\mathcal{M}_X + s; h)\right|$ is strictly smaller than $\frac{\varepsilon}{2}$. We have

$$\Pr\left(\sup_{s\in\mathcal{S}} \left|f_n(\widehat{\mathcal{M}}_{X,n} + s; h) - f_X(\mathcal{M}_X + s)\right| > \varepsilon\right) \xrightarrow{\mathrm{P}} 0 \qquad n \to \infty.$$

The rest of the proof is similar to that of Theorem 2. We can assume that $f_X(x)$ and $f_n(x;h)$ ($n \in \mathbb{N}$) are bounded above by a constant U. Also, $\exists L > 0 : f_X(\mathcal{M}_X) > L$. For n large enough, and when $\left|f_n(\widehat{\mathcal{M}}_{X,n}) - f_X(\mathcal{M}_X)\right| < \frac{L}{2}$, we get

$$\begin{aligned}
\sup_{s\in\mathcal{S}}\left|\widehat{\gamma}_{X,n}(s)-\gamma_X(s)\right| &= \sup_{s\in\mathcal{S}}\left|\frac{f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_n(\widehat{\mathcal{M}}_{X,n}-s;h)}{f_n(\widehat{\mathcal{M}}_{X,n};h)}-\frac{f_X(\mathcal{M}_X+s)-f_X(\mathcal{M}_X-s)}{f_X(\mathcal{M}_X)}\right|\\
&\le \frac{2}{L^2}\sup_{s\in\mathcal{S}}\left|f_X(\mathcal{M}_X)(f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_n(\widehat{\mathcal{M}}_{X,n}-s;h))-f_n(\widehat{\mathcal{M}}_{X,n};h)(f_X(\mathcal{M}_X+s)-f_X(\mathcal{M}_X-s))\right|\\
&= \frac{2}{L^2}\sup_{s\in\mathcal{S}}\Big|f_X(\mathcal{M}_X)f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_n(\widehat{\mathcal{M}}_{X,n};h)f_n(\widehat{\mathcal{M}}_{X,n}+s;h)\\
&\qquad +f_n(\widehat{\mathcal{M}}_{X,n};h)f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_n(\widehat{\mathcal{M}}_{X,n};h)f_X(\mathcal{M}_X+s)\Big|\\
&\quad +\frac{2}{L^2}\sup_{s\in\mathcal{S}}\Big|f_X(\mathcal{M}_X)f_n(\widehat{\mathcal{M}}_{X,n}-s;h)-f_n(\widehat{\mathcal{M}}_{X,n};h)f_n(\widehat{\mathcal{M}}_{X,n}-s;h)\\
&\qquad +f_n(\widehat{\mathcal{M}}_{X,n};h)f_n(\widehat{\mathcal{M}}_{X,n}-s;h)-f_n(\widehat{\mathcal{M}}_{X,n};h)f_X(\mathcal{M}_X-s)\Big|\\
&\le \frac{2}{L^2}\Big[\left|f_n(\widehat{\mathcal{M}}_{X,n};h)-f_X(\mathcal{M}_X)\right|\sup_{s\in\mathcal{S}} f_n(\widehat{\mathcal{M}}_{X,n}+s;h)+f_n(\widehat{\mathcal{M}}_{X,n};h)\sup_{s\in\mathcal{S}}\left|f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_X(\mathcal{M}_X+s)\right|\\
&\qquad +\left|f_n(\widehat{\mathcal{M}}_{X,n};h)-f_X(\mathcal{M}_X)\right|\sup_{s\in\mathcal{S}} f_n(\widehat{\mathcal{M}}_{X,n}-s;h)+f_n(\widehat{\mathcal{M}}_{X,n};h)\sup_{s\in\mathcal{S}}\left|f_n(\widehat{\mathcal{M}}_{X,n}-s;h)-f_X(\mathcal{M}_X-s)\right|\Big]\\
&\le \frac{2U}{L^2}\Big[2\left|f_n(\widehat{\mathcal{M}}_{X,n};h)-f_X(\mathcal{M}_X)\right|+\sup_{s\in\mathcal{S}}\left|f_n(\widehat{\mathcal{M}}_{X,n}+s;h)-f_X(\mathcal{M}_X+s)\right|\\
&\qquad +\sup_{s\in\mathcal{S}}\left|f_n(\widehat{\mathcal{M}}_{X,n}-s;h)-f_X(\mathcal{M}_X-s)\right|\Big] \le \frac{8\varepsilon U}{L^2},
\end{aligned}$$

$\forall\varepsilon > 0$ when $n \to \infty$. The final step holds with probability 1 by (34) and (35) combined with the UCMT. Hence, as $n \to \infty$,

$$\widehat{\gamma}_{X,n}(s) \xrightarrow{\mathrm{P}} \gamma_X(s) \quad \text{uniform in } s. \quad \square$$

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2022.105118.

References

[1] T. Abe, H. Fujisawa, Multivariate skew distributions with mode-invariance through the transformation of scale, Jpn. J. Stat. Data Sci. 2 (2019) 529–544.
[2] T. Amemiya, Advanced Econometrics, Blackwell, Oxford, 1985.
[3] C.N.P.G. Arachchige, L.A. Prendergast, Mean skewness measures, 2019, arXiv:1912.06996v2.
[4] B.C. Arnold, R.A. Groeneveld, Measuring skewness with respect to the mode, Amer. Statist. 49 (1) (1995) 34–38.
[5] J. Averous, M. Meste, Skewness for multivariate distributions: Two approaches, Ann. Statist. 25 (5) (1997) 1984–1997.
[6] A. Azzalini, A class of distributions which includes the normal ones, Scand. J. Stat. 12 (2) (1985) 171–178.
[7] A. Azzalini, The Skew-Normal and Related Families, in: Institute of Mathematical Statistics Monographs, Cambridge University Press, Cambridge, 2014.
[8] A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew-normal distribution, J. R. Stat. Soc. Ser. B Stat. Methodol. 61 (3) (1999) 579–602.
[9] N. Balakrishnan, B. Scarpa, Multivariate measures of skewness for the skew-normal distribution, J. Multivariate Anal. 104 (1) (2012) 73–87.
[10] V. Bengs, H. Holzmann, Uniform approximation in classical weak convergence theory, 2019, arXiv:1903.09864v1.
[11] G.N. Boshnakov, Some measures for asymmetry of distributions, Statist. Probab. Lett. 77 (11) (2007) 1111–1116.
[12] A. Bowley, Elements of Statistics, P. S. King & Son, London, 1920.
[13] T. Cacoullos, Estimation of a Multivariate Density, Technical Report, University of Minnesota, 1964.
[14] J.E. Chacón, T. Duong, Multivariate Kernel Smoothing and Its Applications, CRC Press, New York, 2018.
[15] F. Critchley, M.C. Jones, Asymmetry and gradient asymmetry functions: Density-based skewness and Kurtosis, Scand. J. Stat. 35 (3) (2008) 415–437.
[16] F.N. David, N.L. Johnson, Some tests of significance with ordered variables, J. R. Stat. Soc. Ser. B Stat. Methodol. 18 (1) (1956) 1–31.
[17] A. Eberl, B. Klar, A note on a measure of asymmetry, Statist. Papers 62 (2021) 1483–1497.
[18] A. Eberl, B. Klar, Expectile-based measures of skewness, Scand. J. Stat. 49 (2022) 373–399.
[19] H. Fujisawa, T. Abe, A family of skew distributions with mode-invariance through transformation of scale, Stat. Methodol. 25 (2015) 89–98.
[20] I. Gijbels, R. Karim, A. Verhasselt, On quantile-based asymmetric family of distributions: Properties and inference, Internat. Statist. Rev. 87 (3) (2019) 471–504.
[21] R.A. Groeneveld, G. Meeden, Measuring skewness and kurtosis, J. R. Stat. Soc. Series D (the Statistician) 33 (4) (1984) 391–399.

[22] D.V. Hinkley, On power transformations to symmetry, Biometrika 62 (1) (1975) 101–111.
[23] H.E.T. Holgersson, A modified skewness measure for testing asymmetry, Commun. Stat. Simul. Comput. 39 (2) (2010) 335–346.
[24] T. Isogai, On a measure of multivariate skewness and a test for multivariate normality, Ann. Inst. Statist. Math. 34 (1982) 531–541.
[25] S.R. Jammalamadaka, E. Taufer, G.H. Terdik, On multivariate skewness and Kurtosis, SankhyĀ A (2020) 1–38.
[26] M.C. Jones, Generating distributions by transformation of scale, Statist. Sinica (2014) 749–771.
[27] M.C. Jones, On families of distributions with shape parameters, Internat. Statist. Rev. 83 (2) (2015) 175–192.
[28] M. Kasy, Uniformity and the Delta method, J. Econom. Methods 8 (1) (2019) 1–19.
[29] T. Kollo, Multivariate skewness and Kurtosis measures with an application in ICA, J. Multivariate Anal. 99 (10) (2008) 2328–2338.
[30] V.D. Konakov, On the asymptotic normality of the mode of multidimensional distributions, Theory Probab. Appl. 18 (1974) 794–799.
[31] C. Ley, D. Paindaveine, Multivariate skewing mechanisms: a unified perspective based on the transformation approach, Statist. Probab. Lett. 80 (23–24) (2010) 1685–1694.
[32] H.L. MacGillivray, Skewness and asymmetry: measures and orderings, Ann. Statist. 14 (3) (1986) 994–1011.
[33] J.F. Malkovich, A.A. Afifi, On tests for multivariate normality, J. Amer. Statist. Assoc. 68 (341) (1973) 176–179.
[34] K.V. Mardia, Measures of multivariate skewness and Kurtosis with applications, Biometrika 57 (3) (1970) 519–530.
[35] T.F. Móri, V.K. Rohatgi, G.J. Székely, On multivariate skewness and Kurtosis, Theory Probab. Appl. 38 (3) (1994) 547–551.
[36] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Stat. 33 (3) (1962) 1065–1076.
[37] P.N. Patil, P.P. Patil, D. Bagkavos, A measure of asymmetry, Statist. Papers 53 (4) (2012) 971–985.
[38] K. Pearson, Contributions to the mathematical theory of evolution, Philos. Trans. R. Soc. London Series A 185 (1894) 71–110.
[39] K. Pearson, Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material, Philos. Trans. R. Soc. London Series A 186 (1895) 343–414.
[40] M. Samanta, Nonparametric estimation of the mode of a multivariate density, South Afr. Stat. J. 7 (2) (1973) 109–117.
[41] D.W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, second ed., in: A Wiley-interscience publication, Wiley, New York, 2015.
[42] K. Song, Rényi information, loglikelihood and an intrinsic distribution measure, J. Statist. Plann. Inference 93 (1–2) (2001) 51–69.
[43] M.S. Srivastava, A measure of skewness and Kurtosis and a graphical method for assessing multivariate normality, Statist. Probab. Lett. 2 (5) (1984) 263–267.
[44] G. Terdik, Multivariate Statistical Methods; Going beyond the Linear, Springer Nature; Cham, Switzerland, 2021.
[45] W.R. van Zwet, Convex Transformations of Random Variables, in: Mathematical Centre Tracts 7, Mathematisch Centrum, Amsterdam, 1964.
[46] P.T. von Hippel, Mean, median, and skew: Correcting a textbook rule, J. Stat. Ed. 13 (2) (2005).
[47] H. Yamato, Sequential estimation of a continuous probability density function and mode, Bull. Math. Stat. 14 (1971) 1–12.
[48] G.U. Yule, On the methods of measuring association between two attributes, J. R. Stat. Soc. 75 (6) (1912) 579–652.