Chilean Journal of Statistics
Vol. 5, No. 1, April 2014, 3–18

Parametric Inference
Research Paper

Sub-model identification in the context of flexible distributions obtained by perturbation of symmetric densities

Anna Clara Monti[1,*], and Rosa Fucci[2]

[1]DEMM Department, Piazza Arechi II, 82100 Benevento, Italy,
[2]Gestore dei Servizi Energetici, Viale Maresciallo Pilsudski, 92, 00197 Roma, Italy

(Received: 27 February 2012 · Accepted in final form: 6 January 2014)

Abstract

Distributions obtained by perturbation of symmetric densities provide flexible models suitable to fit the distribution of data affected by departures from normality, in particular when such deviations are due to skewness and/or heavy tails. However, the adoption of these models may lead to inefficient estimators when the data are generated by a simpler distribution. Consequently a testing strategy aimed at finding the most parsimonious model among non nested ones is proposed. The corresponding test statistics are slight modifications of well-known ones, and their asymptotic distributions do not depend on nuisance parameters. The normality test is the final step of the procedure. Analytical results provide the statistical properties of the proposed tests whereas their performance in finite samples is investigated through numerical experiments.

Keywords: Departures from normality · Generalized skew-normal distributions · Kurtosis · Sign test · Skewness · Skew-symmetric distributions.

Mathematics Subject Classification: Primary 62F03 · Secondary 62F05.

1. Introduction

Suppose that a random sample $X_1, \ldots, X_n$ is drawn from an unknown distribution which could be normal as well as skewed and/or leptokurtic. In these circumstances a flexible model, that is a family of distributions suitable to deal with both skewness and heavy tails, can be assumed for the random variable X which represents the population. When deviations from normality occur, flexible models have two main advantages over robust procedures: inference can be carried out through the likelihood function in the standard way and the estimated parameters have a clear meaning with reference to a model (DiCiccio and Monti, 2004). Therefore, when nonnormal features are likely to arise in the distribution of the data, a strategy for model fitting could be the following: initially the most flexible model is assumed, then test procedures are carried out to check whether a sub-model, or even the normal one, is suitable for the data at hand.

*Corresponding author. Email: acmonti@unisannio.it

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

Distributions generated by perturbation of symmetric densities (Azzalini and Capitanio, 2003), also known as skew-symmetric distributions (Wang *et al.*, 2004), provide a wide class of flexible models. Let $g(\cdot)$ be a symmetric density function, and let $\pi(\cdot)$ be a function satisfying $0 \leq \pi(-x) = 1 - \pi(x) \leq 1$, so that $\pi(0) = 1/2$. A flexible model generated by perturbation of symmetry has the density function

$$f(x; \xi, \omega, \pi) = \frac{2}{\omega} g\left(\frac{x-\xi}{\omega}\right) \pi\left(\frac{x-\xi}{\omega}\right). \tag{1}$$

where ξ and ω are the location and scale parameter, respectively. Since (1) can exhibit both skewness and kurtosis, ξ and ω generally do not coincide with the mean and the standard deviations of (1), this is why a notation different from the usual μ and σ is used. Additional parameters may appear in both $g(\cdot)$ and $\pi(\cdot)$, and in particular $g(\cdot)$ often depends on a parameter τ regulating the tickness of the tails.

The function $g(\cdot)$ is also referred to in the literature as the base or kernel density or the central model. The function $\pi(\cdot)$ is instead called perturbing function. When $\pi(\cdot)$ is a constant function identically equal to 1/2, $f(x; \xi, \omega, \pi)$ reduces to the symmetric density function $g(\cdot)$. Azzalini and Capitanio (2003) consider flexible distributions with the parameterization $\pi(x) = H\{w(z)\}$ where $z = (x - \xi)/\omega$, $H(\cdot)$ is the distribution function of a continuous random variable symmetric around zero and $w(\cdot)$ is an odd function, that is $w(-z) = -w(z)$. Often, but not necessarily, $\pi(x) = H(\lambda x)$, in this case $\lambda = 0$ reduces (1) to a symmetric density, hence λ becomes the parameter regulating skewness.

Remarkable examples of flexible models are the skew-normal (Roberts, 1966; Azzalini, 1985, 1986), the skew-t (Branco and Dey, 2001; Azzalini and Capitanio, 2003), and the skew Exponential Power distribution (Azzalini, 1986; DiCiccio and Monti, 2004). See Genton (2004) and Azzalini (2005) for overviews on flexible distributions obtained by perturbation of symmetry.

The selection of the base density $g(\cdot)$ has received special attention in the literature and a very popular choice assumes that the central model is $N(\xi, \omega^2)$. Other interesting models, within the class of distributions defined by (1), are those where there is a specific or limiting (possibly infinite) value τ_0 of the parameter τ such that $g(\cdot)$ is the $N(\xi, \omega^2)$ density. This is the case for instance of the skew-t or the skew Exponential Power distribution. In this case, when $\tau = \tau_0$, the density function $f(x; \xi, \omega, \pi)$ becomes a generalized skew-normal (GSN) density (Loperfido, 2004; Ma *et al.*, 2005). It is given by

$$f_{GSN}(x; \xi, \omega, \lambda, \pi) = \frac{2}{\omega} \phi\left(\frac{x-\xi}{\omega}\right) \pi\left(\frac{x-\xi}{\omega}\right) \tag{2}$$

where $\phi(\cdot)$ is the standard normal density function. If $\pi(x) = \Phi(\lambda x)$, where $\Phi(\cdot)$ denotes the standard normal distribution function, (2) corresponds to the skew normal distribution. Other examples of GSN densities are discussed in Nadarajah and Kotz (2003), whereas inferential issues are investigated by Pewsey (2004) and Ley and Paindaveine (2010a, 2010b).

The class of distributions (2) can arise when the population is normal but a random sample is not available, so that the probability density of the observed sample is distorted by a multiplicative nonnegative weight function. Actually (1) can be applied when the observed data are obtained only from a selected portion of the population of interest and a censoring mechanism regulated by $\pi(\cdot)$ is applied to the samples generated by $g(\cdot)$, hence (1) is called a selection model; see Ma *et al.* (2005), Arellano-Valle *et al.* (2006) and references therein. When the central model is $N(\xi, \omega^2)$, the selective sampling procedure yields a GSN distribution with density (2).

The occurrence of a GSN distribution can be motivated by a robustness argument too. If the central model, which describes the majority of the data, is normal but it is contaminated through the function $\pi(\cdot)$ which can generate also asymmetric outliers, the density of the observed data is given by (2) (Ma *et al.*, 2005).

Either when a selective sampling procedure is implemented or in the robustness context, it is of interest to detect whether the distribution has the density (2), since it involves identification of the central model. Under this perspective, subsequent inference may focus only on the parameters of the base normal density while $\pi(\cdot)$ is regarded as a nuisance component, whose form has no relevance.

Briefly, when $\pi(x) = 1/2$ (1) yields a symmetric density, when the base density is $\phi(\cdot)$ it becomes a GSN density, and when both $\pi(x) = 1/2$ and $g(x) = \phi(x)$ hold (1) reduces to a normal density function. If any of these circumstances occurs, (1) is over-parameterized, and its adoption can produce remarkable losses of efficiency in the estimation of the parameters. To avoid this inconvenience, reliable sub-model testing procedures are required. In particular, there is the need to test the hypotheses $H_0^{\pi} : \pi(x) = 1/2$ for symmetry and $H_0^{GSN} : g(x) = \phi(x)$ for generalized skew-normality. The normality test is the implicit joint outcome of the two tests.

Suppose that T_n^{π} and T_n^{GSN} are the statistics used for testing H_0^{π} and H_0^{GSN}, respectively. When testing H_0^{GSN} the function $\pi(\cdot)$ has a disturbing role (as an infinite-dimensional nuisance parameter) and, analogously, when testing H_0^{π} the form of $g(\cdot)$ has a disturbing role. Widely applicable identification procedures should satisfy the following two conditions:

(i) The (asymptotic) distribution of T_n^{π} under H_0^{π} does not depend on $g(\cdot)$;
(ii) The (asymptotic) distribution of T_n^{GSN} under H_0^{GSN} does not depend on $\pi(\cdot)$.

Let R^{π} and R^{GSN} be the rejection regions for H_0^{π} and H_0^{GSN}, respectively. The significance level of the normality test, with null hypothesis $H_0^{N} : \pi(x) = 1/2$ and $g(x) = \phi(x)$, is

$$1 - \text{pr}\left(T_n^{GSN} \notin R^{GSN}; T_n^{\pi} \notin R^{\pi}\right) \tag{3}$$

evaluated under the normal distribution. The probability (3) typically cannot be computed unless the joint distribution of T_n^{π} and T_n^{GSN} is known or the two statistics are independent under the normal model. In the latter case the significance level of the normality test is

$$1 - \text{pr}\left(T_n^{\pi} \notin R^{\pi}\right) \text{pr}\left(T_n^{GSN} \notin R^{GSN}\right).$$

Hence a further condition, which is required in order to keep under control the significance level of the joint normality test, is

(iii) T_n^{π} and T_n^{GSN} are (asymptotically) independent under the normal model.

Condition (iii) avoids the well-known problems arising in the evaluation of the significance levels of the tests on non-nested hypotheses.

When testing normality the standardized sample skewness and kurtosis are default choices for T^{π} and T^{GSN}, with the advantage of being asymptotically independent under the null hypothesis. However, when either skewness or leptokurtosis occurs, the asymptotic distribution of these statistics depends on the underlying distribution and conditions (i) - (iii) may fail to hold. Therefore an alternative statistic for testing H_0^{π} and a modification of the kurtosis test for H_0^{GSN} will be considered.

The present paper focuses on model identification and proposes some testing procedures based on statistics T^{π} and T^{GSN} which satisfy conditions (i)-(iii). It is structured as follows. The next section outlines the testing strategy. Sections 3 and 4 propose the test procedures for the hypothesis of symmetry and that of generalized skew-normality, respectively. Numerical results are illustrated in section 5, while proofs are confined in the Appendix.

2. Testing Strategy

We shall first consider a symmetry test satisfying condition (i). Attention will be kept on distributions with density of type (1) with a monotone perturbing function. This restriction is indeed very mild since it will be suitable for most of the unimodal distributions obtained through (1). We have the following proposition.

Proposition 2.1 Let X be a random variable with density (1). If $\pi(x)$ is a monotone function, then ξ is the median of X if and only if $\pi(x) = 1/2$.

Consequently, if ξ were known, the hypothesis H_0^π of symmetry, which assume that X has the density $g(\cdot)$, could be verified through the sign test. It is based on the frequency of observations greater than ξ,

$$f = \frac{1}{n} \sum_{i=1}^{n} I\left(X_i > \xi\right)$$

where $I(\delta)$ is an indicator function which takes value 1 when δ holds. The test statistic, given by $2n^{1/2}\left(f - 1/2\right)$, has an asymptotic $N(0, 1)$ distribution. The sign test has an intuitive appeal, it is very robust within the class of symmetric distributions and possesses optimal properties (Rohatgi, 1976, pag. 544-546). The next section will show how to implement this test when, as it usually occurs in practice, ξ needs to be estimated.

We shall now consider a test for generalized skew-normality satisfying condition (ii). It can be derived by exploiting the property of distributional invariance held by the distributions of type (1), which is stated in the following proposition proved by Azzalini and Capitanio (2003, Proposition 2) and Wang *et al.* (2004, Proposition 6).

Proposition 2.2 Let X and Y be two random variables having density function (1) and $g\left(y\right)$, respectively. If $T(\cdot)$ is an even function, that is $T(-x) = T(x)$ for all $x \in R$, then $T(X)$ is identically distributed to $T(Y)$: $T(X) \stackrel{d}{=} T(Y)$.

On the bases of Proposition 2.2, if X has density (1), the random variable $Z = \left(X - \xi\right)/\omega$ is such that $|Z|$ is identically distributed to $|Y|$, where $Y \sim g(y)$, regardless of the form of the perturbing function $\pi(\cdot)$. Hence the problem of testing whether the distribution is GSN is equivalent to testing whether $|Z|$ as the same distribution of the absolute value of a $N(0, 1)$ random variable, i.e. a χ_1 random variable. In this context, the alternative hypothesis assumes departures of $g(\cdot)$ from normality due to the tickness of the tails, therefore a test which focuses on the kurtosis of the distribution seems the most natural choice (see Thode, 2002, pp. 50-54, for the properties of the kurtosis test in the Gaussian context against symmetric alternatives). Consequently a test based on a statistic T_n^{GSN} which is a function of the fourth moment of Z will be proposed in Section 4.

A normality test based on the above test statistics satisfies requirement (iii). If X is symmetric around ξ, $|Z|$ and $\mathrm{sign}(Z)$ are independently distributed (Arellano Valle and Del Pino, 2004). The statistic T_n^π, based on $I\left(X_i > \xi\right)$ for $i = 1 \ldots n$, and the statistic T_n^{GSN}, based on $|Z_1|, \ldots, |Z_n|$, are independent for any underlying symmetric distribution and therefore at the normal model.

In the general case, both ξ and ω are unknown and need to be estimated. Different estimators will be considered for the two tests, such that they are consistent under the corresponding null hypothesis. Consequently consistency is preserved at the normal model, and conditions (i)-(iii) can be expected, at least approximately, to hold for sufficiently large samples.

3. Symmetry test

Suppose X has the symmetric density $\omega^{-1} g\left\{(x-\xi)/\omega\right\}$. Under this assumption, the location parameter ξ is the center of symmetry and coincides with both the median and the mean, therefore the modified sign test proposed by Gastwirth (1971) can be used to verify the hypothesis of symmetry H_0^π. It is based on the following theorem (Gastwirth, 1971, Theorem 1)

Theorem 3.1 Let $X_1, \ldots, X_n$ be i.i.d. observations from an absolutely continuous distribution $G(\cdot)$,with mean ξ, variance σ^2 and density $g(\cdot)$ continuous at ξ. Let

$$\hat{f} = \frac{1}{n}\sum_{i=1}^{n} I\left(X_i > \bar{X}\right)$$

where $\bar{X}$ is the sample mean. The statistic $n^{1/2}\left(\hat{f} - 1/2\right)$ is asymptotically normally distributed with mean zero and variance

$$\sigma_f^2 = \frac{1}{4} + g(\xi)^2 \sigma^2 - 2g(\xi)\gamma$$

where

$$\gamma = \operatorname{cov}\left\{n^{1/2}\left(\bar{X} - \xi\right), n^{1/2}(f - 1/2)\right\} = \int_\xi^\infty (x-\xi)\, g(x) dx.$$

Theorem 3.1 shows that, although $n^{1/2}(\hat{f} - 1/2)$ is asymptotically normal with mean zero, as $n^{1/2}(f - 1/2)$, its asymptotic variance is different from $1/4$, and there is indeed evidence that it is typically quite smaller (see also Theorem 2 by Gastwith, 1971).

In order to implement the test a consistent estimator of σ_f^2 is required. A sample version of γ is given by

$$\hat{\gamma} = \frac{1}{n}\sum_{i=1}^{n} I\left(X_i > \bar{X}\right)\left(X_i - \bar{X}\right),$$

so that the estimator of the variance can be obtained as follow

$$\hat{\sigma}_f^2 = \frac{1}{4} + \hat{g}\left(\bar{X}\right)^2 \hat{\sigma}^2 - 2\hat{g}\left(\bar{X}\right)\hat{\gamma}$$

where $\hat{g}(\cdot)$ is a consistent nonparametric estimator of the density function and $\hat{\sigma}^2$ is the sample variance. The test statistic for H_0^π is given by $T_n^\pi = n^{1/2}\left(\hat{f} - 1/2\right)/\hat{\sigma}_f$.

4. Generalized skew-normality test

This section concerns a procedure for testing H_0^{GSN}, which implies that X is GSN with density (2). Since, under the null hypothesis, Z is identically distributed to the absolute value of a $N(0,1)$ random variable, $\kappa = E\left(Z^4\right)$ takes value 3. Consequently if ξ were known, the test could be based on the following result.

Proposition 4.1 Let the statistics $\hat{\omega}$ and $\hat{\kappa}_\xi$ be defined as follows

$$\hat{\omega}_\xi^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\xi)^2, \qquad \hat{\kappa}_\xi = \frac{1}{n}\sum_{i=1}^{n}\frac{(X_i-\xi)^4}{\hat{\omega}_\xi^4}$$

where $X_1,\ldots,X_n$ is a random sample drawn from a population with density (2) and location parameter ξ. Then the asymptotic distribution of $n^{1/2}(\hat{\kappa}_\xi-3)/24$ is standard normal.

In practical applications unfortunately the value of ξ is not available. However the following proposition provides a statistic useful to test H_0^{GSN}.

Proposition 4.2 Let $\hat{\xi}$ be an estimator for ξ such that, when $X_1,\ldots X_n$ have density (2), $\hat{\xi}-\xi = O_p\left(n^{-1/2}\right)$, $E\left(\hat{\xi}\right) = \xi + O\left(n^{-1}\right)$, and $n^{1/2}\left(\hat{\xi}-\xi\right)$ is asymptotically normally distributed with variance σ_ξ^2. Consider the statistics

$$\hat{\omega}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\hat{\xi}\right)^2, \qquad \hat{\kappa} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(X_i-\hat{\xi}\right)^4}{\hat{\omega}^4},$$

then $n^{1/2}\left(\hat{\kappa}-3\right)$ is asymptotically normally distributed with variance

$$\sigma_\kappa^2 = 24 + \frac{16\eta^2}{\omega^2}\sigma_\xi^2 + \frac{8\eta}{\omega^5}\gamma_{\xi,m_4} - \frac{48\eta}{\omega^3}\gamma_{\xi,m_2} \tag{4}$$

where $\eta = \left(3\mu_1/\omega - \mu_3/\omega^3\right)$, $\mu_r = E\left(m_r\right)$, $m_r = (1/n)\sum_{i=1}^{n}\left(X_i-\xi\right)^r$, and

$$\gamma_{\xi,m_r} = \operatorname{cov}\left\{n^{1/2}\left(\hat{\xi}-\xi\right), n^{1/2}\left(m_r-\mu_r\right)\right\}.$$

Proposition 4.2 suggests that the asymptotically normal statistic $T_n^{GSN} = n^{1/2}\left(\hat{\kappa}-3\right)/\sigma_\kappa$ can be used to test the generalized skew-normal hypothesis H^{GSN}. To implement the test, an estimator $\hat{\xi}$ of ξ (which in this context is the mean of the central $N\left(\xi,\omega^2\right)$ model) is required. In order to satisfy condition (ii) of Section 1, such estimator should be semiparametric, in the sense that it should not depend on the form of $\pi(\cdot)$. It needs to be derived by focusing on the base normal density, whereas $\pi(\cdot)$ is treated nonparametrically. Estimators which satisfy this requirement have been proposed by Ma *et al.* (2005) and Azzalini *et al.* (2010). In particular, Ma *et al.* (2005) have proposed a locally efficient semiparametric estimator of the parameters of the base density based on the theory of regular asymptotically linear (RAL) estimators, which are consistent and asymptotically normal regardless of the possible misspecification of the perturbing function $\pi(\cdot)$ (see also Ma and Hart, 2007). The estimator proposed by Azzalini *et al.* (2010)

relies on a conceptually simpler motivation based on the invariance property of Proposition 2.2, leading to the invariance-based estimating equations (IBEE). The IBEE approach can be considered as a special case of the RAL formulation, and therefore it inherits the same properties. Further details on the IBEE will be discussed in Section 5.

To apply the test, suggested by Proposition 4.2, an estimate of σ^2_κ is required. For the variance of the estimator σ^2_ξ, either an analytical formula (if available) or a bootstrap estimator of the variability of $\hat{\xi}$ can be employed. The covariances γ_{ξ,m_r} instead can be estimated by bootstrap. Let $X^*_1, \ldots X^*_n$ be a sample drawn, with replacement, from the original sample $X_1, \ldots, X_n$, which yields the bootstrap replication $\hat{\xi}^*$ of $\hat{\xi}$. In the bootstrap context the role of the true parameter is played by $\hat{\xi}$ while the sample is given by $X^*_1, \ldots X^*_n$. It follows that the statistics $m^*_r = (1/n)\sum_{i=1}^n \left(X^*_i - \hat{\xi}\right)^r$ are the boostrap versions of the m_r's, for $r = 2, 4$. Consequently estimates of γ_{ξ,m_r} are obtained from the bootstrap covariances between the m^*_r's and $\hat{\xi}^* - \hat{\xi}$.

Alternatively to obtain the covariances which appear in (4) we can approximate $\hat{\xi} - \xi$ by means of the influence function. Under regularity conditions we have the Von Mises expansions (Serfling, 1980, pag. 212)

$$n^{1/2}\left(\hat{\xi} - \xi\right) = \frac{1}{n^{1/2}}\sum_{i=1}^n IF(X_i, \xi) + O_p\left(n^{-1/2}\right)$$

where $IF(X, \xi)$ is the influence function (see Maronna *et al.*, 2006, pag. 55-57; Hampel *et al.*, 1986, pag. 84). The asymptotic variance of $n^{1/2}\left(\hat{\xi} - \xi\right)$ is given by $\sigma^2_\xi = E\left\{IF(X, \xi)^2\right\}$ (see 3.17 of Maronna *et al.*, 2006). The covariances γ_{ξ,m_r} for $r = 2, 4$ are given by

$$\gamma_{\xi,m_r} = E\left\{n^{1/2}\left(\hat{\xi} - \xi\right)\frac{1}{n^{1/2}}\sum_{i=1}^n (X_i - \xi)^r\right\} = E\left\{IF(X, \xi)\,(X - \xi)^r\right\}.$$

Sample versions of σ^2_ξ and γ_{ξ,m_r} are obtained as follows

$$\hat{\sigma}^2_\xi = \frac{1}{n}\sum_{i=1}^n \hat{IF}(X_i, \xi)^2 \qquad (5)$$

$$\hat{\gamma}_{\xi,m_r} = \frac{1}{n}\sum_{i=1}^n \hat{IF}(X_i, \xi)\left(X_i - \hat{\xi}\right)^r \qquad for \quad r = 2, 4 \qquad (6)$$

where $\hat{IF}(X_i, \xi)$ is an empirical version of the influence function.

Finally μ_1 and μ_3 can be estimated by their sample versions.

Of course one may also consider estimating σ^2_κ directly by bootstrap. However the results are likely to be unsatisfactory. The accuracy of the bootstrap estimators of the standard errors depends on various factors (Efron & Tibshirani, 1993, pp. 50-53, 272-273, 280-281): the variability of the standard error obtained from infinite bootstrap replications se_∞, the sample size n and the actual number of bootstrap replications B. Estimating σ^2_κ is equivalent to estimating the variance of a fourth standardized cumulants, that is a moment of 8th order, which typically has a very high variability, i.e. se_∞ is remarkably big. It follows that unless n is extremely high, whatever B is, the estimator may be considerably inaccurate. On the contrary, the strategy of estimating the terms, σ_ξ^2 and γ_{ξ,m_r} for $r = 2, 4$, separately allow to reduce the order of the moments to be estimated, which turns out to be at most 5. Furthermore, this strategy takes advantage of the knowledge of the constants appearing in (4) which would, otherwise, be uselessly estimated.

5. Numerical results

This section contains results of simulations carried out to evaluate the performance of the proposed test procedures.

For the test of symmetry, the $N(0, 1)$ distribution, the Student t distribution with degrees of freedom $\nu = 3, 5, 7, 10, 15$, the uniform and the logistic distribution were considered. For each of them, and for sample sizes n varying between 50 and 50·000, 10·000 samples were generated, and the test was implemented according to the procedure described in Section 3. Table 1 shows the actual level of the symmetry test when the nominal size is either 5% or 10%. The simulated level of the test is indeed pretty close to the nominal one.

Table 1. Frequency of rejections of H_0^π - 10·000 simulations

Nominal significance level 0.05

	Density							
n	t_3	t_5	t_7	t_{10}	t_{15}	*Logistic*	*Uniform*	*Normal*
50	0.060	0.045	0.045	0.041	0.037	0.042	0.032	0.032
100	0.060	0.054	0.049	0.048	0.047	0.045	0.049	0.049
200	0.060	0.055	0.050	0.054	0.053	0.050	0.050	0.046
1·000	0.055	0.053	0.057	0.048	0.052	0.052	0.050	0.049
2·000	0.057	0.052	0.056	0.052	0.049	0.053	0.050	0.050
5·000	0.057	0.054	0.052	0.050	0.049	0.051	0.049	0.053
10·000	0.054	0.050	0.054	0.050	0.049	0.053	0.047	0.049
50·000	0.050	0.050	0.052	0.049	0.056	0.050	0.052	0.049

Nominal significance level 0.10

	Density							
n	t_3	t_5	t_7	t_{10}	t_{15}	*Logistic*	*Uniform*	*Normal*
50	0.121	0.105	0.105	0.101	0.099	0.105	0.085	0.092
100	0.115	0.102	0.987	0.095	0.092	0.089	0.105	0.099
200	0.124	0.110	0.996	0.101	0.097	0.098	0.106	0.097
1·000	0.108	0.105	0.112	0.100	0.101	0.104	0.102	0.104
2·000	0.114	0.103	0.106	0.101	0.100	0.102	0.096	0.100
5·000	0.110	0.104	0.106	0.099	0.101	0.100	0.098	0.100
10·000	0.103	0.105	0.103	0.099	0.099	0.102	0.097	0.099
50·000	0.102	0.098	0.102	0.098	0.103	0.104	0.104	0.099

To investigate the performance of the GSN test, samples were drawn from the skew normal distribution (Azzalini, 1985, 1986), with density function

$$f_{SN}(x; \xi, \omega, \lambda) = \frac{2}{\omega}\,\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\lambda\frac{x-\xi}{\omega}\right). \tag{7}$$

It returns the $N\left(\xi, \omega^2\right)$ distribution when $\lambda = 0$, it has positive skewness for $\lambda > 0$ and it is negatively skewed otherwise. The values of the shape parameter, which were taken into account, are $\lambda = 1.5, 3, 5, 10$; only positive values were considered for λ since the shape of the distribution is specular for negative values. Finally 10·000 samples were drawn from (7) for various sample sizes n between 100 and 5·000.

The parameters ξ and ω of the normal central model were estimated through the IBEE approach of Azzalini *et al.* (2010). Because of the invariance property of Proposition 2.2, under (2) $Z = |X - \xi|/\omega$ has the same moments of a χ_1 random variable, and in particular $E|Z| = (2/\pi)^{1/2}$ and $E|Z|^2 = 1$. The estimating equations are obtained by setting the sample averages of $|Z|$ and Z^2 equal to their expected values. The estimator $\hat{\xi}$ of ξ is given by the solution t of

$$\frac{1}{n}\sum_{i=1}^{n}|X_i - t| - \left(\frac{2}{\pi}\right)^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{n}(X_i - t)^2\right\}^{1/2} = 0$$

whereas the estimator of ω^2 is given by $\hat{\omega}^2$ of Proposition 4.2 (see Azzalini *et al.* (2010) for computational issues related to this method). The corresponding estimating equations are $\psi(x; \xi, \omega) = \left\{|z| - (2/\pi)^{1/2}, z^2 - 1\right\}$ where $z = (x - \xi)/\omega$.

Four methods were considered for estimating σ_κ^2. The first method ($M1$) estimates the covariances γ_{ξ,m_r} for $r = 2,4$ by 100 bootstrap replications while σ_ξ^2 is estimated by (5) which coincides with formula (16) of Azzalini *et al.* (2010). The empirical version of the influence function is obtained by replacing the parameters by $\hat{\xi}$ and $\hat{\omega}^2$ in (3.10) of Maronna et al. (2006). In the second method ($M2$), both σ_ξ^2 and γ_{ξ,m_r} for $r = 2,4$ are approximated through 100 bootstrap replications. The third method ($M3$) uses (5) and (6) for estimating σ_ξ^2 and γ_{ξ,m_r}, respectively. Finally, in the last method ($M4$), σ_κ^2 is estimated directly through 100 bootstrap replications. Tables from 2 to 5 show the simulated significance level of the GSN test with the four method used to estimate σ_κ^2. All the tables consider the nominal levels 0.10 and 0.5.

Table 2. Frequency of rejections of H_0^{GSN}. Test $M1$ performed by estimating σ_ξ^2 by (5) and γ_{ξ,m_r} for $r = 2,4$ by 100 bootstrap replications - 10·000 simulations.

Nominal Level	0.10				0.05			
	α				α			
n	1.5	3	5	10	1.5	3	5	10
100	0.036	0.055	0.054	0.058	0.025	0.039	0.037	0.040
200	0.052	0.069	0.071	0.074	0.037	0.048	0.051	0.053
500	0.085	0.093	0.089	0.083	0.061	0.065	0.064	0.059
1000	0.115	0.105	0.099	0.100	0.082	0.074	0.071	0.066
2000	0.133	0.110	0.116	0.107	0.094	0.075	0.077	0.069
5000	0.133	0123	0.115	0.114	0.091	0.080	0.070	0.065

Table 3. Frequency of rejections of H_0^{GSN}. Test $M2$ performed by estimating σ_ξ^2 and γ_{ξ,m_r} for $r = 2,4$ by 100 bootstrap replications - 10·000 simulations.

Nominal Level	0.10				0.05			
	α				α			
n	1.5	3	5	10	1.5	3	5	10
100	0.032	0.046	0.045	0.048	0.023	0.033	0.032	0.036
200	0.047	0.054	0.063	0.068	0.033	0.039	0.048	0.049
500	0.063	0.081	0.086	0.080	0.045	0.060	0.061	0.058
1000	0.079	0.095	0.098	0.098	0.059	0.067	0.070	0.065
2000	0.092	0.102	0.112	0.104	0.066	0.069	0.072	0.067
5000	0.101	0.114	0.107	0.110	0.065	0.070	0.063	0.062

Table 4. Frequency of rejections of H_0^{GSN}. Test $M3$ performed by estimating σ_ξ^2 by (5) and γ_{ξ,m_r} for $r = 2, 4$ by (6) - 10·000 simulations.

Nominal Level	0.10				0.05			
	α				α			
n	1.5	3	5	10	1.5	3	5	10
100	0.026	0.039	0.039	0.039	0.016	0.027	0.026	0.025
200	0.040	0.046	0.052	0.054	0.026	0.031	0.035	0.039
500	0.057	0.066	0.070	0.069	0.039	0.045	0.045	0.044
1000	0.079	0.080	0.098	0.088	0.053	0.053	0.070	0.054
2000	0.096	0.091	0.103	0.094	0.064	0.058	0.064	0.059
5000	0.103	0.107	0.100	0.106	0.066	0.066	0.058	0.058

Table 5. Frequency of rejections of H_0^{GSN}. Test $M4$ performed by estimating σ_κ^2 through 100 bootstrap replications - 10·000 simulations.

Nominal Level	0.10				0.05			
	α				α			
n	1.5	3	5	10	1.5	3	5	10
100	0.203	0.184	0.206	0.230	0.140	0.123	0.146	0.161
200	0.161	0.181	0.200	0.208	0.107	0.121	0.139	0.147
500	0.132	0.168	0.165	0.176	0.081	0.108	0.111	0.119
1000	0.115	0.148	0.144	0.154	0.069	0.098	0.095	0.098
2000	0.119	0.134	0.138	0.132	0.070	0.079	0.083	0.078
5000	0.103	0.117	0.115	0.122	0.054	0.065	0.066	0.071

For $n \leq 200$ and σ_κ^2 estimated by $M1$, $M2$, and $M3$, the test appears to be conservative, however the significance level gets satisfactorily close to the nominal one for larger sample sizes. This outcome is due to the slow convergence of the distribution of the test statistic to the normal distribution. A large sample is required for the normal approximation to the distribution of T_n^{GSN} to be entirely satisfactory. In this regard, it is to be considered that even the kurtosis test in the Gaussian context tends to be pretty conservative in small samples. Indeed it needs sample sizes even larger than 5·000 to be accurate (Thode, 2002, pag. 51) for at least a couple of reasons: the mean and variance of the test statistic converge very slowly to their asymptotic values and the convergence of the distribution of the test statistic to the $N(0, 1)$ distribution is very slow. Analogous events occur in this context too, with the additional inconvenience that also the distribution of the estimator of ξ converges very slowly to normality. Such inconveniences are exacerbated when λ is close to zero, that is in a neighborhood of the normal model, where inference in the skew normal model is known to be quite problematic (see Azzalini, 1985; Azzalini and Capitanio, 2003; Pewsey, 2004).

The results of tables 2, 3, and 4, on the whole, appear quite satisfactory. When γ_{ξ,m_r} is estimated by bootstrap, unless the sample is very large, it appears more convenient to estimate σ_ξ^2 by (5) rather than by bootstrap, i.e. $M1$ is slightly preferable to $M2$. Nevertheless when the sample sizes n increases the test performed through $M1$ becomes a little too liberal especially for small values of λ. On the contrary the test performed by applying $M2$ shows a more stable asymptotic behaviour. A very satisfactory asymptotic behaviour is also achieved by $M3$.

As anticipated in Section 4, estimating σ_κ^2 directly by bootstrap, i.e. applying $M4$, can lead to a very poor performance of the test, which appear extremely liberal unless the sample size is remarkably large. In any case, for large values of n, $M3$ is definitely to be preferred to $M4$ either in terms of accuracy and for the smaller computing time required.

Table 6 shows the simulated significance level of the joint test on symmetry and generalized skew-normality, that is the normality test. Again the four methods for estimating

σ_κ^2 are compared with each other. The overall nominal level is either $\alpha = 0.10$ or $\alpha = 0.05$. This implies that the nominal level of the symmetry test and that of the generalized skew-normality test is approximately 0.05132 when $\alpha = 0.10$ and 0.02532 when $\alpha = 0.05$. For small sample sizes the test seems to be conservative, with the exception of the case when $M4$ is applied and the test becomes way too liberal, as already noticed for the GSN test too. However when the sample size n increases, the simulated level does get close to the nominal one.

Table 6. Frequency of rejections of the hypothesis of normality $H_0^N : X \sim N(\xi, \omega^2)$ - 10·000 simulations.

Nominal significance level 0.05

	n						
Estimation of σ_κ^2	100	200	500	1·000	2·000	5·000	10·000
$M1$	0.030	0.030	0.035	0.037	0.040	0.042	0.044
$M2$	0.030	0.029	0.033	0.035	0.038	0.036	0.037
$M3$	0.029	0.029	0.033	0.038	0.041	0.040	0.041
$M4$	0.211	0.181	0.148	0.130	0.115	0.096	0.083

Nominal significance level 0.10

	n						
Estimation of σ_κ^2	100	200	500	1·000	2·000	5·000	10·000
$M1$	0.056	0.057	0.066	0.070	0.073	0.077	0.080
$M2$	0.056	0.056	0.065	0.074	0.076	0.079	0.082
$M3$	0.055	0.056	0.066	0.071	0.077	0.080	0.081
$M4$	0.283	0.255	0.217	0.196	0.180	0.157	0.147

Figure 1 shows the scatter plot of the two test statistics T^π and T^{GSN} when the distribution is normal and $n = 1·000$. No considerable dependence appears to occur between T^π and T^{GSN}. This outcome supports the proposal of the joint test, which relies on the, at least approximate, independence between the two statistics when $X \sim N(\xi, \omega^2)$.

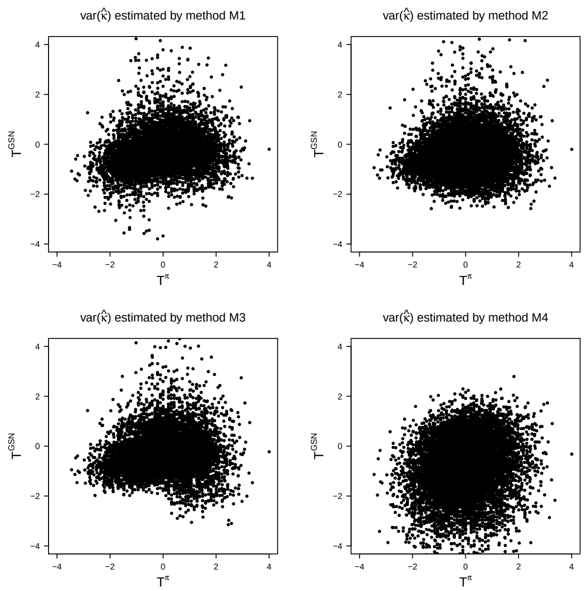

Figure 1. Scatter plot of T^π and T^{GSN} for sample size $n = 1·000$ under normality

6. Appendix

Proof of Proposition 2.1

Consider the integral

$$\int_{-\infty}^{\xi} f(x;\xi,\omega,\pi)\,dx = \frac{2}{\omega}\int_{-\infty}^{\xi} g\left(\frac{x-\xi}{\omega}\right)\pi\left(\frac{x-\xi}{\omega}\right)dx\ ; \tag{8}$$

by integrating (8) by parts we have

$$\int_{-\infty}^{\xi} f(x;\xi,\omega,\pi)\,dx = 2G(0)\,\pi(0) - \frac{2}{\omega}\int_{-\infty}^{\xi} g\left(\frac{x-\xi}{\omega}\right)\pi'\left(\frac{x-\xi}{\omega}\right)dx$$

where $G(\cdot)$ is the distribution function corresponding to $g(\cdot)$ and $\pi'\{(x-\xi)/\omega\}$ is the derivative of $\pi\{(x-\xi)/\omega\}$ with respect to x. Since $G(0)=1/2$ and $\pi(0)=1/2$, it yields

$$\int_{-\infty}^{\xi} f(x;\xi,\omega,\pi)\,dx = \frac{1}{2} - \frac{2}{\omega}\int_{-\infty}^{\xi} g\left(\frac{x-\xi}{\omega}\right)\pi'\left(\frac{x-\xi}{\omega}\right)dx.$$

Hence ξ is the median of X if and only if $\pi'\{(x-\xi)/\omega\}=0$, which implies $\pi\{(x-\xi)/\omega\} = 1/2$.

Proof of Proposition 4.1

Without loss of generality we shall assume $\omega = 1$. Let $m_4 = n^{-1}\sum_{i=1}^{n}(X_i-\xi)^4$ so that $\hat{\kappa}_\xi = m_4 \big/ \hat{\omega}_\xi^4$. A Taylor expansion of $1/\hat{\omega}_\xi^4$ around $\hat{\omega}^2 = 1$ gives $1\big/\hat{\omega}_\xi^4 = 1 - 2\left(\hat{\omega}_\xi^2 - 1\right) + O_p\left(n^{-1}\right)$. Consequently

$$\hat{\kappa}_\xi = m_4 - 2\left(\hat{\omega}_\xi^2 - 1\right)m_4 + O_p\left(n^{-1}\right) \tag{9}$$

where the remaining terms will contribute terms of order n^{-1} in $E(\hat{\kappa}_\xi)$ and terms of order lower than n^{-1} in $\operatorname{var}(\hat{\kappa}_\xi)$. In order to compute the mean and the variance, we consider the expansion of the second term of (9)

$$\begin{aligned}\hat{\kappa}_\xi &= m_4 - 2\left\{E\left(\hat{\omega}_\xi^2\right)-1\right\}E(m_4) - 2\left(\hat{\omega}_\xi^2-1\right)E(m_4)\\ &\quad -2\left\{E\left(\hat{\omega}_\xi^2\right)-1\right\}\{m_4 - E(m_4)\} + O_p\left(n^{-1}\right)\\ &= m_4 - 2\left(\hat{\omega}_\xi^2-1\right)E(m_4) + O_p\left(n^{-1}\right)\end{aligned}$$

since, under (2), $E\left(\hat{\omega}_\xi^2\right) = 1$. Thus $E(\hat{\kappa}_\xi) = E(m_4) + O\left(n^{-1}\right)$ and

$$\operatorname{var}(\hat{\kappa}_\xi) = \operatorname{var}(m_4) + 4E\{(m_4)\}^2\operatorname{var}\left(\hat{\omega}_\xi^2\right) - 4E(m_4)\operatorname{cov}\left(\hat{\omega}_\xi^2, m_4\right) + o\left(n^{-1}\right)\ .$$

Under (2), $\hat{\omega}_\xi^2$ and m_4 are the second and fourth sample central moment, respectively, of a random sample drawn from a $N(0,1)$ distribution, hence $E(m_4) = 3$, $\text{var}\left(\hat{\omega}_\xi^2\right) = 2/n$, $\text{var}(m_4) = 96/n$ and $\text{cov}\left(\hat{\omega}_\xi^2, m_4\right) = 12/n$ (Serfling, 1980, pag. 68). It yields $E(\hat{\kappa}_\xi) = 3 + O\left(n^{-1}\right)$ and $\text{var}(\hat{\kappa}_\xi) = 24/n + O\left(n^{-2}\right)$. Since $\hat{\omega}_\xi^2$ converges in probability to $\omega^2 = 1$, $\hat{\kappa}_\xi$ has the same asymptotic distribution of $n^{-1}\sum_{i=1}^n \{(X_i - \xi)/\omega\}^4$. Consequently, by the central limit theorem, the asymptotic distribution of $n^{1/2}(\hat{\kappa}_\xi - 3)/24^{1/2}$ is $N(0,1)$.

Proof of Proposition 4.2

A Taylor expansion of $\hat{w}^2$ with respect to $\hat{\xi}$ around ξ yields

$$\hat{\omega}^2 = \hat{\omega}_\xi^2 - 2\left(\hat{\xi} - \xi\right) m_1 + O_p\left(n^{-1}\right) \ . \tag{10}$$

Furthermore we have

$$\frac{1}{\hat{\omega}^4} = \frac{1}{\hat{\omega}_\xi^4} - 2\frac{1}{\hat{\omega}_\xi^6}\left(\hat{\omega}^2 - \hat{\omega}_\xi^2\right) + O_p\left(n^{-1}\right) \ , \tag{11}$$

and by replacing (10) in (11) it yields

$$\frac{1}{\hat{\omega}^4} = \frac{1}{\hat{\omega}_\xi^4} + 4\frac{1}{\hat{\omega}_\xi^6}\left(\hat{\xi} - \xi\right) m_1 + O_p\left(n^{-1}\right) \ . \tag{12}$$

Moreover, since $\hat{\omega}_\xi^2 - \omega^2 = O_p\left(n^{-1/2}\right)$, we have

$$\frac{1}{\hat{\omega}_\xi^4} = \frac{1}{\omega^4} - \frac{2}{\omega^6}\left(\hat{\omega}_\xi^2 - \omega^2\right) + O_p\left(n^{-1}\right) \ , \qquad \frac{1}{\hat{\omega}_\xi^6} = \frac{1}{\omega^6} + O_p\left(n^{-1/2}\right) \ . \tag{13}$$

By replacing (13) in (12) it yields

$$\frac{1}{\hat{\omega}^4} = \frac{1}{\omega^4} - \frac{2}{\omega^6}\left(\hat{\omega}_\xi^2 - \omega^2\right) + \frac{4}{\omega^6}\left(\hat{\xi} - \xi\right) m_1 + O_p\left(n^{-1}\right) \ .$$

Let $\hat{m}_r = (1/n)\sum\limits_{i=1}^n \left(X_i - \hat{\xi}\right)^r$, we get

$$\hat{m}_4 = m_4 - 4\left(\hat{\xi} - \xi\right) m_3 + O_p\left(n^{-1}\right) \ ,$$

consequently

$$\hat{\kappa} = \frac{\hat{m}_4}{\hat{w}^4} = \frac{m_4}{\omega^4} - \frac{2m_4}{\omega^6}\left(\hat{\omega}_\xi^2 - \omega^2\right) + \frac{4}{\omega}\left(\frac{m_4 m_1}{\omega^5} - \frac{m_3}{\omega^3}\right)\left(\hat{\xi} - \xi\right) + O_p\left(n^{-1}\right) \ .$$

Let $\mu_r = E(m_r)$; since the quantities m_r, for $r = 1, \ldots, 4$ are sample means, we have $m_r - \mu_r = O_p\left(n^{-1/2}\right)$. A Taylor expansion of $\hat{\kappa}$ with respect to the m_r's around their means yields

$$\hat{\kappa} = \frac{\mu_4}{\omega^4} + \frac{1}{\omega^4}\left(m_4 - \mu_4\right) - 2\frac{\mu_4}{\omega^6}\left(\hat{\omega}_\xi^2 - \omega^2\right) + \frac{4}{\omega}\left(\frac{\mu_4\mu_1}{\omega^5} - \frac{\mu_3}{\omega^3}\right)\left(\hat{\xi} - \xi\right) + O_p\left(n^{-1}\right) \ .$$

If the distribution has the density (2), $\mu_4/\omega^4 = 3$. Hence

$$\hat{\kappa} = 3 + \frac{1}{\omega^4}\left(m_4 - \mu_4\right) - \frac{6}{\omega^2}\left(\hat{\omega}_\xi^2 - \omega^2\right) + 4\eta\frac{1}{\omega}\left(\hat{\xi} - \xi\right) + O_p\left(n^{-1}\right).$$

Since $E\left(\hat{\xi}\right) = \xi + O\left(n^{-1}\right)$ and $\hat{\omega}_\xi^2$ and m_4 are unbiased, $E\left(\hat{\kappa}\right) = 3 + O\left(n^{-1}\right)$. The variance of $\hat{\kappa}$ is

$$\begin{aligned}\mathrm{var}\left(\hat{\kappa}\right) = {} & \mathrm{var}\left(\frac{m_4}{\omega^4}\right) + 36\mathrm{var}\left(\frac{\hat{\omega}_\xi^2}{\omega^2}\right) - 12\mathrm{cov}\left(\frac{m_4}{\omega^4}, \frac{\hat{\omega}_\xi^2}{\omega^2}\right) + \frac{16\eta^2}{\omega^2}\mathrm{var}\left(\hat{\xi}\right) \\ & + \frac{8\eta}{\omega^5}\mathrm{cov}\left(\hat{\xi}, m_4\right) - \frac{48\eta}{\omega^3}\mathrm{cov}\left(\hat{\xi}, \hat{\omega}_\xi^2\right) + o\left(n^{-1}\right) \ .\end{aligned}$$

Under (2), because of the invariance property of Proposition 2.2, $\hat{\omega}_\xi^2/\omega^2$ and m_4/ω^4 have the same distribution of the sample means of $(Y - \xi)^2$ and $(Y - \xi)^4$ where $Y \sim N(\xi, 1)$. Therefore we have $\mathrm{var}\left(\hat{\omega}_\xi^2/\omega^2\right) = 2n^{-1}$, $\mathrm{var}\left(m_4/\omega^4\right) = 96n^{-1}$ and $\mathrm{cov}\left(\hat{\omega}_\xi^2/\omega^2, m_4/\omega^4\right) = 12n^{-1}$, which yields (4).

Finally, since $n^{1/2}\left(\hat{\kappa} - 3\right)$ is a function of asymptotically normal statistics, it is itself asymptotically normally distributed.

Acknowledgements

Research partially supported by the Ministero dell'Istruzione, dell'Universitá e della Ricerca with funds for Progetti di Ricerca di Interesse Nazionale 2008, prot. 2008AHWTJ4-003.
The research of the first author greatly benefited by stimulating discussion with Nicola Maria Rinaldo Loperfido.

References

Arellano-Valle, R. B., del Pino, G. E., 2004. From symmetric to asymmetric distributions: a unified approach, in Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality, Genton, M. G., Editor, Chapman & Hall CRC, Boca Raton, FL.

Arellano-Valle, R. B., Branco, M. D., Genton M. G., 2006. A unified view on skewed distributions arising from selections. The Canadian Journal of Statistics, 34, 581-601.

Azzalini, A., 1985. A class of distributions which includes the Normal ones. Scandinavian Journal of Statistics, 12, 171-78.

Azzalini, A., 1986. Further results on a class of distributions which includes the Normal ones. Statistica, 46, 199-208.

Azzalini, A, 2005. The skew-normal distributions and related multivariate families. With discussion by Marc G. Genton and a rejoinder by the author. Scandinavian Journal of Statistics, 32, 159-200.

Azzalini, A., Capitanio, A, 2003. Distribution generated by a perturbation of symmetry with emphasis on a multivariate skew t distribution. Journal of the Royal Statistical Society: Series B, 65, 367-89.

Azzalini, A., Genton, M. G., Scarpa, B., 2010. Invariance-based estimating equations for skew-symmetric distributions. Metron, Special issue on: Skew-symmetric and flexible distributions, 68, 275-98.

Branco, M.D., Dey, D.K., 2001. A general class of skew-elliptical distributions. Journal of Multivariate Analysis, 79, 99-113.

DiCiccio, T.J., Monti, A.C., 2004. Inferential aspects of the skew-exponential power distribution. Journal of the American Statistical Association, 99, 439-450.

Efron, B, Tibshirani, R. J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Gastwirth, J.L., 1971. On the Sign Test for Symmetry. Journal of the American Statistical Association, 66, 821-823.

Genton, M. G., 2004. Skew-symmetric and generalized skew-elliptical distributions, in Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality, Genton, M. G., Editor, Chapman & Hall / CRC, Boca Raton, FL.

Genton, M. G., Loperfido N. M. R., 2005. Generalized Skew-Elliptical Distributions and Their Quadratic Forms, Annals of the Institute of Statistical Mathematics, 57, 389-401.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel W. A., 1986. Robust Statistics: The Approach based on the Influence Functions, J. Wiley, New York.

Ley, C., Paindaveine, D., 2010a. On the singularity of multivariate skew-symmetric models. Journal of Multivariate Analysis, 101, 1434-1444.

Ley, C., Paindaveine, D., 2010b. On Fisher information matrices and profile log-likelihood functions in generalized skew-elliptical models. Metron, Special Issue on: Skew-symmetric and flexible distributions, 68, 235-250.

Loperfido, N. M. R., 2004. Generalized Skew-Normal distributions, in Skew-Elliptical distributions and their applications, Genton M. G., Editor, Chapman & Hall / CRC, Boca Raton, FL., 65-80.

Ma, J., Genton, M. G., Tsiatis, A. A., 2005. Locally Efficient Semiparametric Estimators for Generalized Skew-Elliptical Distributions. Journal of the American Statistical Association, 100, 980-989.

Ma, J., Hart, J. D., 2007. Constrained Local Likelihood Estimators for Semiparametric Skew-normal Distributions. Biometrika, 87, 823-836.

Maronna, R.A., Martin, R.D., Yohai, V.J., 2006. Robust Statistic, Theory and Methods. J. Wiley, New York.

Nadarajah, S., Kotz, S., 2003. Skewed distributions generated by the normal kernel. Statistics & Probability Letters, 65, 269-277.

Pewsey, A., 2004. Some observations on a simple means of generating skew distributions. International Conference on Distribution Theory, Order Statistics and Inference in Honor of Barry C. Arnold, June 16-18, 2004.

Roberts, C., 1966. A correlation model useful in the study of twins. Journal of the American Statistical Association, 61, 1184-1190.

Rohatgi, V. K., 1976. An Introduction to Probability Theory and Mathematical Statistics, J. Wiley, New York.

Serfling, R., 1980. Approximation Theorems of Mathematical Statistics. J. Wiley, New York.

Thode, H. C., 2002. Testing for normality. Marcel Dekker, New York.

Wang, J., Boyer, J., Genton, M. G., 2004. A skew-symmetric representation of multivariate distributions. Statistica Sinica, 14, 1259-1270.