Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Default Bayesian goodness-of-fit tests for the skew-normal model

S. Cabras [a] & M. E. Castellanos [a]

[a] Dipartimento di Matematica e Informatica, Università Degli Studi di Cagliari, Cagliari, Italy

[b] Departamento de Estadística e Invesgtigación Operativa, Universidad Rey Juan Carlos, Móstoles, Spain

Version of record first published: 16 Dec 2008.

To cite this article: S. Cabras & M. E. Castellanos (2009): Default Bayesian goodness-of-fit tests for the skew-normal model, Journal of Applied Statistics, 36:2, 223-232

To link to this article: http://dx.doi.org/10.1080/02664760802443988

Journal of Applied Statistics
Vol. 36, No. 2, February 2009, 223–232

Default Bayesian goodness-of-fit tests for the skew-normal model

S. Cabras[a]* and M.E. Castellanos[b]

[a] *Dipartimento di Matematica e Informatica, Università Degli Studi di Cagliari, Cagliari, Italy;* [b] *Departamento de Estadística e Invesgtigación Operativa, Universidad Rey Juan Carlos, Móstoles, Spain*

(*Received 9 April 2008; final version received 11 August 2008*)

In this paper we propose a series of goodness-of-fit tests for the family of skew-normal models when all parameters are unknown. As the null distributions of the considered test statistics depend only on asymmetry parameter, we used a default and proper prior on skewness parameter leading to the prior predictive p-value advocated by G. Box. Goodness-of-fit tests, here proposed, depend only on sample size and exhibit full agreement between nominal and actual size. They also have good power against local alternative models which also account for asymmetry in the data.

Keywords: EDF test; model checking; prior predictive distribution; power; p-values; size of test

1. Introduction

The class of skew-normal densities ($\mathcal{SN}$ hereafter) has appeared independently several times in statistical literature as in [26,28]. The present name is given by Azzalini [3] and it has been generalized to the multivariate case in [4,6]. In this paper we consider only the univariate case.

One of the main features of the $\mathcal{SN}$ distribution is the ability to model the skewness through a shape parameter, λ. This model exhibits remarkable properties in terms of mathematical tractability and it has been applied, as a reference model, in several scientific disciplines as, for example, epidemiology, geology and economics (i.e., [7,10,18,19]).

The use of a parametric model has to be supported by an adequate goodness-of-fit (GOF hereafter) on the observed data. The model checking problem, addressed in this paper, is a preliminary analysis in that if data are compatible with the assumed model, then the full (and difficult) process of model elaboration and model selection (or averaging) can be avoided. Model selection is not addressed here and we refer to [8,9,25] for a comprehensive discussion on the relation between model checking and model selection.

*Corresponding author. Email: s.cabras@unica.it.es

ISSN 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/02664760802443988
http://www.informaworld.com

Works on GOF for the $\mathcal{SN}$ model can be found in [29] where normality is tested within the $\mathcal{SN}$ class by showing that the coefficient of skewness is locally the most powerful location and scale invariant statistic; Pewsey [27] used the same test statistics for testing departures from the half-normal distribution; Gupta and Chen [20] studied the Kolmogorov–Smirnov test to assess GOF of the $\mathcal{SN}$ model assuming all parameters known; where instead Mateu–Figueras et al. [24] derived the distribution of several test statistics when all unknown parameters are estimated with the maximum likelihood estimators (MLE), this work reports tables with critical values of GOF test for fixed values of λ and different sample sizes, n, obtaining an agreement between nominal and actual size of the test only for large $n(n > 400)$. In a recent work, Dalla Valle [15] proposed a test based on the Anderson–Darling test statistic (AD statistic, in the sequel) to analyse the fit of the $\mathcal{SN}$ model. To avoid the dependence of the approximated quantiles on the λ parameter in [15], a polynomial regression is considered, on the quantiles of the AD statistic as function of the skewness parameter. Unfortunately, in this latter work, there is a lack of agreement between nominal and actual size of the test, at least, for the analysed sample sizes ($n = 50$ and $n = 100$, see [15, Tables 4 and 5]). From this, it seems that in current literature a suitable GOF procedure for $\mathcal{SN}$ is not available, in terms of test size and power, as also required in the discussion section of [2].

In this paper we follow, instead, the Bayesian methodology proposed by Box [12] and extensively discussed in [8,9]. We approximate by Monte Carlo simulations the predictive prior distribution under the $\mathcal{SN}$ model, with null distribution, for some empirical distribution function test statistics (EDF hereafter). This is achieved by integrating out λ with respect to the Jeffreys' prior distribution discussed in [22]. In this way we can assess the compatibility of whole $\mathcal{SN}$ class with the data, moreover, the null distribution does not depend on unknown parameters and it has full agreement between nominal and actual test size for every $n \geq 3$. The full agreement is assessed by the fact that the sample null distribution of the p-value is uniform. The key point is that marginal prior predictive distributions of the used EDF test statistics: (i) do not depend on location and scale parameters when they are efficiently estimated in the sense of [16] and (ii) the Jeffreys' prior, in this case, is a proper distribution.

The rest of the paper is organized as follows: Section 2 reviews the $\mathcal{SN}$ model, Section 3 discusses the EDF statistics and proposes a simple algorithm to derive their respective marginal predictive distributions. In Section 4 presents a power study against local alternative models. To illustrate how to use our procedure to check the GOF of the $\mathcal{SN}$ model two applications are included in Section 5, while conclusions are summarized in Section 6.

2. The skew-normal model

We say that $X \sim \mathcal{SN}$ if X has density

$$f(x \mid \xi, \eta, \lambda) = \frac{2}{\eta}\phi\left(\frac{x-\xi}{\eta}\right)\Phi\left(\lambda\frac{x-\xi}{\eta}\right), \quad -\infty < x < \infty,$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, density and cumulative distribution function (cdf) of the standard normal distribution. Parameters $-\infty < \xi < \infty$ and $0 < \eta < \infty$ are, respectively, location and scale and $-\infty < \lambda < \infty$ is the shape parameter. When $\lambda = 0$, the $\mathcal{SN}$ corresponds to the normal model, while when $\lambda \to \pm\infty$ the distribution tends to the half-normal model. It is interesting to note that the skewness, γ, varies in the interval $(-0.995, 0.995)$ (see [27]), and this may cause incompatibility with data that present larger skewness.

MLE of $\mathcal{SN}$ parameters present the following problems: (i) MLE of λ, $\hat{\lambda}$, may be infinite with high probability for small n and large λ (see [22]); (ii) the Fisher information matrix is singular at $\lambda = 0$ (see [3]); (iii) the profile-likelihood function for λ may have different stationary points, with one at $\lambda = 0$ independently of the observed sample [27].

In [4,22,27] they analysed different estimation techniques in order to mitigate or avoid the mentioned problems. In particular, problem (ii) can be avoided by working with the central parametrization [3], denoted by $\mathcal{SN}_{\text{CP}}(\mu, \sigma, \gamma)$, where μ, σ and γ are mean, standard deviation and skewness, respectively (see [27], Formula (1)). Problem (iii) is important for small n and apparently it has no straightforward solution in a frequentist framework (see [4, Section 6.3]). Liseo and Loperfido [22] proposed to use the Jeffrey's and reference prior [11] on λ (or equivalently on γ) to mitigate the irregularities of the likelihood function. As explained the following, in our setup we need only the scalar prior on λ. We use the Jeffrey's or reference prior reported in [22, Section 2] and the fact that it is a proper prior distribution.

3. GOF tests for the skew-normal model

In order to construct a GOF test we need three elements: (1) a diagnostic test statistic, T, to quantify incompatibility with the data, using the convention that large observed values, t_{obs}, indicate incompatibility; (2) a completely specified distribution for T, $h(t)$, under the null model; (3) a way to measure the conflict between t_{obs} and $h(t)$ as the popular tail area p-value used in this work.

3.1 *Definitions of EDF statistics*

EDF statistics measure the difference between the empirical distribution, F_n, and a theoretical one, F, which in our case must be some element of the $\mathcal{SN}$ class. We fix F using: MLE of γ, $\hat{\gamma}$, sample mean, $\tilde{\mu}$, and sample standard deviation, $\tilde{\sigma}$, for μ and σ, respectively. Therefore, the theoretical distribution is

$$F \equiv \mathcal{SN}_{\text{CP}}(\tilde{\mu}, \tilde{\sigma}, \hat{\gamma}).$$

Note that our F differs from the one used in [24] only in the estimation of μ and σ. As commented in [4], even using the central parametrization for moderate sample sizes, there are still problems with the MLE estimator; in fact, the estimation of λ can occur in the frontier of λ support, that is, $\hat{\lambda} \to \pm\infty$ ($\hat{\gamma} \to \pm 0.99527$). For the samples in which this occurs we adopt the strategy mentioned in [4] when the maximization of λ is diverging, we choose the smallest $\hat{\lambda}$ whose likelihood is not significantly smaller than the likelihood with $\lambda = \infty$ according to a likelihood-ratio test.

The EDF statistics we use are discussed in [30]. They are the Kolmogorov–Smirnov, D, and the Kuiper, V,

$$D^+ = \sup_x \{F_n(x) - F(x)\}, \quad D^- = \sup_x \{F(x) - F_n(x)\},$$
$$D = \max\{D^+, D^-\}, \quad V = D^+ + D^-,$$

Cramér-von Mises, W^2, and Watson's U^2 statistics is given by

$$W^2 = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 \, \mathrm{d}F(x),$$
$$U^2 = n \int_{-\infty}^{+\infty} \left(F_n(x) - F(x) - \int_{-\infty}^{+\infty} (F_n(x) - F(x)) \, \mathrm{d}F(x) \right)^2 \mathrm{d}F(x).$$

Their corresponding sample version is obtained using the following formulas:

$$D^+ = \max_i \left(\frac{i}{n} - z_i\right), \quad D^- = \max_i \left(z_i - \frac{i-1}{n}\right), \quad D = \max\{D^+, D^-\},$$
$$V = D^+ + D^-, \quad W^2 = \sum_{i=1}^{n} \left(z_i - \frac{2i-1}{2n}\right)^2 + \frac{1}{12n}, \quad U^2 = W^2 - n\left(\bar{z} - \frac{1}{2}\right)^2, \tag{1}$$

where $z_i = F(x_{(i)})$, $\bar{z} = \sum_{i=1}^{n} z_i/n$ and $x_{(i)}$ denotes the i order statistics ($x_{(i)} < x_{(i+1)}$, $i = 1, \dots, n$).

As μ and σ are unknown we replace them with efficient estimators in the sense of [16] such as $\tilde{\mu}$ and $\tilde{\sigma}$. As noted in [30, Section 4.3.2], when unknown parameters are location or scale, and if these are estimated by appropriate methods, the distribution of EDF statistics will not depend on their true values.

However, $h(t)$ still depends on γ and in the frequentist paradigm $h(t)$ it is usually obtained by plugging in a point estimator of unknown parameters (i.e., $\hat{\gamma}$) obtaining the corresponding plug-in p-value, p_{plug}. This type of p-value is known to be conservative in small samples because of the double use of the data: first to estimate $h(t)$ and second, to calculate the p-value using again the data through t_{obs} (see [8,9]), where p_{plug} is used in [24], and because of the double use of the data, it results in less power than the one proposed here.

Our Bayesian approach avoids this issue as the null distribution $h(t)$ is the marginal prior predictive distribution of T obtained by integrating γ with respect to the Jeffreys' prior. This procedure is not feasible, in general, as the Jeffreys' prior is usually improper.

3.2 *Bayesian checking of the $\mathcal{SN}$ model*

We use the marginal prior predictive distribution of T by averaging γ in the conditional distribution $h(t \mid \gamma)$ with respect to the prior $\pi(\gamma)$:

$$m(t) = \int h(t \mid \gamma)\pi(\gamma)\,\mathrm{d}\gamma, \tag{2}$$

where $\pi(\gamma)$ is

$$\pi(\gamma) \propto \pi(\lambda)\left|\frac{\mathrm{d}\lambda}{\mathrm{d}\gamma}\right|, \quad \pi(\lambda) = K^{-1}\sqrt{\int_{-\infty}^{\infty} 2z^2\phi(z)\frac{\phi^2(\lambda z)}{\Phi(\lambda z)}\,\mathrm{d}z}, \tag{3}$$

and K is obtained by using adaptive quadrature integration for the integrals with respect to z and λ. Using the numerical approximation of prior cdf we can easily simulate from $\pi(\lambda)$. Note that it is sufficient to approximate prior (3) only once and distribution (2) just once for a given n. This approach is more general than those in [24] where one has to know γ and look at the corresponding $h(t \mid \gamma)$. Prior $\pi(\lambda)$ plays an important rule in this model checking procedure and it shows useful properties for GOF: (i) $m(t)$ always exists; (ii) it is invariant under reparametrization allowing to switch easily between central and direct parametrization because $\pi(\gamma)$ is obtained by applying the corresponding transformation rule (see [27, Formula (1)]) on simulated values of λ. To approximate $h(t)$ for a particular n we used $M = 10^6$ draws from $m(t)$ using the following 3 steps algorithm:

Step 1 Draw $\gamma^{(1,\dots,m,\dots,M)} \sim \pi(\gamma)$, for $M = 10^6$;
Step 2 for each $\gamma^{(m)}$ generate a random sample of size n from the $\mathcal{SN}$ model with $\mu = 0, \sigma = 1$ ($h(t)$ does not depend on μ,σ) and $\gamma = \gamma^{(m)}$;
Step 3 for the mth sample calculate the EDF statistic on the random sample.

The p-values are uniformly distributed for every $n \geq 3$ as showed in Figure 1.

Table 1 provides critical quantiles of the prior predictive distribution (2) for each EDF statistics and different n. We can see that each $m(t)$ has an asymptotical limiting distribution for $n \to \infty$ whose quantiles are approximately those of the last right column ($n = 500$).

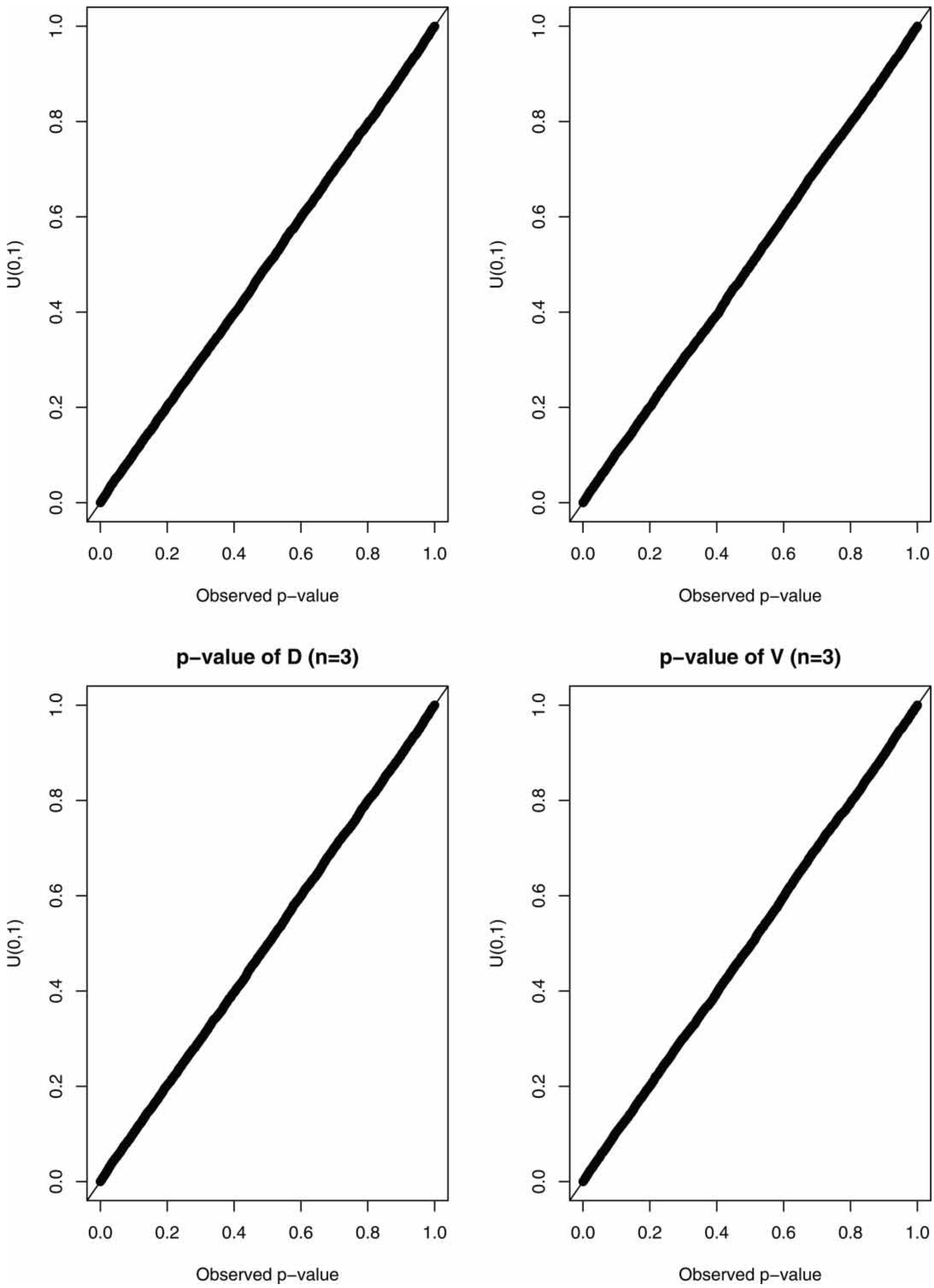

Figure 1. Quantile–Quantile-plot of p-value distribution (1000 samples) under the null model $\mathcal{SN}(x \mid \lambda)$, $\lambda \sim \pi(\lambda)$ against $U(0, 1)$ for $n = 3$. and different test statistics.

Table 1. Critical quantiles of $m(t)$ for EDF test statistics.

	Sample size (n)										
Probability	5	10	20	30	50	100	150	200	300	400	500
						W^2					
0.100	0.073	0.079	0.087	0.092	0.095	0.095	0.095	0.096	0.096	0.096	0.096
0.050	0.089	0.098	0.106	0.113	0.118	0.120	0.120	0.121	0.120	0.121	0.120
0.025	0.105	0.118	0.128	0.134	0.142	0.147	0.147	0.147	0.147	0.148	0.147
0.010	0.128	0.148	0.161	0.166	0.176	0.187	0.186	0.187	0.186	0.186	0.183
						U^2					
0.100	0.073	0.079	0.086	0.091	0.094	0.095	0.095	0.096	0.095	0.096	0.095
0.050	0.088	0.098	0.106	0.112	0.117	0.119	0.119	0.120	0.120	0.120	0.119
0.025	0.104	0.118	0.128	0.133	0.141	0.146	0.146	0.146	0.146	0.147	0.146
0.010	0.128	0.147	0.161	0.166	0.175	0.185	0.185	0.186	0.185	0.185	0.181
						$\sqrt{n}D$					
0.100	0.608	0.679	0.723	0.745	0.762	0.775	0.781	0.783	0.787	0.790	0.789
0.050	0.658	0.734	0.784	0.812	0.831	0.847	0.850	0.858	0.857	0.863	0.860
0.025	0.702	0.785	0.845	0.870	0.892	0.915	0.922	0.926	0.923	0.930	0.930
0.010	0.753	0.853	0.915	0.945	0.971	1.000	1.010	1.013	1.021	1.018	1.015
						$\sqrt{n}V$					
0.100	1.143	1.228	1.295	1.335	1.362	1.383	1.391	1.398	1.403	1.406	1.404
0.050	1.232	1.332	1.401	1.445	1.480	1.504	1.512	1.522	1.525	1.530	1.530
0.025	1.310	1.433	1.510	1.548	1.592	1.624	1.633	1.641	1.641	1.648	1.642
0.010	1.426	1.554	1.636	1.674	1.722	1.776	1.789	1.787	1.797	1.793	1.792

4. Power study under some alternative models

We approximate the power of GOF tests for different EDF statistics against alternative distributions of the $\mathcal{SN}$. All the considered distributions account for asymmetry in the data. Figure 2 reports the power for a test of size 95% against the following models:

Lognormal: Lognormal($\mu = 0, \sigma$) with location μ and shape parameter σ (see [21, p. 208]). We consider σ equals 0.3, as small values of σ makes Lognormal distribution similar to $\mathcal{SN}$.

Gamma: Gamma(α, β) with shape α and rate β. Large values of α makes this model similar to the $\mathcal{SN}$. For seeking of comparison with [24] we consider these special cases: exponential distribution and the χ^2 distribution with four degrees of freedom.

Skew-T: ST($\mu = 0, \sigma = 1, \gamma = 0.5, df$) where μ, σ and γ are location, scale and shape parameter respectively, and df indicates the degrees of freedom. This model, introduced by Azzalini and Capitanio [5], encompasses the $\mathcal{SN}$ through the df parameter: ST(μ, σ, γ, df) $\to \mathcal{SN}(\mu, \sigma, \gamma)$ as $df \to \infty$. We consider the Skew-T model with 1, 3 and 10 df.

Weibull: Weibull($c = 2, \alpha = 1$), where c and α are shape and scale parameters, respectively (see [22, p. 641]). Large values of c makes the Weibull model similar to the $\mathcal{SN}$.

Statistic W^2 is generally the most powerful, followed by the U^2. We also used the AD test statistic, which is not shown here (see [13]) as its power was unsatisfactory for moderate sample sizes ($n < 100$) in general, except for Weibull alternatives with small α. As expected, Figure 2 shows that the power reduces when alternative models are close to the $\mathcal{SN}$. This occurs for the $\mathcal{ST}$ with $df \geq 10$, the Gamma model with $\alpha \geq 4$, the Weibull model and for the Lognormal model. We can see that GOF tests proposed here are more powerful than those in [24]; in particular,

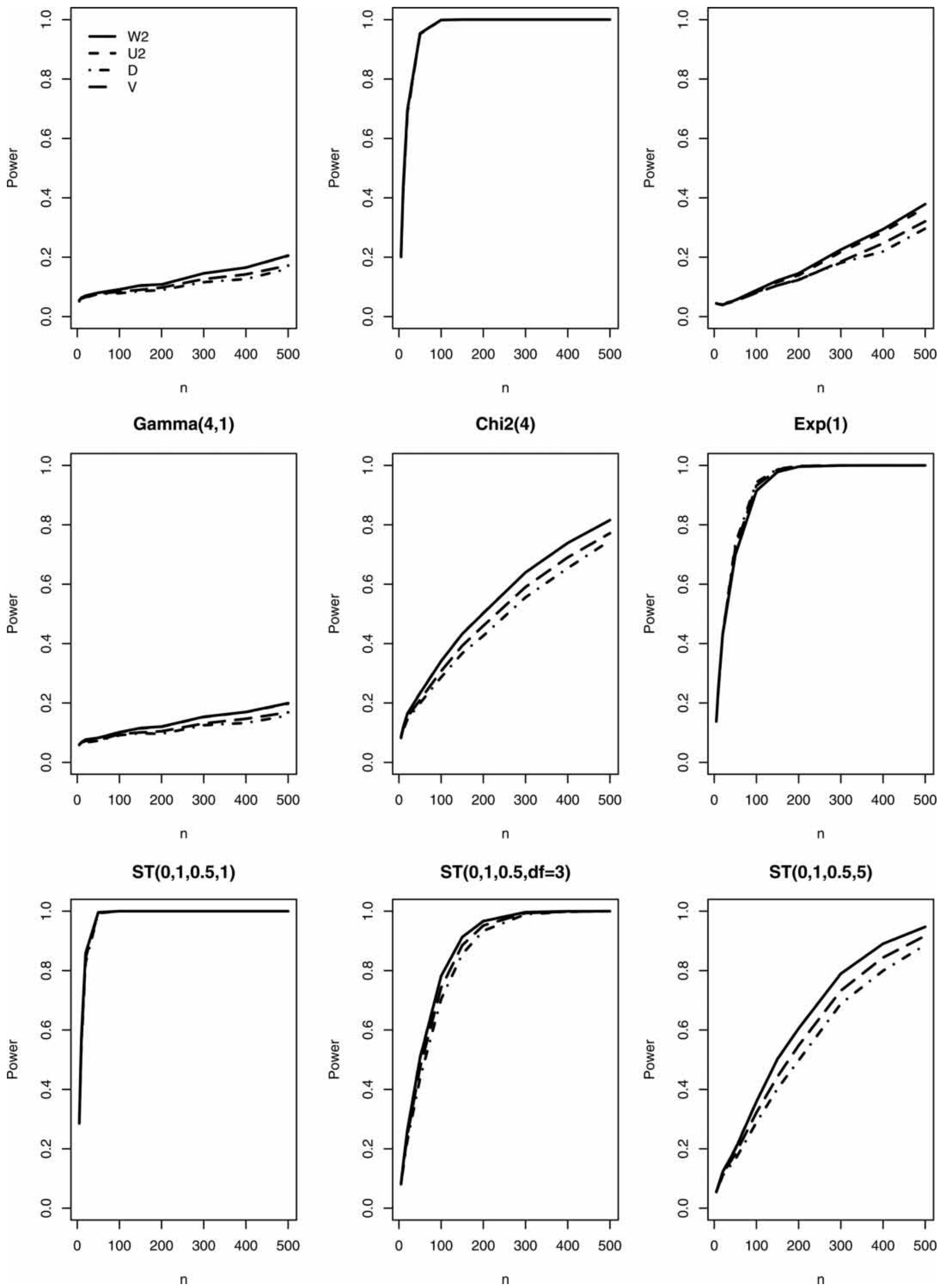

Figure 2. Power functions for the alternative models.

for $n \leq 20$. This power difference, generally, reduces asymptotically, in fact for $n = 500$ and Weibull(2, 1) the power is lower: 47% in [24] against 38% here. However, there are cases, such as χ_4^2, where our tests still have considerably more power even for $n = 500$: 75% against 52% in [24]. Also, our proposal is more powerful than the one in [15]. As appears in [15, Tables 6 and 7], our tests are more powerful for the Lognormal, Gamma and Weibull models for $n = 100$ and $n = 500$ (sample sizes there considered).

Table 2. Estimations and test statistic W^2 for AIS data.

Variables	Sex	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\gamma}$	W^2	Quantile 0.95	$\mathcal{SN}$ fitting
WT	Female	67.343	10.915	−0.215	0.034	0.120	Yes
	Male	82.524	12.406	0.339	0.049	0.120	Yes
RCC	Female	4.404	0.321	0.696	0.047	0.120	Yes
	Male	5.027	0.351	0.293	0.176	0.120	No

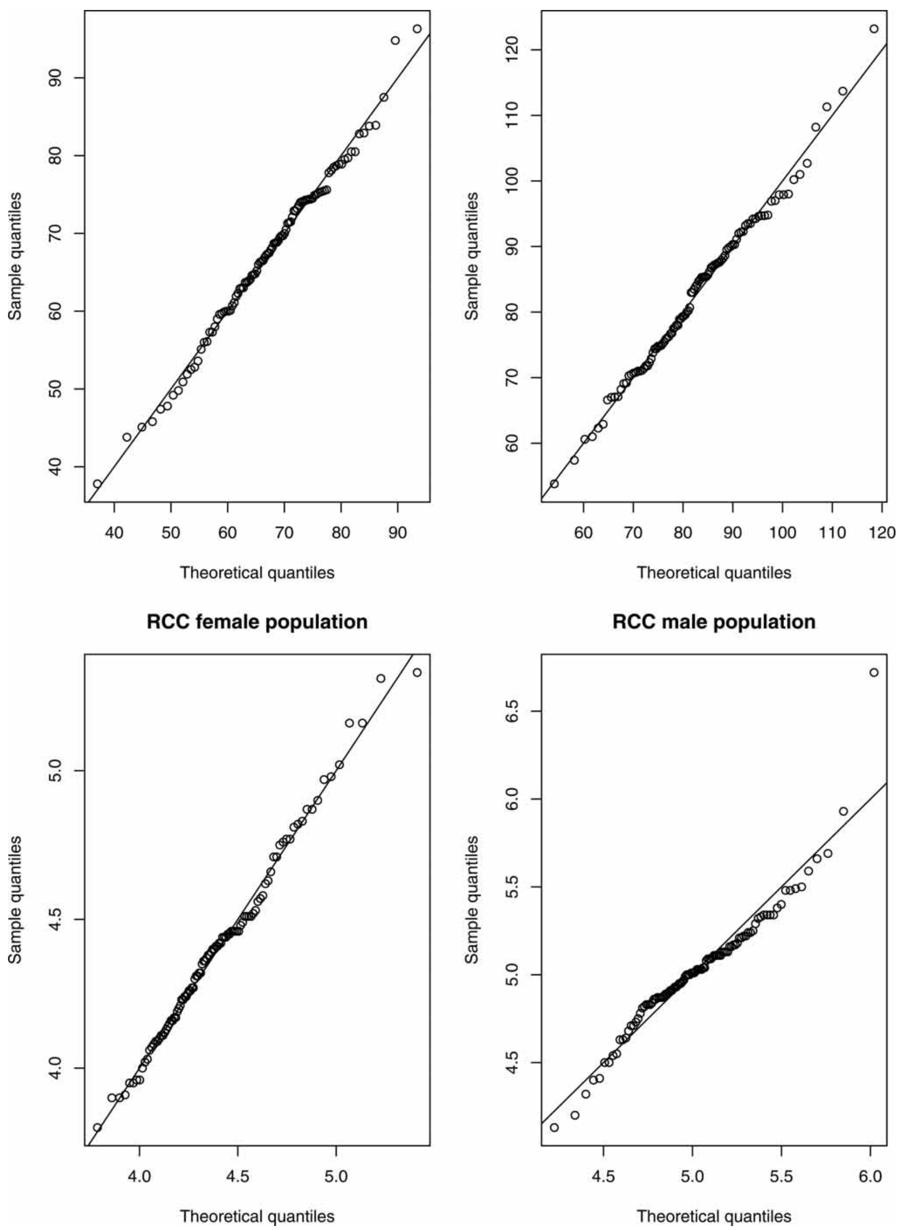

Figure 3. Quantile–Quantile-plot of sample female (left) and male (right) weights and RCC against the theoretical distribution F.

5. Application to real data

In this section we consider two datasets in order to show how tests work. The first consists of an artificial dataset, while the second is a real dataset where the $\mathcal{SN}$ has been used as a reference model.

First we consider the Frontier dataset used in [3] to show a case when MLE of λ does not exist. Using the strategy mentioned in Section 3.1, the adjusted MLE of λ is 8.14, while the moment estimators of μ and σ are 0.88 and 0.76, respectively. Substituting the parameters by these estimations and computing the statistics as mentioned in Equations (1), we get $W^2 = 0.026$, $U^2 = 0.025$, $\sqrt{n}D = 0.441$ and $\sqrt{n}V = 0.791$. All these values are well below the quantile 90% showing that the Frontier data are compatible with the $\mathcal{SN}$ model.

The second dataset is a subset of the Australian Institute of Sport (AIS) data first examined by Cook and Weisberg [14] within a framework of regression analysis. The AIS data consist of several biomedical measurements on 102 male and 100 female athletes. In [5,6] the $\mathcal{SN}$ is used to model some measurements. These data are also analysed in [1,17] using different skew distributions. In [15] the fit of the $\mathcal{SN}$ to several variables using a modification of the AD test is done. Here we consider two of the four variables analysed in this last paper, but fitting a separate $\mathcal{SN}$ to males and females. The two variables considered are the weight (WT) (in kg), and the number of red cells count (RCC). Table 2 shows parameter estimations for each variable and sex, the value of W^2 (the most powerful EDF statistic), 95% quantile for $n = 100$, and the conclusion about the fitting of the $\mathcal{SN}$. Table 2 shows the $\mathcal{SN}$ results to be compatible for all measurements but the RCC in males. In this case another model is needed. Quantile–Quantile-plots in Figure 3 informally confirm the results of the test. These results partially agree with those in [15, Table 10] where the sex factor has been ignored. In fact, the $\mathcal{SN}$ does not fit the RCC, while here it fits only for females. This also may be an evidence of sexual dimorphism as discussed in [23] and references therein.

6. Conclusions

We propose GOF tests that depend only on sample size. This procedure, based on $m(t)$ fully agrees with the Neymann hypothesis testing, as rejection regions are fixed prior to observing the data. The tests are based on the Jeffreys' prior Equation (3), which is used to eliminate the nuisance parameter γ present in the distribution of the test statistics. This may be viewed in contrast to the methodology in [15], where the dependence on γ is avoided using a post-processing of the quantiles of the null distribution of the test statistics. Goodness-of-fit tests used here as benchmarks are less powerful [15,24]. This is true, in general, for tests based on a double use of the data: parameter estimation, $\hat{\gamma}$, and calculus of p-value on observed t_{obs}. Another open field of research could be the assessing of the GOF for the Skew-T model. There, degrees of freedom is another nuisance parameter present in the null distribution of the test statistic.

Acknowledgements

This research has been supported by M.I.U.R. (Ministero dell'Istruzione, dell'Univesità e della Ricerca) of Italy and by the Spanish Ministry of Science and Technology, under Grant MTM2007-61554. Part of the work was done while the second author was visiting the Informatics and Mathematics Department at University of Cagliari (Italy).

References

[1] B. Arnold and R. Beaver, *The skew-Cauchy distribution*, Statist. Probab. Lett. 49 (2000), pp. 285–290.

[2] ———, *Skewed multivariate models related to hidden truncation and/or selective reporting*, Test 11 (2002), pp. 7–54.

[3] A. Azzalini, *A class of distribution which includes the normal ones*, Scand. J. Statist. 12 (1985), pp. 171–178.
[4] A. Azzalini and A. Capitanio, *Statistical applications of the multivariate skew normal distribution*, J. Roy. Statist. Soc. Ser. B 61 (1999), pp. 579–602.
[5] ———, *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution*, J. Roy. Statist. Soc. Ser. B 65 (2003), pp. 367–389.
[6] A. Azzalini and A. Dalla Valle, *The multivariate skew normal distribution*, Biometrika 83 (1996), pp. 715–726.
[7] L. Ball and N.G. Mankiw, *Relative price changes as aggregate supply shocks*, Quar. J. Econom. CX (1995), pp. 161–193.
[8] M.J. Bayarri and J. Berger, *Quantifying surprise in the data and model verification*, in *Bayesian Statistics 6*, M.J. Bayarri, J.O. Berger, J.M. Bernardo, A. Dawid, and A. Smith, eds., Oxford University Press, London, 1999, pp. 53–82.
[9] ———, *P-values for composite null models*, J. Amer. Statist. Assoc. 95 (2000), pp. 1127–1142.
[10] J.L. Bazán, M.D. Branco, and H. Bolfarine, *A skew item response model*, Bayesian Anal. 1 (2006), pp. 861–892.
[11] J. Berger and J.M. Bernardo *On the development of reference priors*, in *Bayesian Statistics 4*. J.O. Berger, J.M. Bernardo, A. Dawid, and A. Smith, eds., Oxford University Press, London, 1992, pp. 35–60.
[12] G. Box, *Sampling and Bayes inference in scientific modeling and robustness*, J. Roy. Statist. Soc. Ser. A 143 (1980), pp. 383–430.
[13] M.E. Castellanos and S. Cabras, *Testing the skew-normal model from a Bayesian perspective*, Tech. Rep. 2005-11, Rey Juan Carlos University, 2005. Available at http://bayes.escet.urjc.es/publicaciones.php.
[14] R. Cook and S. Weisberg, *An introduction to regression graphics*, John Wiley and Sons, New York, 1994.
[15] A. Dalla Valle, *A test for the hypothesis of skew-normality in a population*, J. Statist. Comput. Simul. 77 (2007), pp. 63–77.
[16] J. Durbin, *Weak convergence of the sample distribution function when parameters are estimated*, Ann. Statist. 1 (1973), pp. 279–290.
[17] M. Genton and N. Loperfido, *Generalized skew-elliptical distributions and their quadratic forms*, Ann. Inst. Statist. Math. 57 (2005), pp. 389–401.
[18] P. Ghosh, M.D. Branco, and H. Chakraborty, *Bivariate random effect model using skew-normal distribution with application to HIVRNA*, Statist. Med. 26 (2006), pp. 1255–1267.
[19] R.C. Gupta and N. Brown, *Reliability studies of the skew-normal distribution and its application to a strength-stress model*, Commun. Statist. Theory Methods 30 (2001), pp. 2427–2445.
[20] A. Gupta and T. Chen, *Goodness-of-fit tests for the skew-normal distribution*, Commun. Statist. Simulation Comput. 30 (2001), pp. 907–930.
[21] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, John Wiley & Sons, New York, 1994.
[22] B. Liseo and N. Loperfido, *A note on reference priors for the scalar skew-normal distribution*, J. Stat. Plann. Inference 136 (2006), pp. 373–389.
[23] E. Marini et al., *Sex differences in skinfold variability across human populations and during the life cycle*, Ann. Human Biol. 34(3) (2007), pp. 373–392.
[24] G. Mateu-Figueras, P. Puig, and A. Pewsey, *Goodness-of-fit tests for the skew- normal distribution when the parameters are estimated from the data*, Commun. Statist. Theory Methods 36 (2007), pp. 1735–1755.
[25] A. O'Hagan, *HSSS model criticism* (*with discussion*), in *Highly Structured Stochastic Systems*, P.J. Green, N.L. Hjort, and S.T. Richardson, eds., Oxford University Press, Oxford, UK, 2003, pp. 423–445.
[26] A. O'Hagan and T. Leonard, *Bayes estimation subject to uncertainty about parameter constraints*, Biometrika 63 (1976), pp. 201–202.
[27] A. Pewsey, *Problems of inference for Azzalini's skew-normal distribution*, J. Appl. Statist. 27 (2000), pp. 1184–1190.
[28] C. Roberts, *A correlation model useful in the study of twins*, J. Amer. Statist. Assoc. 61 (1966), pp. 1184–1190.
[29] A. Salvan, *A locally most powerful invariant tests of normality*, in *Atti della XXXIII Riunione Scientifica della Societ Italiana di Statistica 2*, Bari, Italy, 1986, pp. 173–179.
[30] M.A. Stephens, *Tests Based on EDF Statistics*, Chapter 4, in *Goodness-of-fit Techniques*, R.B. D'Agostino and M.A. Stephens, eds., Marcel Dekker, New York, USA, 1986, pp. 52–71.