POWER SKEW SYMMETRIC DISTRIBUTIONS: TESTS FOR SYMMETRY

R.N. Rattihalli[1], M. Raghunath[2]
[1]*Department of Statistics, Central University of Rajasthan*
[2]*Department of Statistics, SDSOS, NMIMS (Deemed to be) University*
Bandarsindri, Mumbai, INDIA
e-mail: [1]`rnr5@rediffmail.com`, [2]`raghummr@gmail.com`

Abstract

In this article a semi-parametric class of skew-symmetric distributions is considered. We call this class as Power-Skew-Symmetric ($\mathcal{PSS}$) distributions, being obtained by considering a positive power of a distribution function symmetric about 0. Based on U-statistics we develop two nonparametric tests for symmetry in the $\mathcal{PSS}$ class. Performances of the proposed tests are evaluated using efficacy and empirical power.
Keywords: efficiency, empirical power, semi-parametric class, skew-distributions, U-statistics

1 Introduction

In the literature parametric/nonparametric classes of skew-symmetric distributions have been generated by introducing an additional parameter to a class of symmetric distributions. Azzalini (1985) introduced a class of skew-symmetric distributions based on normal distribution and Gómez et. al (2006) have generalized this class by introducing one more addition parameter. Further extensions based on t, Lapalce, Cauchy, Uniform, Logistic distributions have been considered by Gupta et. al (2002) and their distributional properties have been studied by Nadarajah and Kotz (2003, 2006). Mudolkar and Hutson (2000) have proposed Epsilon-Skew normal family by using normal density and a skewness parameter ε.

Lehmann (1953) proposed a family of distributions

$$\mathbb{F}_F(x,\alpha) = \{F^{\alpha}(x), \alpha \in (0,\infty)\} \tag{1}$$

where F is a distribution function. In the context of testing the null hypothesis that F is the true distribution one may confine to the class (1) and the subclass of (1) with $\alpha \neq 1$ is referred to as the class of *Lehmann alternatives.* If F is absolutely continuous then the corresponding density function is

$$\varphi_F(x,\alpha) = \alpha f(x)\left\{F(x)\right\}^{\alpha-1}, x \in \mathbb{R}, \alpha > 0. \tag{2}$$

The class (1) is used for data analysis by considering F to be a specified parametric family (usually taken to be symmetric), for example Durrans (1992), Gupta and Gupta (2008), Pewsey et. al (2012).

In this article we consider

$$\{F^{\alpha}(x); F \text{ is symmetric about } 0\ , \alpha \in (0, \infty)\}\,. \tag{3}$$

called the *Power Skew Symmetric* family of distributions and is denoted by $\mathcal{PSS}$. Thus the family of all distribution functions which are positive powers of a continuous symmetric distribution function symmetric about 0. It is a semiparametric family.

It is to be noted that a member of $\mathcal{PSS}$ is symmetric if and only if $\alpha = 1$. Based on U-statistics theory, We propose two nonparametric tests for testing symmetry in $\mathcal{PSS}$ class.

In section 2 the class of $\mathcal{PSS}$ is defined and some of its properties, graphs of Distribution Function (DF), Probability Density Function (PDF) for certain members generated from some well know symmetric models are given. Two U-statistics type statistics for testing symmetry in this class ($\alpha = 1$) are proposed in section 3. Asymptotic null distributions of the proposed statistics are discussed in section 4. The efficacies of the tests are derived in section 5. In section 6, empirical powers of the proposed tests are computed for different subclasses of $\mathcal{PSS}$ generated from some well known symmetric models.

2 The Class of Power-Skew-Symmetric Distributions

In this section we define the class of *Power-Skew-Symmetric*($\mathcal{PSS}$) distributions and study some of its properties.

Let $F(t) = P(T \le t)$ be the distribution function of random variable (r.v.) Y and $F(t-) = P(Y < t)$. If $F(.)$ is continuous at t the $F(t) = F(t-)$. The distribution function F (or the r.r. Y) is said to be symmetric about 0 if $F(t) = 1 - F(t-), -\infty < t < \infty$. The class $\mathcal{PSS}$ is defined as,

$$\mathcal{PSS} = \{F^{\alpha}(x) : x \in \mathbb{R}, F \ \ is \ \ symmetric \ \ about \ \ 0\}\,. \tag{4}$$

In the following we show that $(F(.), \alpha)$ constitutes the parameter for the class $\mathcal{PSS}$. If $G(.) \in \mathcal{PSS}$ then $G(x) = F^{\alpha}(x)$ for some distribution function F symmetric about 0 and some $\alpha > 0$. To be precise $G(x)$ is $G_{F,\alpha}(x)$, but for notational simplicity, unless otherwise required, we write it as $G(x)$, The class $\mathcal{PSS}$ is a semi-parametric family and it can be extend by introducing the location and the scale parameters. In the following we shall show that $(F(.), \alpha)$ constitutes the parameter for the $\mathcal{PSS}$ class.

Lemma 1. $(F(.), \alpha)$ *constitutes the parameter for the class* $\mathcal{PSS}$

Proof. For if, $F_1^{\alpha_1}(x) = F_2^{\alpha_2}, \forall x$, then we have,

$$F_1(x) = F_2^{\frac{\alpha_2}{\alpha_1}}(x), \forall x. \tag{5}$$

We note that if F is symmetric about 0 then $F^{\alpha}(x)$is symmetric about 0 if and only if $\alpha = 0$. Hence as $F_1(.)$ and $F_2(.)$ are symmetric about 0, (4) holds if and only if $\alpha_1 = \alpha_2$, which in turn also implies $F_1(x) = F_2(x), \forall x$.
Hence the proof. □

In the following we give some of the properties of members of $\mathcal{PSS}$ class distributions. Let X be an r.v. with continuous distribution function (DF) $F^{\alpha}(x)$ (denoted by $X \sim F^{\alpha}(x)$).

P.1: If $X \sim F^{\alpha}(x)$ then $-X \sim 1 - F^{\alpha}(-x)$.

P.2: If $X \sim G(.) \in \mathcal{PSS}$, then $G(0) = (1/2)^{\alpha}$ for any $F(.)$ symmetric about 0.

P.3: Let $X \sim F^{\alpha}$(say) , $Y \sim F(.)$ and $G_{\alpha}(.)$ be the distribution function of X then

- **i)** The probability density function of X is given by $g_{\alpha}(x) = \alpha F^{\alpha-1} f(x)$. provided it exists.
- **ii)** The supports of $G_{\alpha}(.)$ and $F(.)$ are the same.
- **iii)** X is stochastically larger(sampler) than Y according as $\alpha \leq 1 (\alpha \geq 1)$.
- **iv)** The inverse function of the CDFs satisfy the relation,

$$G_{\alpha}^{-1}(u) = F^{-1}\left(u^{(1/\alpha)}\right), 0 < u < 1, \alpha > 0.$$

P.4: (Gupta and Gupta (2008)) If the $X_i \sim PSS(F, \alpha_i), i = 1, 2, ..., n$ are independent $X_{(n)} = max(X_1, X_2, ..., X_n) \sim PSS\left(F, \sum_{i=1}^{n} \alpha_i\right)$.

The graphs of DF and PDF of the $\mathcal{PSS}$ distributions generated from Cauchy, Laplace, Logistic, Normal and Uniform distributions for some values of α are given in the Appendix.

3 Proposed Classes of Tests

Let $X_1, X_2, ..., X_n$ be be independent identically distributed random variables with common DF $\in G \in \mathcal{PSS}$. The problem of interest is to test the hypothesis H_0: G is symmetric about 0 against the alternative H_1: G is not symmetric about 0, that is to test

$$H_0 : \alpha = 1 \text{ against } H_1 : \alpha \neq 1. \tag{6}$$

Here F is a nuisance parameter.

Motivated from Mehra et. al (1990) and Rattihalli and Raghunath (2012), we propose two U-test-statistics to test the above hypothesis. The kernel function depends on a constant to be chosen so as to maximize the efficacy of the test. This is possible as the efficacies of the tests do not depended upon nuisance parameter $F(.)$. The two U-test statistics are given by

$$T_a = \frac{\sum_{C_1} \psi_a(x_i, x_j)}{\binom{n}{2}} \tag{7}$$

$$S_b = \frac{\sum_{C_2} \psi_b(x_i, x_j, x_k)}{\binom{n}{3}} \tag{8}$$

where the summations C_1 and C_2 are respectively over the $\binom{n}{2}, \binom{n}{3}$ combinations of integers from $\{1, 2, .., n\}$ and

$$\psi_a(x_i, x_j) = \begin{cases} a & if \ \min\{x_i, x_j\} > 0 \\ 1(-1) & if \ x_i x_j < 0, x_i + x_j > (<)0 \\ -a & if \ \max\{x_i, x_j\} > 0 \\ 0 & otherwise, \end{cases} \tag{9}$$

$$\psi_b(x_i, x_j, x_k) = \begin{cases} b & if \ x_{(1)} > 0 \\ 1(-1) & if \ x_{(1)} < 0 < x_{(2)}(x_{(2)} < 0 < x_{(3)}), x_{(1)} + x_{(3)} > (<)0 \\ -b & if \ x_{(3)} < 0 \\ 0 & otherwise, \end{cases} \tag{10}$$

where $x_{(i)}$ is the r^{th} order statistic from a sub-sample of size 3.

A test rejects H_0 in favour of H_1 for the large absolute value of the corresponding test statistic.

4 Asymptotic null distribution of the proposed test statistics

Since the statistics T_a and S_b are one sample U-statistics, then from the theorem of Hoeffding (1948), we have the following theorem.

Theorem 1. *Let $\sigma_a^2 = Var\left[E_{H_0}(\psi_a(X_1, X_2)|X_1 = x_1)\right]$. Then under the H_0 $\sqrt{n}\left[T_a - E_{H_0}(T_a)\right]$ converges in distribution as $n \to \infty$ to $N(0, 4\sigma_a^2)$ r.v.*

Thus to obtain the asymptotic null distribution of T_a, it is enough to find $E\left[\psi_a(X_1, X_2)\right]$, $E_{H_0}\left[\psi_a(X_1, X_2)|X_1 = x_1)\right]$ and are obtained in the following.

$$\begin{aligned} E[T_a] &= E[\psi_a(X_1, X_2)] \\ &= aP[X_{(1)} > 0] + P[X_{(1)} < 0 < X_{(2)}, X_{(1)} + X_{(2)} > 0] \\ &\quad - P[X_{(1)} < 0 < X_{(2)}, X_{(1)} + X_{(2)} < 0] - ap[X_{(2)} < 0] \\ &= a\{P[E_1] - p[E_4]\} + P[E_2] - P[E_4] \end{aligned}$$

and the probabilities of the above events are,

$$\begin{aligned} P[E_1] &= \left(1 - 2^{-\alpha}\right)^2 \\ P[E_2] &= 2^{1-\alpha} - 2\alpha \int_0^{1/2} (1-u)^\alpha u^{\alpha-1} du \\ P[E_3] &= 2\alpha \int_0^{1/2} (1-u)^\alpha u^{\alpha-1} du - 2^{1-2\alpha} \\ P[E_4] &= 2^{-2\alpha}. \end{aligned}$$

It is to be noted that, the above probabilities do not depend on the underlying symmetric model $F(.)$.

Thus we get,

$$\nu_a(\alpha) = E[\psi_1(X_1, X_2)] = a(1 - 2^{1-\alpha}) + 2^{1-\alpha} + 2^{2\alpha-1} - 4\alpha \int_0^{1/2} (1-u)^\alpha u^{\alpha-1} du. \quad (11)$$

Under $H_0 : \alpha = 1$, we have,

$$E_{H_0}[T_a] = 0.$$

Further,

$$E_{H_0}(\psi_a(X_1, X_2)|X_1 = x_1) = \begin{cases} 2F(x_1) - \left(\frac{a+1}{2}\right) & when \quad x_1 \leq 0 \\ \left(\frac{a-3}{2}\right) + 2F(x_1) & when \quad x_1 \geq 0 \end{cases} \quad (12)$$

Hence the asymptotic variance $4\sigma_a^2$ is,

$$\begin{aligned} 4\sigma_a^2 &= 4Var[E_{H_0}(\psi_a(X_1, X_2)|X_1 = x_1)] \\ &= 4\left\{\int_0^\infty \left[\left(\frac{a-3}{2}\right) + 2F(x_1)\right]^2 dF(x_1) + \int_{-\infty}^0 \left[2F(x_1) - \left(\frac{a+1}{2}\right)\right]^2 dF(x_1)\right\}. \end{aligned} \quad (13)$$

Thus,

$$4\sigma_a^2 = \left(\frac{1}{3} + a^2\right). \quad (14)$$

Similarly the asymptotic distribution of T_b is given by,

Theorem 2. *Let $\sigma_b^2 = Var[E_{H_0}(\psi_b(X_1, X_2, X_3)|X_1 = x_1)]$. Then under H_0 $\sqrt{n}[S_b - E_{H_0}(S_b)]$ converges in distribution as $n \to \infty$ to $N(0, 9\sigma_b^2)$ r. v.*

The expectation and asymptotic variance of T_b are given by,

$$\begin{aligned} \nu_b(\alpha) = b\left[(1 - 2^{-\alpha})^3 - 2^{-3\alpha}\right] + 3\Big\{&2^{-\alpha}(1 - 2^{1-\alpha} + 2^{-2\alpha}) \\ &-\alpha \int_0^{1/2} ((1-u)^\alpha - 2u^\alpha)(1-u)^\alpha u^{\alpha-1} du\Big\}. \end{aligned} \quad (15)$$

Under $H_0 : \alpha = 1$, we have,

$$E_{H_0}[S_b] = 0.$$

It is easy to verify that,

$$E_{H_0}(\psi_b(X_1, X_2, X_3)|X_1 = x_1) = \begin{cases} 2F(x_1) - 2F^2(x_1) - \left(\frac{b+2}{4}\right) & when \quad x_1 \leq 0 \\ \left(\frac{b+2}{4}\right) + 2F^2(x_1) - 2F(x_1) & when \quad x_1 \geq 0 \end{cases} \quad (16)$$

The asymptotic variance $9\sigma_b^2$ is given by,

$$9\sigma_b^2 = 9\left(\frac{b^2}{16} + \frac{b}{12} + \frac{1}{20}\right). \quad (17)$$

In the next section we obtain the constants 'a' and 'b', so that the efficacies of the tests are maximal.

5 Efficacies of the proposed tests

Let $T = \{T_n\}$ be a sequence of test statistics for testing the hypothesis that $H_0 : \theta = \theta_0$ against the suitable alternative. Let $E(T_n) = \mu_n(\theta)$ and $Var(T_n) = \sigma_n^2(\theta)$. Under certain regularity conditions (see Randles and Wolfe (1979)) the efficacy of T is given by,

$$eff[T] = \lim_{n\to\infty} \frac{\mu'_n(0)}{\sqrt{n}\sigma_n(0)}. \tag{18}$$

By considering $T_n = T_a$, $\mu_n(\theta) = \nu_a(\alpha)$, where $\nu_a(\alpha)$ is given in (10) and

$$\nu_a'(\alpha)|_{\alpha=1} = (2-a)\ln(1/2) - (3/2) - 4\int_0^{1/2}(1-u)\ln(u(1-u))du,$$

The efficacy of T_a is,

$$eff^2[T_a] = \frac{3\left[(2-a)\ln(1/2) - (3/2) - 4I_1\right]^2}{(3a^2+1)} \tag{19}$$

where $I_1 = \int_0^{1/2}(1-u)\ln(u(1-u))du$.

The optimal value a^* of a is obtained by solving $(d/da)eff^2(T_a) = 0$ and verifying $(d^2/(da^2))eff^2(T_a) < 0$ at the solution obtained. Here the value obtained is,

$$a^* = \frac{2\ln(1/2)}{24I_1 + 9 - 12\ln(1/2)}, \tag{20}$$

where I_1 is define above and by numerical integration it can be shown that $I_1 = -0.7983$. Hence from (19) we have,

$$a^* = 0.7528.$$

Thus the efficacy of T_{a^*} is,

$$eff^2[T_{a^*}] = 0.763.$$

Similarly, the efficacy of the test T_b is,

$$eff^2[S_b] = \frac{15\left[8I_2 - (2b+1)3\ln(1/2)\right]^2}{(60b^2+80b+48)} \tag{21}$$

where $I_2 = \int_0^{1/2}(1-u)[(3u-1) + 2(2u-1)\ln(1-u) + (5u-1)\ln(u)]du = 0.1122$ (by numerical integration).

The optimal value of b^* of b is,

$$b^* = \frac{8I_2 - \ln(1/2)}{2\ln(1/2)}. \tag{22}$$

Substituting the value of I_2 in (21), we get,

$$b^* = 0.1476.$$

and

$$eff^2[S_{b^*}] = 0.7912.$$

The efficacy of the test S_{b^*} is more than that of T_{a^*}.

6 Monte-Carlo Simulation

In this section we carry out the empirical power study to assess the performances of the proposed test statistics T_{a^*} and S_{b^*}. For simulation study samples were drawn from $G(.)$, when F corresponds to Cauchy, Laplace, logistic, normal, triangular and uniform.

Under H_0 test statistics T_{a^*} and S_{b^*} are asymptotically normal with mean 0 and variances given in (13) and (16) respectively. Then corresponding to the size γ, the criteria for rejection are

a) to reject H_0 if $|T_{a^*}| \geq \frac{Z_{(\gamma/2)} 2\sigma_{a^*}}{\sqrt{n}}$

b) to reject H_0 if $|S_{b^*}| \geq \frac{Z_{(\gamma/2)} 3\sigma_{b^*}}{\sqrt{n}}$

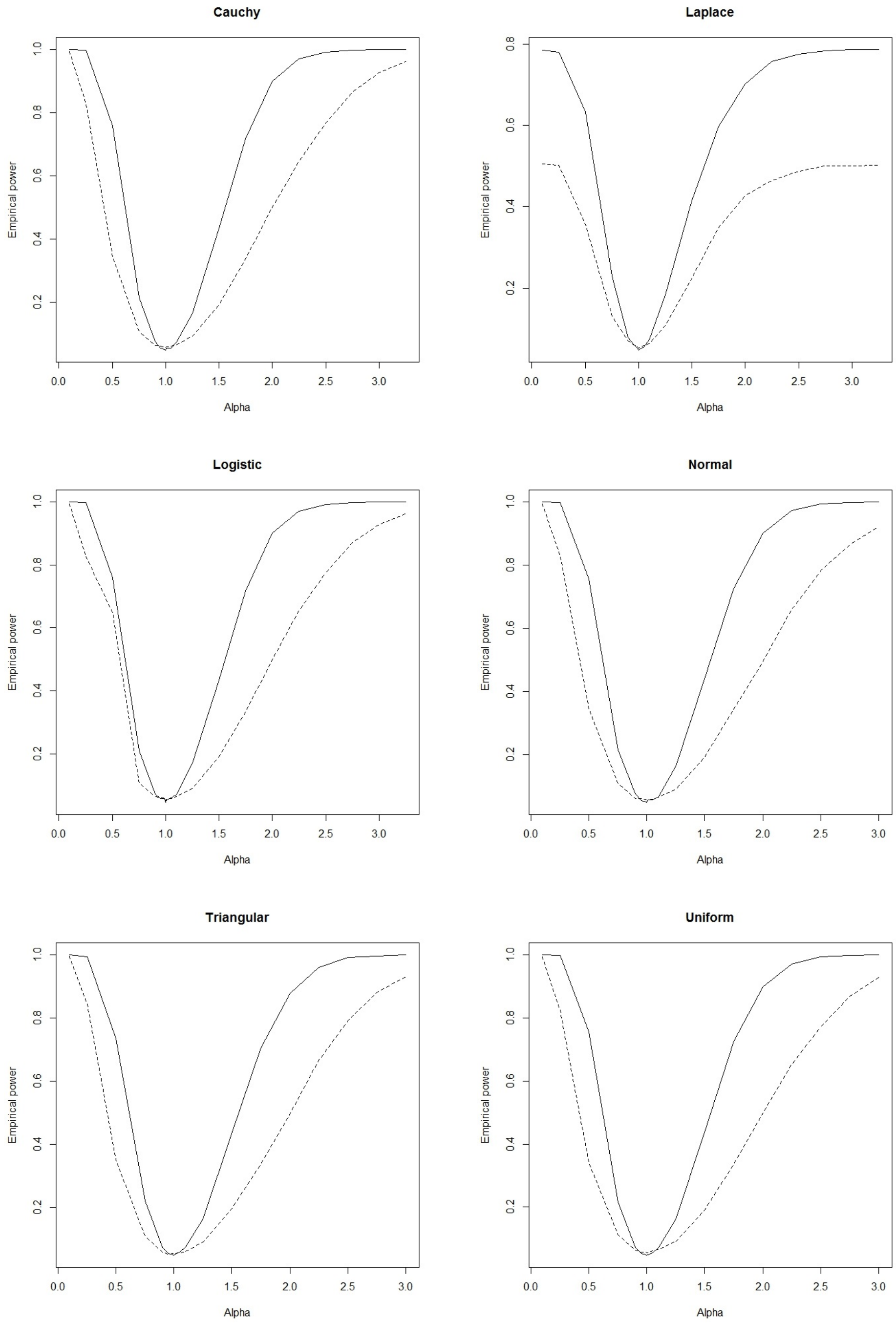
An empirical power study for both the tests was carried out for moderate sample size $n = 25$ with $\gamma = 0.05$. The results based 10000 Monte Carlo simulations are tabulated in Table 1.

Table 1: *Empirical Powers of T_{a^*} and S_{b^*} for various values of alpha with* $\gamma = 0.05$ *and number of Monte Carlo simulations* 10000.

α	Tests	Cauchy	Laplace	Logistic	Normal	Triangular	Uniform
0.1	T_{a^*}	1.0	0.7846	1.0	1.0	1.0	1.0
	S_{b^*}	0.9943	0.505	0.9949	0.9947	0.9964	0.9951
0.5	T_{a^*}	0.7578	0.6319	0.7597	0.756	0.7344	0.7569
	S_{b^*}	0.3447	0.3561	0.6493	0.3462	0.3532	0.3425
0.95	T_{a^*}	0.0550	0.0617	0.0586	0.0550	0.0558	0.0526
	S_{b^*}	0.0610	0.0600	0.0611	0.0610	0.0542	0.0567
0.99	T_{a^*}	0.0495	0.0515	0.0587	0.0499	0.0509	0.0488
	S_{b^*}	0.0581	0.0549	0.0588	0.0581	0.0561	0.0581
1.0	T_{a^*}	0.0488	0.0480	0.0460	0.0495	0.0503	0.0480
	S_{b^*}	0.0555	0.0514	0.0508	0.0514	0.0558	0.0549
1.01	T_{a^*}	0.0543	0.0493	0.0543	0.0572	0.0530	0.0495
	S_{b^*}	0.0587	0.0545	0.0556	0.0569	0.0564	0.0555
1.05	T_{a^*}	0.0548	0.0546	0.0601	0.0569	0.0588	0.0548
	S_{b^*}	0.0615	0.0589	0.0571	0.0598	0.0573	0.0615
1.5	T_{a^*}	0.4329	0.4159	0.4349	0.4410	0.4347	0.4380
	S_{b^*}	0.1920	0.2246	0.1919	0.191	0.1957	0.1907
2.0	T_{a^*}	0.9002	0.7016	0.9008	0.9016	0.8785	0.8988
	S_{b^*}	0.4989	0.4274	0.497	0.4953	0.4984	0.4989
2.5	T_{a^*}	0.9928	0.7751	0.9924	0.9932	0.9918	0.9944
	S_{b^*}	0.7685	0.487	0.7759	0.7849	0.7913	0.7719
3.0	T_{a^*}	1.0	0.7859	1.0	1.0	1.0	1.0
	S_{b^*}	0.9273	0.5004	0.9285	0.9225	0.9318	0.9283

Table 2: *

Graphs of the Empirical Powers of T_{a^*} (solid line) and S_{b^*}(longer dashing line) for various values of alpha.



From the table 1 and the above graphs we observe the following.

a) The proposed test statistics are maintaining the level of significance.

b) the empirical powers of S_{b^*} are larger in the neighborhood of the null. T_{a^*} performs better when we move away from the null hypothesis.

7 Conclusion

In this article we have considered a semi-parametric class of skew-symmetric distributions called Power-Skew-Symmetric ($\mathcal{PSS}$) distributions. We have developed two tests for symmetry, based on the theory of U-Statistics for testing symmetry in this class. The kernel functions depend on arbitrary constants, which are chosen so that efficacies of the test are maximal. Though they are asymptotic tests, based on simulation study, from Table 1, we observe that for each test the attained levels for all the models are almost equal to the nominal level. The efficacy of the test S_{b^*} is higher than that of T_{a^*}. As expected the empirical powers of S_{b^*} are larger in the neighborhood of the null, of course T_{a^*} is better than S_{b^*} if the values of α are much away from the null value 1.

Remark: Similar to the class $\mathcal{PSS}$ of power skew-symmetric distribution functions one can define the $\mathcal{PSS}^{\mathbf{s}}$ the class of power skew-symmetric survival functions, by considering the survival functions instead of the distribution functions. All the related properties and tests can be obtained in the similar way. Properties related to maximum of random variables will be now related with minimum of random variables.

References

[1] Azzalini A. (1985). A Class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*. Vol. **12**, pp. 171-178.

[2] Durrans S. R. (1992). Distributions of fractional order statistics in hydrology. *Water Resources Research*. Vol. **28**, pp. 1649-1655.

[3] Gupta A. K., Chang F. C., Huang W. J. (2002). Some Skew-Symmetric Distributions. *Random Operators and Stochastic Equations*. Vol. **10**, pp. 133-140.

[4] Gupta R. D., Gupta R. C. (2006). Analyzing skewed data by power normal model. *Test*. Vol. **17**, pp. 197-210.

[5] Gomez H. W., Salinas H. S., Bolfarine H. (2006). Generalized skew-normal models: properties and inference. *Statistics*. Vol. **40**, pp. 495-505.

[6] Hoeffding, W. (1948). A Class of Statistics with Asymptotically Normal Distribution. *Annals of Mathematical Statistics*. Vol. **19**, pp. 223-325.

[7] Lehmann E. L. (1953). The power of rank tests. *Annals of Mathematical Statistics*. Vol. **24**, pp. 23-43.

[8] Mehra, K. L., Prasad, N. G. N., Madhava Rao, K. S. (1990). A Class of Non-parametric Tests for the One-Sample Location Problem, *Australian Journal of Statistics*, **32**, pages 373-392.

[9] Mudholkar, G.S., Hutson, A. D. (2000). The epsilon-skew normal distribution for analyzing near normal data, *Journal of Statistical Planning and Inference*, **83**, pages 291-309.

[10] Nadarajah, S., Kotz, S. (2006). Skew Distributions Generated from Different Families, *Acta Applicandae Mathematica*, **91**, pages 1-37.

[11] Pewsey, A., Gomez, H. W., Bolfarine, H. (2012). Likelihood-based inference for power distributions, *Test*, **21**, pages 775-789.

[12] Randles, R. H., Wolfe, D. A. (1979). *Introduction to the Theory of Nonparametric Statistics,* Wiley, New York.

[13] Rattihalli, R.N. (2004). *Statistical Models derived from Well-Known distributions: An unpublished Report of a Research Project,* Submitted to University Grant Commission, New Delhi, India.

[14] Rattihalli, R. N., Raghunath, M. (2012). Generalized Nonparametric Tests for One-Sample Location Problem based on Sub-Samples, *ProbStat Forum*, **05**, pages 112-123.

Appendix

Graphs of the DF and PDF of Power-skew-symmetric distributions derived from various $F(.)s$, with $\alpha = 1.5$ (longer dashing line), $\alpha = 0.5$ (dotted line) and $\alpha = 1$ (solid line).