A Goodness of Fit Test for Testing Spherical Symmetry of a Bivariate Distribution

A. A. Al-Shiha
Department of Statistics & Operations Research
College of Science, King Saud University
P. O. Box 2455, Riyadh 11451
Saudi Arabia
E-mail: aalshiha@ksu.edu.sa

Abstract

Al-Shiha (2007) has proposed a simple test for testing spherical symmetry of a bivariate distribution. Even though, this test is powerful for some cases, it has a low power for rejecting the hypothesis of spherical symmetry for non spherically symmetric distributions that have an equal probability in each of the four quadrants. In this paper, we propose a simple test that avoids the deficiency of the test of Al-Shiha (2007). The proposed test is based on a goodness-of-fit chi-square test. The sampling distribution of the proposed test statistic follows approximately a chi-square distribution. We conduct a simulation study to evaluate the performance of the proposed test for testing spherical symmetry of a continuous bivariate distribution with known means.

Key words: Chi-square test, Circular symmetry, Spherical symmetry, Bivariate symmetry.

1. Introduction

Several tests for testing symmetry of a univariate distribution are proposed in the literature, see for example, Gastwirth (1971), Hill and Rao (1977), and Randles et. al (1980).

McWilliams (1990) showed the use of a runs statistic for testing symmetry of a distribution. Modarres and Gastwirth (1996) presented a modified runs test for symmetry. Tajuddin (1994) proposed a test for symmetry based on Wilcoxon two-sample test. Khalique and Tajuddin (1996) have studied the use of a sign test statistic in testing asymmetry of a distribution.

The problem of testing symmetry in a bivariate or a multivariate distribution is tackled in different ways. Several papers such as Bell and Haller (1969) and Hollander (1971) have considered the problem of testing the bivariate symmetry as testing H_O: $F(x,y) = F(y,x) \; \forall (x,y) \in \Re^2$. This problem is regarded as a test of interchangeability by Ernst and Schucany (1999) among others. Beran (1979), Alzaid et. al (1990) and others have studied the elliptical symmetry of multivariate distributions, which also requires the interchangeability of random variables. Heathcote et. al (1995) have talked about the diagonal symmetry of a p-variate distribution, whereas Small (1990) has discussed various kinds of symmetry including the concept of an angular symmetry.

Ahmad and Cerrito (1991) have discussed the requirements of a bivariate symmetry. They have proposed two different definitions of symmetry. Both the definitions do not regard the interchangeability property of random variables under bivariate symmetry.

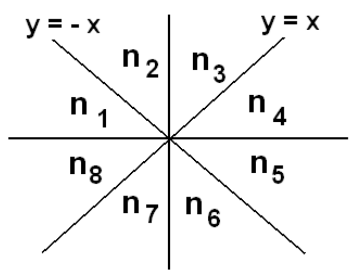
Oja and Nyblom (1989) have given several bivariate generalizations of the sign test for a univariate distribution. These tests include the bivariate sign tests proposed by Blueman (1958), Brown et. al (1992) and Hodges (1955). However, all the above mentioned authors and others such as Chaudhuri and Sengupta (1993) have given the extensions of the sign test for testing the location parameter. Koltchinskii and Li (1998) have given a test for testing spherical symmetry of a distribution in $\Re^p$, $p \geq 2$, with unknown location parameter. Romano (1989), Baringhaus (1991) and Smith (1977) have considered testing spherical symmetry for a distribution with known centre.

We present a simple test based on a chi-square goodness-of-fit test for testing spherical symmetry of a bivariate continuous distribution with known location parameter. This test can easily be extended to a multivariate case. We conduct a simulation study similar to the one carried on by Koltchinskii and Li (1998) to examine the performance of our proposed test statistic. We present the simulation results in Section 3 and we end up with the conclusion of this paper in Section 4.

2. The Proposed Test Statistic

Let $(x_i, y_i), i = 1, 2, \ldots, n,$ be a random sample of n observations from a bivariate distribution $F_{X,Y}(x,y)$ with known mean vector, $\boldsymbol{\mu} = (\mu_x, \mu_y)'$. With no loss of generality, we can assume $\mu_x = \mu_y = 0$. Let the number of observations falling in the i-th octant, A_i, be n_i $(i = 1, 2, \ldots, 8)$, respectively, as described in Figure 1, where $n = \sum_{i=1}^{8} n_i$. For example, n_1 is the number of observations falling in the 1-st octant, $A_1 = \{(x, y) \in \Re^2 : x < 0, y > 0, x < -y\}$.

Figure 1.



Let p_i be the proportion of observations falling in the i-th octant. It is well known that the distribution of $(n_1, n_2, \ldots, n_8)$ follows a multinomial distribution. If we assume that the given bivariate distribution $F_{X,Y}(x,y)$ is spherically symmetric about the origin, then the distribution of $(n_1, n_2, \ldots, n_8)$ will be multinomial distribution with equal parameters, $p_1 = p_2 = \ldots = p_8 = 1/8$. Our goal is to test:

H_o : The bivariate distribution is spherically symmetric,

H_1 : The bivariate distribution is not spherically symmetric,

or equivalently,

$H_o: \; p_i = 1/8 \; ; \; for \;\; all \;\; i = 1,2,\ldots,8 \, ,$

$H_1: \; p_i \neq 1/8 \; ; \; for \;\; some \;\; i \, .$

Under H_o, the expected number of observations falling in the i-th octant is $E_i = n/8$. The test statistic is the following goodness-of-fit chi-square statistic:

$$\chi_O^2 = \sum_{i=1}^{8} \frac{(n_i - E_i)^2}{E_i}$$

$$= \frac{8}{n} \sum_{i=1}^{8} (n_i - \frac{n}{8})^2 \qquad (1)$$

Under H_o, and for large values of n, the sampling distribution of χ_O^2 follows approximately a chi-square distribution with 7 degrees of freedom. The test rejects H_o (the hypothesis of spherically symmetry) if the value of χ_O^2 exceeds $\chi_\alpha^2(7)$, the $(1-\alpha)100^{th}$ percentile of a chi-square distribution with 7 degrees of freedom, at a specified value of α .

The procedure of construction this test statistic is motivated by the procedure proposed by Al-Shiha (2007) where the test statistic was given by:

$$S = \frac{4}{n} \sum_{i=1}^{4} (n_i^* - \frac{n}{4})^2 \qquad (2)$$

where n_i^* is the number of observations falling in the i-th quadrant. Under H_o, and for large values of n, the sampling distribution of S follows approximately a chi-square distribution with 3 degrees of freedom.

3. Simulation Study

3.1 Empirical Size of the Test

We compare the empirical α with the nominal α values of the proposed test statistic for the choices of $\alpha = 0.01, 0.05$ and 0.1 and various sample sizes ($n = 10(10)100, 200$) from different underlying distributions. We consider all the sampling distributions considered by Koltchinskii and Li (1998). They have considered the following distributions under the null hypothesis that the distribution is spherically symmetric.

(1) $H_o^{(1)}$ = the standard bivariate normal distribution, i.e.,

$$f(x,y) = \frac{1}{2\pi} \exp\{-(x^2 + y^2)/2\}; \;\; (x,y) \in \Re^2 \, . \qquad (3)$$

(2) $H_o^{(2)}$ = the uniform distribution on the unit disk, i.e.,

$$f(x,y) = \frac{1}{\pi} \; ; \; 0 < x^2 + y^2 < 1; \;\; (x,y) \in \Re^2 . \qquad (4)$$

(3) $H_o^{(3)}$ = the uniform distribution on the unit circle, i.e.,

$$f(x,y) = \frac{1}{2\pi} \; ; \; x^2 + y^2 = 1; \;\; (x,y) \in \Re^2 . \qquad (5)$$

According to Koltchinskii and Li (1998), all distributions considered under H_o are spherically symmetric.

We draw 10000 random samples of various sizes from each of the above mentioned distributions and we observe the number of times the observed test statistic exceeds the critical value for different choices of level of significance. The results obtained are presented in Tables 1 and Figure 2. Table 1 and Figure 2 compare the empirical α values with the nominal α values for different choices of the sample size, n. From Table 1 and Figure 2, we observe that the nominal values of α tend to be very close to the empirical α for all cases considered.

We also compute the empirical confidence intervals for the nominal α values for different choices of the sample size, n. For each combination (α, n), we simulated 100 random samples, computed the sample mean of the empirical sizes, $\overline{\alpha}$ and the sample standard deviation, S_α, and then we computed the confidence interval $\overline{\alpha} \pm 2S_\alpha$. The results obtained are presented in Tables 2 and Figure 3.

From Table 2 and Figure 3, we observe that the nominal values of α lie within the confidence intervals for all cases considered excepts for $\alpha = 0.1$ at $n = 20$ and $n = 40$. However, for these cases, the nominal values of α tend to be larger than the upper confidence limit indicating that the proposed test tends to be slightly conservative for these cases. Consequently, this behavior will strengthen the procedure for the cases where H_o is rejected.

Based on the simulation study of this section, we conclude that that the proposed test is a valid test.

Table 1: Empirical α at different nominal values of α
for different underlying distributions under H_o

Sample Size n	Underlying Distributions								
	$H_o^{(1)}$			$H_o^{(2)}$			$H_o^{(3)}$		
	Nominal Alpha (α)								
	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
10	0.0081	0.0343	0.0913	0.0087	0.0318	0.0929	0.0095	0.0334	0.0932
20	0.0082	0.0445	0.0801	0.0084	0.0493	0.0811	0.0079	0.0434	0.0786
30	0.0112	0.0478	0.0965	0.0106	0.0481	0.0972	0.0087	0.0477	0.0982
40	0.0084	0.0453	0.0913	0.0097	0.0447	0.0859	0.0078	0.0424	0.0881
50	0.0104	0.0468	0.0979	0.0106	0.0483	0.1019	0.0109	0.0484	0.0993
60	0.0078	0.0474	0.0943	0.0095	0.0512	0.0960	0.0094	0.0497	0.0953
70	0.0104	0.0539	0.1041	0.0088	0.0493	0.0975	0.0082	0.0492	0.0998
80	0.0097	0.0454	0.0917	0.0099	0.0509	0.0974	0.0100	0.0496	0.0977
90	0.0087	0.0464	0.0996	0.0088	0.0479	0.0989	0.0110	0.0504	0.1010
100	0.0114	0.0475	0.0932	0.0084	0.0483	0.0918	0.0087	0.0503	0.0970

200	0.0105	0.0520	0.1037	0.0090	0.0482	0.0952	0.0096	0.0503	0.0939

Figure 2: Empirical α for $H_o^{(1)}$, $H_o^{(2)}$, and $H_o^{(3)}$.

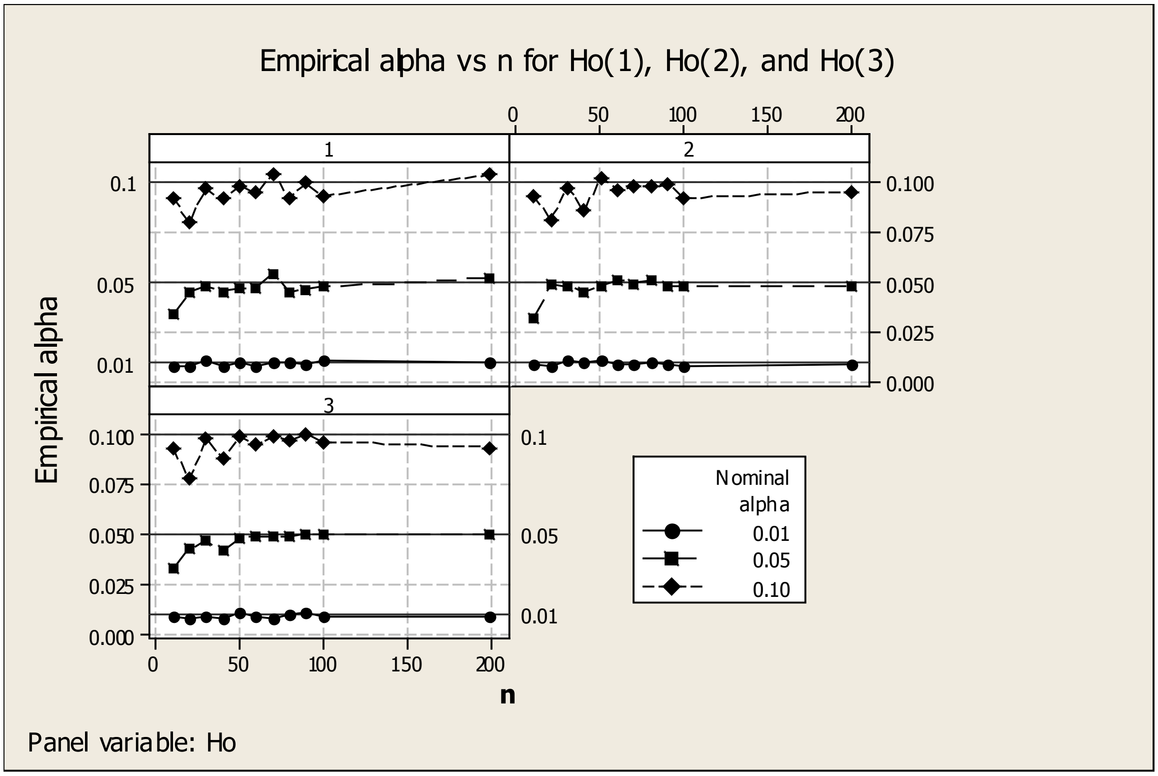


Table 2-a: Confidence interval for the size of the test α at different nominal values of α for the underlying distribution under $H_o^{(1)}$

Sample Size n	Nominal Alpha (α)					
	0.01		0.05		0.10	
	Lower	Upper	Lower	Upper	Lower	Upper
10	0.0077	0.0118	0.0313	0.0387	0.0888	0.1006
20	0.0064	0.0100	0.0415	0.0494	0.0738	0.0845
30	0.0085	0.0124	0.0435	0.0523	0.0920	0.1051
40	0.0074	0.0110	0.0416	0.0501	0.0841	0.0955
50	0.0084	0.0124	0.0421	0.0504	0.0937	0.1076
60	0.0073	0.0117	0.0446	0.0544	0.0868	0.0991
70	0.0074	0.0114	0.0454	0.0541	0.0936	0.1057
80	0.0077	0.0118	0.0440	0.0528	0.0891	0.1010
90	0.0079	0.0113	0.0442	0.0531	0.0945	0.1068
100	0.0078	0.0117	0.0457	0.0537	0.0896	0.1017

200	0.0080	0.0118	0.0458	0.0542	0.0929	0.1044

Table 2-b: Confidence interval for the size of the test α at different nominal values of α for the underlying distribution under $H_o^{(2)}$

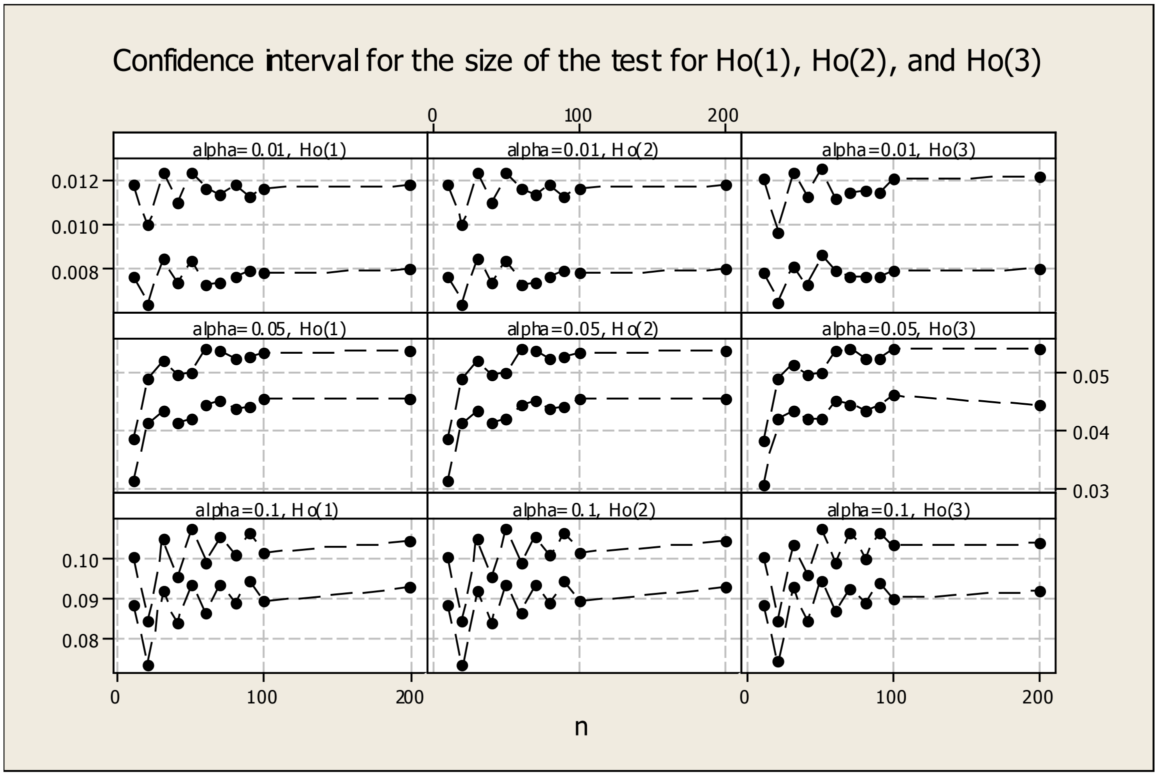
Sample Size n	Nominal Alpha (α)					
	0.01		0.05		0.10	
	Lower	Upper	Lower	Upper	Lower	Upper
10	0.0077	0.0117	0.0308	0.0385	0.0882	0.1009
20	0.0062	0.0103	0.0414	0.0499	0.0744	0.0850
30	0.0088	0.0121	0.0442	0.0513	0.0930	0.1036
40	0.0074	0.0111	0.0422	0.0494	0.0841	0.0949
50	0.0085	0.0127	0.0418	0.0509	0.0957	0.1064
60	0.0079	0.0113	0.0454	0.0533	0.0873	0.0982
70	0.0078	0.0115	0.0463	0.0540	0.0943	0.1060
80	0.0078	0.0116	0.0439	0.0517	0.0884	0.1012
90	0.0077	0.0118	0.0442	0.0529	0.0955	0.1057
100	0.0080	0.0121	0.0448	0.0545	0.0893	0.1026
200	0.0080	0.0120	0.0453	0.0544	0.0916	0.1039

Table 2-c: Confidence interval for the size of the test α at different nominal values of α for the underlying distribution under $H_o^{(3)}$

Sample Size n	Nominal Alpha (α)					
	0.01		0.05		0.10	
	Lower	Upper	Lower	Upper	Lower	Upper
10	0.0078	0.0121	0.0308	0.0384	0.0889	0.1008
20	0.0065	0.0097	0.0421	0.0493	0.0745	0.0847
30	0.0081	0.0124	0.0436	0.0518	0.0931	0.1038
40	0.0073	0.0113	0.0421	0.0501	0.0847	0.0960
50	0.0087	0.0126	0.0423	0.0502	0.0946	0.1078
60	0.0079	0.0112	0.0455	0.0540	0.0870	0.0992
70	0.0077	0.0115	0.0448	0.0546	0.0926	0.1067
80	0.0077	0.0116	0.0436	0.0529	0.0890	0.1003
90	0.0077	0.0115	0.0445	0.0527	0.0941	0.1065
100	0.0079	0.0121	0.0463	0.0544	0.0903	0.1034

200	0.0080	0.0122	0.0447	0.0546	0.0921	0.1039

Figure 3: Confidence intervals for the size of the test α at different nominal values of α for the underlying distribution under $H_o^{(1)}$, $H_o^{(2)}$, and $H_o^{(3)}$.



3.2 The Power of the Test

To evaluate the power of the test statistic, the following distributions are considered under the alternative hypothesis that the distribution is not spherically symmetric:

(1) $H_1^{(1)}$ = the distribution of random vector with two independent exponential variates with parameters $\lambda_1 = 1$ and $\lambda_2 = 2$, respectively; i.e.,

$$f(x,y) = 2\exp\{-x-2y\};\ \ x>0,\ y>0. \qquad (6)$$

(2) $H_1^{(2)}$ = the distribution of the random vector with two independent components, exponential with parameter $\lambda = 1$ and standard normal; i.e.,

$$f(x,y) = \frac{1}{\sqrt{2\pi}}\exp\{-x-y^2/2\};\ \ x>0,\ y\in\Re. \qquad (7)$$

(3) $H_1^{(3)}$ = the mixture (with parameter 1/2) of two normal distributions in $\Re^2$ with unit covariances and with means (0,0) and (3,0); i.e.,

$$f(x,y) = \frac{1}{2\sqrt{2\pi}}[\exp\{-x^2/2\}+\exp\{-(x-3)^2/2\}]\exp\{-y^2/2\};\ \ (x,y)\in\Re^2. \qquad (8)$$

(4) $H_1^{(4)}$ = the uniform distribution in equilateral triangle with center point (0,0); i.e.,

$$f(x,y) = 1;\ \ -\sqrt{2} \le 2y-1 \le |x| \le \sqrt{2}. \qquad (9)$$

According to Koltchinskii and Li (1998), the distributions considered under H_1 are not spherically symmetric although the last three are symmetric in some forms.

We draw 10000 random samples of various sizes from each of the above mentioned distributions and we observe the number of times the observed test statistic exceeds the critical value for different choices of level of significance. The results obtained are presented Table 3 and Figure 4. Table 3 and Figure 4 present the power of the test against various alternatives for nominal $\alpha = 0.01, 0.05$ and 0.1 and different choices of sample size, $n = 10(10)100, 200$.

From Table 3 and Figure 4, we observe that the test is quite powerful against all alternatives $H_1^{(1)}$, $H_1^{(2)}$, $H_1^{(3)}$ and $H_1^{(4)}$. In particular, the test is most powerful for the case $H_1^{(4)}$ and is less powerful for the case $H_1^{(2)}$.

4. Conclusion

In Section 3, we noticed that the test statistic χ_O^2 is a valid test for testing spherical (circular in this case) symmetry of a bivariate distribution. The test χ_O^2 is more powerful than the test proposed by Al-Shiha (2007). Specifically, the test χ_O^2 has high power for rejecting all four alternatives that considered, whereas Al-Shiha's (2007) test has low power for rejecting the alternatives $H_1^{(3)}$ and $H_1^{(4)}$.

It is expected that the test will have low power for rejecting false H_o when the distribution has an equal probability in each of the eight octants, i.e., when the following condition is fulfilled:

$$P\{(X,Y) \in A_i\} = 1/8 \text{, for all } i = 1, 2, \ldots, 8. \quad (10)$$

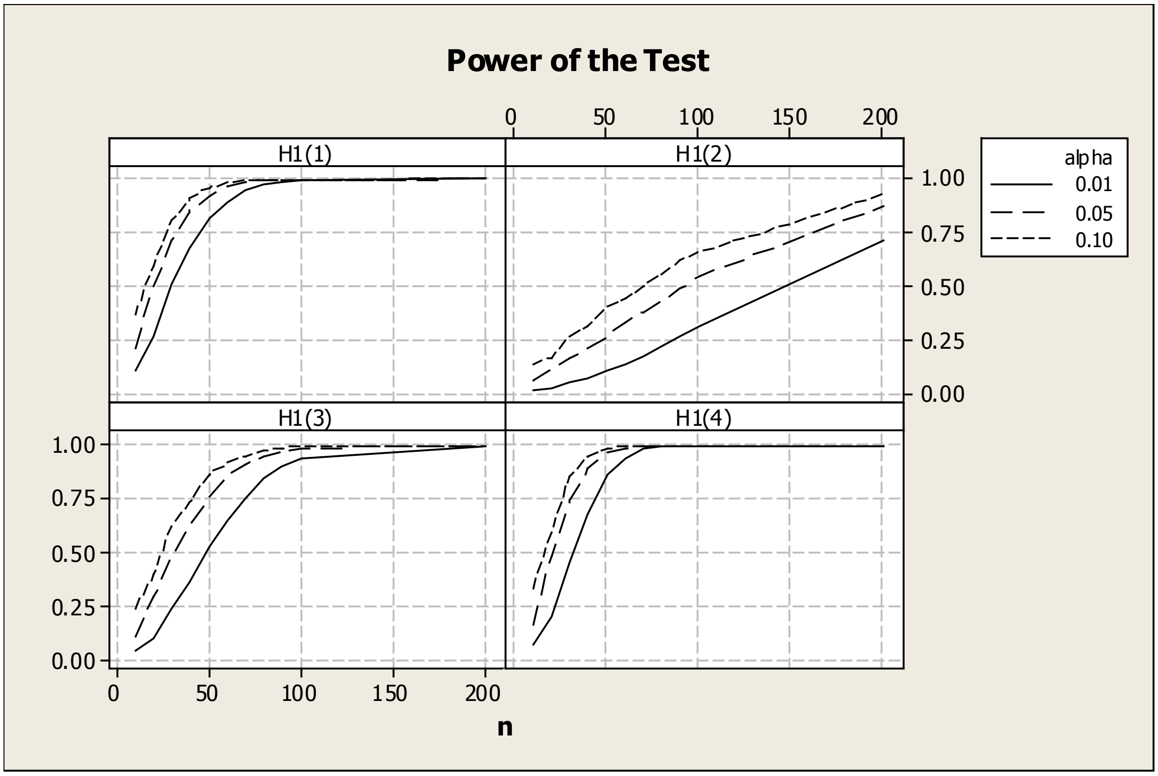
The proposed test is very simple and its sampling distribution is known. In addition, the test can be easily extended to test the spherical symmetry of a p-variate distribution.

Table 3: Power of the test against different alternatives and different choices of α and sample sizes

Sample Size	Distribution under Alternatives											
	$H_1^{(1)}$			$H_1^{(2)}$			$H_1^{(3)}$			$H_1^{(4)}$		
	Nominal Alpha (α)											
n	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
10	0.1079	0.2167	0.3697	0.0206	0.0603	0.1429	0.0410	0.1069	0.2437	0.0732	0.1630	0.3372
20	0.2699	0.5018	0.5924	0.0306	0.1142	0.1700	0.0994	0.2935	0.3959	0.2027	0.4840	0.6009
30	0.5089	0.7121	0.8092	0.0574	0.1650	0.2710	0.2401	0.4762	0.6258	0.4581	0.7409	0.8541
40	0.6745	0.8497	0.9074	0.0751	0.2100	0.3163	0.3629	0.6224	0.7379	0.6777	0.8903	0.9486
50	0.8163	0.9226	0.9622	0.1147	0.2689	0.4021	0.5301	0.7578	0.8624	0.8608	0.9638	0.9885
60	0.8971	0.9660	0.9821	0.1420	0.3324	0.4415	0.6527	0.8572	0.9171	0.9402	0.9899	0.9962
70	0.9498	0.9861	0.9936	0.1740	0.3818	0.5060	0.7527	0.9106	0.9537	0.9826	0.9974	0.9996

80	0.9753	0.9957	0.9984	0.2256	0.4311	0.5596	0.8458	0.9525	0.9776	0.9925	0.9995	0.9999
90	0.9893	0.9975	0.9988	0.2661	0.4891	0.6208	0.9004	0.9759	0.9900	0.9976	0.9998	0.9999
100	0.9957	0.9991	0.9996	0.3134	0.5471	0.6589	0.9426	0.9880	0.9954	0.9995	0.9999	1.0000
200	1.0000	1.0000	1.0000	0.7180	0.8742	0.9295	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000

Figure 4: Power of the Test for the alternatives $H_1^{(1)}$, $H_1^{(2)}$, $H_1^{(3)}$, and $H_1^{(4)}$



REFERENCES

Ahmad, I. A. and Cerrito. P. B. (1991), Bivariate Symmetry: Definitions and Basic Properties. *Nonparametric Statistics*, Vol. 1, 165-169.

Al-Shiha, A.A. (2007), Bivariate Symmetry and a Simple Generalization of the Sign Test for Testing Spherical Symmetry of a Bivariate Distribution. Third Saudi Science Conference, Riyadh, 10 – 13th of March 2007.

Alzaid, A. A., Rao, C. R. and Shanbhag, D. N. (1990), Elliptical Symmetry and Exchangeability with Characterizations, *J. Mult. Anal.*, 33, 1-16.

Baringhaus, L. (1991), Testing for Spherical Symmetry of a Multivariate Distribution. *Ann. Statist.*, 19, 899-917.

Bell, C. B. and Haller. H. S. (1969), Bivariate Symmetry Tests: Parametric and Nonparametric, *Ann. Math. Statist.*, 40, 259-269.

Beran, R. (1979). Testing for Elliptical Symmetry of Multivariate Density, *Ann. Statist.*, *7*, 150-162.

Blueman, I. (1958), A new Bivariate Sign Test, *JASA*. 448-456.

Brown, B. M., Hettmansperger, T. P., Nyblom, J. and Oja, H. (1992), On certain Bivariate Sign Tests and Medians, *JASA*, 87, 127-135.

Chaudhuri. P. and Sengupta, D. (1993), Sign Tests in Multidimension: Inference based on the Geometry of the Data Cloud, *JASA*, 1363-1370.

Ernst, M. D. and Schucany, W. R. (1999), A class of Permutation Tests of Bivariate Interchangeability. *JASA*, 273-284.

Gastwirth, J. L. (1971), On the Sign Test for Symmetry, *JASA*, 821-823.

Heathcote, C. R., Rachev, S. T. and Cheng, B. (1995), Testing Multivariate Symmetry, *J. Mult. Anal.*, 91-112.

Hodges, J. L. (1955), A Bivariate Sign Test, *Ann. Math. Statist.*, 26, 523-527.

Hill, D. L. and Rao, P. V. (1977), Tests of Symmetry Based on Cramer-Von Mises Statistics, *Biometrika*, 64, 489-494.

Hollander, M. (1971), A Nonparametric Test for Bivariate Symmetry, *Biometnka*. 58, 1. 203-212.

Khalique, A. and Tajuddin, I. H. (1996), A Test of Symmetry based on Sign Statistic, *Pak. J. Statistics*, 243-249.

Koltchinskii, V. I. and Li, L. (1998), Testing for Spherical Symmetry of a Multivariate Distribution, *J. Mult. Anal.*, 228-244.

McWilliams, T. P. (1990), A Distribution-Free Test for Symmetry Based on a Runs Statistic, *JASA*, 1130-1133.

Modarres, R. and Gastwirth. J. L. (1996), A Modified Runs Test for Symmetry, *Statistics and Probability letters,* 31, 107-112.

Oja, H. and Nyblom, J. (1989), Bivariate Sign Tests, *JASA*, 249-259.

Romano, J. P. (1989), Bootstrap and Randomization Tests of some Nonparametric Hypotheses, *Ann. Statist.*. 141-159.

Randles, R. H., Flinger, M. A.. Policello, G. E and Wolfe, D. A. (1980), An Asymptotically Distribution-Free Test for Symmetry versus Asymmetry, *JASA*, 75, 163-172.

Small, C. G. (1990). A Survey of Multidimensional Medians, *International Statist. Rev.*, 58, 3, 263-277.

Smith, P. J. (1977). A Nonparametric Test for Bivariate Circular Symmetry, *Commu. Statist. - Theor - meth.*, A6(3), 209-220.

Tajuddin, I. H. (1994), Distribution-Free Test for Symmetry based on Wilcoxon two-sample Test, *J. Appl. Statist.*, Vol. 21, No. 5, 409-415.