IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

A Study on Goodness of Fit for Normality

Tripakshi Borthakur[1] and Bipin Gogoi[2]

Research Scholar, Dept of Statistics, Dibrugarh University, Assam[1]

Professor, Dept of Statistics, Dibrugarh University, Assam[2]

Abstracts: One of the important assumptions of data is the normality on which most of the statistical model and procedures rely on regarding the validity of given data hypothesis. Assuming the normality assumption blindly may affect the accuracy of inferences and estimation procedures. As observed, the collected data from real field are not always follow the normality assumption. So, data must be verified with adequate statistical test before used. There are various kinds of goodness of fit tests in literature. Some of them are special purpose tests, so that they are suitable and perform well for some special situations. Others are omnibus tests that are applicable to general cases. Most commonly used tests are Pearson’s chi-squared test and EDF (empirical distribution function) tests, such as Kolmogorov-Smirnov, Cramer-Von-Mises and Anderson-Darling test. The chi-squared test is easy to use but they are generally less powerful than EDF tests. In this paper we want to study the performance of twelve different tests for normality including the above mentioned tests. Considering various sample sizes and different alternative hypotheses results are obtained and displayed in different tables. Finally, discussions are made on the basis of the results.

Keywords: normality test; power comparison; simulation method, alternative of the form symmetrical and asymmetrical distribution

1. INTRODUCTION

Fitting of a probability model to observed data is an important statistical problem from both theory and application point of view. There is a multitude of statistical models and procedures that rely on the validity of a given data hypothesis, being the normality of the data assumption one of the most commonly found in statistical studies. As observed in many models and in research on applied statistics and economics, following the normal distribution assumption blindly may affect the accuracy of inference and estimation procedures. The evaluation of this distributional assumption has been addressed, for example, in Min (2007) where the conditional normality assumption in the sample selection model applied to housing demand is examined. The definition of adequate normality tests can, therefore, be seen to be of much importance since the acceptance or rejection of the normality assumption of a given data set plays a central role in numerous research fields. As such, the problem of testing normality has gained considerable importance in both theoretical and empirical research and has led to the development of a large number of goodness-of-fit tests to detect departures from normality. Given the importance of this subject and the widespread development of normality tests over the years, comprehensive descriptions and power comparisons of such tests have also been the focus of attention, thus helping the analyst in the choice of suitable tests for his particular needs. Examples of such comprehensive reviews on the effectiveness of many normality tests towards a wide range of non-normality alternatives may be found, for example, in Shapiro and Wilk (1965), Stephens(1974), D’Agostino (1971), Bonett and Seier(2002), Farrell and Rogers-Stewart(2006), Yazici and Yolacan(2007) and in the references cited therein. Since the tests that have been developed are based on different characteristics of the normal distribution, it can be seen from these comparison studies that their power to detect departures from normality can be significantly different depending on the nature of the non-normality. Furthermore, although the referred comparison studies have been appearing over the years, it is worth mentioning that some of the more recent ones, e.g. Farrel and Stewart (2006), Yazici and Yolacan(2007), do not include several interesting and more recently developed tests. Moreover, power results presented in Yazici and Yolacan(2007) appear to contradict those resulting from previous studies. A further comparison of normality tests, such as the one proposed herein, can therefore be considered to be of foremost interest. A simulation study is presented herein to estimate the power of twelve tests aiming to assess the validity of the univariate normality assumption of a data set. The selected tests include a group of well-established normality tests as well as more recently developed ones. Section 2 presents a general description of the normality tests selected for the study. The effects on the power of the tests due to the sample size, the selected significance level and the type of alternative distribution are also considered in the proposed study.

The study is carried out for various sample sizes n and considering several significance levels α. With respect to the considered alternative distributions, the study considers a number of statistical distributions that are categorized into three sets. The first set includes several types of symmetric non-normal distributions, the second set includes several types of asymmetric distributions and the third set comprises modified normal distributions with various shapes. Section 3 presents the simulation approach considered in the study and the power results of the normality tests for the different alternative distribution sets, which are then discussed.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

Finally, conclusions and recommendations resulting from the study are provided in Section 4.

2. GOODNESS-OF-FIT TESTS FOR NORMALITY

The selected normality tests are considered for testing the composite null hypothesis for the case where both location and scale parameters, μ and σ, respectively, are unknown. Normality test formulations differ according to the different characteristics of the normal distribution they focus. The goodness-of-fit tests considered in the proposed study are grouped into four general categories viz. based on Empirical distribution function, based on moments, based on regression and correlation and others and a brief review of each test is presented herein. In the following review, it is considered that $x_1, x_2, ..., x_n$ represent a random sample of size n; $x_{(1)}, x_{(2)}, ..., x_{(n)}$ represent the order statistics of that sample; $\bar{x}, s^2, \sqrt{b_1}$, and b_2 are the sample mean, variance, skewness and kurtosis, respectively, given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}} \quad , \quad b_2 = \frac{m_4}{(m_2)^2} \quad . . . \ (1)$$

where the jth central moment m_j is given by

$$m_j = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^j$$

2.1 Goodness of Fit test Based on Empirical Distribution Functions

2.1.1 The Kolmogorov-Smirnov test modified by Lilliefors and Stephens

Kolmogorov and Smirnov (1933) developed a one sample goodness of fit test based on empirical distribution function (EDF). Lilliefors (1967) proposed a modification of Kolmogorov-Smirnov test for normality when the mean and the variance are unknown, and must be estimated from the data. The test statistic K-S is defined as

$$KS = \max_{1\le i\le n} \left[\Phi(x_i; \bar{x}; s^2) - \frac{(i-1)}{n}; \frac{i}{n} - \Phi(x_i; \bar{x}, s^2) \right] \quad . . . \quad (2.1)$$

Where $\Phi(x_i; \bar{x}, s^2)$ is the cumulative distribution function of the normal distribution with parameters estimated from the data. The normality hypothesis of the data is then rejected for large values of K-S. Table of percentage points are found in Lilliefors (1967). Modification of K-S statistic given by Stephens (1969) from the Lilliefors form is as follows;

$$KS^* = KS\,(\sqrt{n} - 0.01 + 0.85/\sqrt{n}) . . . (2.2)$$

Comparing with the upper tail significance points of the distribution on the null hypothesis; may be rejected the null hypothesis if value of KS* exceeds the table value at corresponding significance levels. Table of percentage point is available in Stephens (1969).

2.1.2. The Anderson- Darling test

Anderson and Darling (1952, 1954) introduced a new class of quadratic test statistics. These are given by

$$Q_n(\psi) = n. \int_{-\infty}^{\infty} [F_n(x) - \Phi(x)]^2 \psi(x) dF(x)$$

Where $F_n(x)$ is empirical distribution function (EDF), $\Phi(x)$ is the cumulative distribution function of the standard normal distribution and $\psi(x)$ is a weight given by

[$\Phi(x).(1-\Phi(x))]^{-1}$. It can be seen from Anderson-Darling (1954) that AD can be written as

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[In\Phi(z_{(i)}) + In(1-\Phi(z_{(n+1-i)})] \quad . . . \quad (2.3)$$

Where $z_{(i)} = (x_{(i)} - \bar{x})/s$. In order to increase its power when μ and σ are estimated from the sample, a modification factor has proposed for AD by Stephens (1974) resulting in new statistic AD*:

$$AD^* = AD\,(1 + \frac{0.75}{n} + \frac{2.25}{n^2}) \quad . . . \ (2.4)$$

The normality hypothesis of the data is then rejected for large values of the test statistic.
Table of percentage points of this statistic is given by D'Agostino (1986).

1.1.3 The Zhang-Wu Zc

Zhang and Wu (2005) recently proposed a new class of EDF test statistics Z_C and Z_A of the general form

$$Z = \int_{-\infty}^{\infty} 2n\{F_n(x) In\left(\frac{F_n(x)}{F_0(x)}\right) + (1-F_n(x))\, In\left[\frac{(1-F_n(x))}{(1-F_0(x))}\right]\} dx(x)$$

Where $F_0(x)$ is a hypothetical distribution function completely specified and w(x) is a weight function. In the case where dw(x) is considered to be $[1/F_0(x)]$. $[1/(1-F_0(x))]dF_0(x)$ and $F_0(x)$ is $\Phi(x)$, the test statistic is obtained by

2.1.3(a) Zhang Z_C test

$$Z_C = \sum_{i=1}^{n}[In\frac{(1/\Phi(z_{(i)}) - 1)}{(n-0.5)/(i-0.75) - 1}]^2 . . . (2.5)$$

2.1.3(b) Zhang Z_A test

In the case where dw(x) is considered to be $[1/F_n(x)].[1/(1-F_n(x))]dF_n(x)$, the test statistic Z_A is the obtained by

$$Z_A = -\sum_{i=1}^{n}\left[\frac{In\Phi(z_{(i)})}{n-i+0.5} + \frac{In[1-\Phi(z_{(i)})]}{i-0.5}\right] \quad \ldots (2.6)$$

For both tests, the normality hypothesis of the data is rejected for large values of the test statistic. Table values of statistics Z_C, and Z_A are available in Zhang (2001).

2.2 Goodness of Fit test Based on Moments

2.3

2.2.1 The Jarque-Bera test

The Jarque-Bera test is a popular goodness of fit test in the field of economics. It has been first proposed by Bowman and Shenton (1975) but is mostly known from the proposal of Jarque and Bera (1980). The test statistic JB is defined by

$$JB = \frac{n}{6}\left(b_1 + \frac{(b_2-3)^2}{4}\right) \quad \ldots \quad (2.7)$$

The normality hypothesis of the data is rejected for large values of the test statistic. In addition, according to Bowman and Shenton (1975), it can be seen that JB is asymptotically chi-squared distributed with two degrees of freedom.

2.2.2 The Doornik- Hansen test

Various modifications of the Jarque-Bera test have been proposed over the years in order to increase its efficiency. For example, Urzua (1996) introduced a modification consisting of a different standardization process for b_1 and b_2 , though Thadewald and Buning(2007) showed that such modification did not improve the power of the original formulation. A less known formulation is that of Doornik and Hansen (1994), which suggests the use of the transformed skewness and the use of a transformed kurtosis according to the proposal in Bowman and Shenton (1977). The statistic of Doornik-Hansen test is thus given by

$$DH = [Z(\sqrt{b_1})]^2 + [z_2]^2 \quad \ldots (2.8)$$

Where the transform skewness Z ($\sqrt{b_1}$) and kurtosis z_2 are obtained by Bowman and Shenton (1977)

$$Z(\sqrt{b_1}) = \frac{In(Y/c + \sqrt{(Y/c)^2+1})}{\sqrt{In(w)}}$$

$$z_2 = \left[\left(\frac{\xi}{2a}\right)^{1/3} - 1 + \frac{1}{9a}\right](9a)^{1/2}$$

With ξ and a obtained by

$$\xi = (b_2 - 1 - b_1).2k \ ;$$

$$k = \frac{(n+5)(n+7)(n^3+37n^2+11n-313)}{12(n-3)(n+1)(n^2+15n-4)}$$

$$a_1 = \frac{(n+5)(n+7)}{6(n-3)(n+1)(n^2+15-4)}$$

$$a_2 = [(n-2)(n^2+27n-70) + b_1(n-7)(n^2+2n-5)]$$

$$a = a_1.a_2$$

The normality hypothesis of the data is rejected for large values of the test statistic and DH is also approximately chi-squared distributed with two degrees of freedom.

2.2.3 The Gel-Gastwirth robust Jarque-Bera test

Gel and Gastwirth (2008) recently proposed a robust version of the Jarque-Bera test. Stemming from the fact that sample moments are, among other things, known to be sensitive to outliers. Gel and Gastwirth have proposed a modification of JB that uses a robust estimate of the dispersion in the skewness and kurtosis definitions given in equation (1) instead of the second order central moments m_2 . The selected robust dispersion measure is the average absolute deviation from the median and leads to the following statistic of the robust Jarque-Bera test RJB given by

$$RJB = \frac{6}{n}\left(\frac{m_3}{J_n^3}\right)^2 + \frac{n}{64}\left(\frac{m_4}{J_n^4} - 3\right)^2 \quad \ldots (2.9)$$

With J_n obtained by

$$J_n = \frac{\sqrt{\pi/2}}{n}\sum_{i=1}^{n}|x_i - M|$$

In which M is the sample median. The normality hypothesis of the data is rejected for large values of the test statistic and RJB asymptotically follows the chi-square distribution with two degrees of freedom.

2.2.4 The Bonett- Seier test

Bonett and Seier (2002) have suggested a modified measure of kurtosis for testing normality, which is based on a modification of Geary's proposal (1936). The test statistic of new kurtosis measure T_w is thus given by

$$T_w = \frac{\sqrt{n+2}\,(\hat{\omega}-3)}{3.54} \quad \ldots \quad (2.10)$$

In which $\hat{\omega}$ is set by

$$\hat{\omega} = 13.29[\text{In}\sqrt{m_2} - In(\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|)].$$

The normality hypothesis of the data is rejected for both small and large values of T_w using a two-sided test and it is suggested that T_w approximately follows a standard normal distribution.

2.3. Regression and Correlation tests

2.3.1 The Shapiro-Wilk test

The Shapiro and Wilk (1965) W statistic is a well-established and powerful test of normality. The statistic W represents the ratio of two estimates of the variance of a normal distribution and is obtained by

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{n.m^2} \quad \ldots \quad (2.11)$$

Where the vectors of weights a is obtained by $(a_1, a_2, \ldots, a_n) = m.V^{-1}.(m.V^{-1}.V^{-1}.m^t)^{-0.5}$, in which m and V are the mean vector and covariance matrix of the order statistics of the standard normal distribution. The computation of the vector of weights a considered herein is defined according to the improved algorithm presented by Royston (1995), which considers the methodology described in Royston (1992) and Royston (1993). Given the definition of W, it is intuitive to observe the normality hypothesis of the data is rejected for small values of W. in order to simplify the application of this test, transformations g have been defined in Royston (1993) for different sample sizes such that g(W) approximately follows a standard normal distribution.

2.3.2 The D'Agostino D test

D'Agostino (1971) proposed the test statistic D as an extension of the Shapiro-Wilk test. The test statistic D is given by

$$D = \frac{T}{\sqrt{n^3 SS}} \quad \ldots \quad (2.12)$$

where, $T = \sum (i - \frac{n+1}{2}) X_i$

$$SS = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Here i is the order or rank of observation X. The test statistic D gives an upper and lower critical values. For each significance level, if the calculated D is less than or equal to the first member of the pair of critical values, or greater than or equal to the second member, then the normality hypothesis is rejected.

2.3.3 The Filliben correlation test

Filliben (1975) described the probability plot correlation coefficient r as a test for normality. The correlation coefficient is defined between the sample order statistics and the estimated median values of the theoretical order statistics.

Considering that $m_{(1)}, m_{(2)}, \ldots, m_{(n)}$ represent the estimated median values of the order statistics from a uniform distribution U(0,1), each $m_{(i)}$ is obtained by

$$m_{(i)} = \begin{cases} 1 - 0.5^{(1/n)}, i = 1 \\ \frac{(i - 0.3175)}{(n + 0.365)}, 1 < i < n \\ 0.5^{(1/n)}, i = n \end{cases}$$

upon which the estimated median values of the theoretical order statistics can be obtained using the transformation $M_{(i)} = \Phi^{-1}(m_{(i)})$. The correlation coefficient r is then defined as

$$r = \frac{\sum_{i=1}^{n} x_{(i)}.M_{(i)}}{\sqrt{\sum_{i=1}^{n} M_{(i)}^2} \sqrt{\sum_{i=1}^{n} x_{(i)}^2}} \quad \ldots \quad (2.13)$$

leading to the rejection of the normality hypothesis of the data for small values of r.

2.4 Other tests

2.4.1 The Gel-Miao-Gastwirth test

Gel, Miao and Gastwirth (2007) have recently proposed a directed normality test, which focuses on detecting heavier tails and outliers of symmetric distributions. The test is based on the ratio of the standard deviation and on the robust measure of dispersion Jn defined in equation (2). The normality test statistic R_{sJ} is therefore given by

$$R_{sJ} = s/J_n \quad \ldots \quad (2.14)$$

which should tend to one under a normal distribution. According to Gel, Miao and Gastwirth (2007) , the normality hypothesis of the data is rejected for large values of R_{sJ} and the statistic $\sqrt{n}\,(R_{sJ} - 1)$ is seen to asymptotically follow the normal distribution $N(0; \frac{\pi}{2} - 1.5)$. However, it has been empirically found that rejecting the normality hypothesis using a two-sided test extends the range of application of this test, namely to light-tailed distributions, without a significant reduction of its power towards heavy-tailed distributions. Given its enhanced behaviour, the two-sided test is the primary choice for the proposed study.

3. SIMULATION STUDY

To study the empirical level and power of twelve tests statistics we have generated sample from different distributions. The study was carried out for six different (n = 10,20,25,30,50 and 100) sample sizes and considering significance levels 0.10,0.05 and 0.01 (for 1 percent level not shown in table due to space) considering the

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

alternative of non-normal symmetric, asymmetric and contaminated normal (80% observations from N(0,1) and the remaining 20% from N(0,3)) distributions. Results obtained are shown in different tables given below. Here, normal observations are generated using Box-Muller (1958) formula and for the other distributions, method of inverse integral transformations is used. For each result 10,000 repetitions are made. The ratio of number of test statistic value greater than critical value divided by the total number of repetition gives the empirical level of test statistic under null case and power of the test statistic under the alternative hypothesis.

4. RESULTS

Table 1(a) Empirical levels of test under Normal (0,1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.1062	.0532	.1025	.0521	.1010	.0501	.1011	.0510	.1066	.0527	.0996	.0495
20	.1040	.0541	.1043	.0510	.1026	.0505	.1018	.0495	.1089	.0503	.1045	.0486
25	.1034	.0542	.0976	.0518	.0922	.0506	.0970	.0500	.0966	.0490	.1044	.0533
30	.1032	.0545	.0984	.0526	.1022	.0499	.0969	.0501	.1026	.0518	.0953	.0534
50	.1039	.0529	.0960	.0450	.0951	.0461	.0913	.0443	.0951	.0504	.1028	.0544
100	.1061	.0568	.0970	.0525	.1012	.0514	.0916	.0450	.1014	.0530	.0965	.0544

Table 1(b) Empirical levels of tests under Normal (0,1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.0853	.0450	.0731	.0554	.1032	.0421	.1032	.0449	.1470	.0782	.0953	.0450
20	.0890	.0480	.0826	.0623	.0993	.0443	.0935	.0447	.1195	.0555	.1000	.0530
25	.0910	.0490	.0837	.0633	.1002	.0497	.0984	.0477	.1090	.0506	.0989	.0502
30	.0930	.0490	.0865	.0657	.0995	.0478	.0936	.0491	.0995	.0464	.0969	.0484
50	.0970	.0510	.0847	.0584	.1002	.0497	.0905	.0485	.0936	.0378	.0979	.0473
100	.0720	.0460	.0887	.0611	.0989	.0503	.1000	.0485	.0847	.0318	-	-

Table 2(a) Empirical power of tests under Cauchy (0,1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.6639	.5928	.7251	.6418	.8565	.8023	.8919	.8697	.6621	.5892	.7026	.6329
20	.8847	.8463	.9090	.8450	.9731	.9554	.9829	.9776	.9064	.8120	.9424	.8753
25	.9355	.9105	.9537	.9104	.9880	.9804	.9942	.9910	.9262	.8830	.9552	.9363
30	.9619	.9458	.9784	.9494	.9963	.9924	.9981	.9975	.9566	.9290	.9755	.9637
50	.9966	.9944	.9993	.9964	.9999	.9998	1.000	.9999	.9977	.9928	.9989	.9984
100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2(b) Empirical power of tests under Cauchy (0,1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.4850	.4370	.6839	.6554	.5463	.4928	.6988	.6098	.7677	.6632	.6534	.5878
20	.8438	.8188	.9218	.9097	.8884	.8630	.9098	.8736	.9481	.9004	.8911	.8614
25	.9144	.8958	.9622	.9539	.9444	.9303	.9500	.9276	.9731	.9476	.9394	.9197
30	.9509	.9396	.9809	.9768	.9740	.9663	.9728	.9568	.9886	.9739	.9596	.9458
50	.9962	.9941	.9992	.9989	.9993	.9990	.9978	.9962	.9998	.9989	.9952	.9928
100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	-	-

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

Table 3(a) Empirical power of tests under Logistic (0,1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.1350	.0772	.6158	.4851	.6740	.5562	.7439	.6897	.2735	.1865	.1567	.0893
20	.1465	.0849	.8031	.6942	.8672	.8004	.9101	.8719	.3892	.2817	.2155	.1394
25	.1588	.0922	.8640	.7700	.9203	.8722	.9488	.9269	.4406	.3246	.2424	.1666
30	.1648	.0973	.9069	.8344	.9537	.9192	.9680	.9549	.4849	.3669	.2601	.1881
50	.1967	.1241	.9823	.9587	.9945	.9897	.9960	.9942	.6325	.5201	.3488	.2597
100	.2526	.1621	.9998	.9994	1.000	1.000	1.000	1.000	.8591	.7860	.4632	.3780

Table 3(b) Empirical power of tests under Logistic (0,1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.0375	.0250	.1259	.1027	.1174	.0581	.1555	.0773	.2198	.0947	.1344	.0749
20	.1188	.0912	.2078	.1727	.1617	.1048	.2147	.1363	.2504	.1187	.1797	.1177
25	.1509	.1177	.2339	.1971	.1801	.1203	.2352	.1566	.2625	.1235	.1932	.1272
30	.1821	.1458	.2609	.2240	.2001	.1377	.2602	.1779	.2851	.1386	.1991	.1312
50	.2688	.2235	.3507	.3004	.2689	.1974	.3294	.2470	.3471	.1749	.2017	.1353
100	.4336	.3763	.5023	.4470	.4196	.3238	.4727	.3806	.4789	.2605	-	-

Table 4(a) Empirical power of tests under Double Exponential (0,1)Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.2254	.1475	.2746	.1759	.3839	.2735	.4707	.4004	.2227	.1446	.2654	.1804
20	.3166	.2193	.3308	.2106	.5170	.3958	.6295	.5484	.3883	.2773	.4249	.3155
25	.3646	.2641	.3574	.2274	.5651	.4532	.6806	.6113	.4420	.3410	.4842	.3763
30	.4050	.3067	.3802	.2509	.6275	.5042	.7291	.6618	.5069	.3694	.5301	.4337
50	.5596	.4448	.4898	.3722	.7950	.7002	.8612	.8184	.6856	.5929	.6989	.6059
100	.8151	.7191	.7244	.5650	.9639	.9322	.9775	.9665	.9178	.8703	.8909	.8403

Table 4(b) Empirical power of tests under Double Exponential (0,1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.0897	.0587	.2435	.2082	.1733	.1154	.2741	.1633	.3706	.2038	.2216	.1460
20	.2614	.2103	.4270	.3824	.3514	.2756	.4083	.3040	.5216	.3250	.3483	.2628
25	.3311	.2767	.4981	.4502	.4346	.3549	.4604	.3574	.5858	.3843	.3904	.3024
30	.3869	.3300	.5584	.5096	.5037	.4185	.5078	.4011	.6486	.4381	.4216	.3252
50	.5693	.5124	.7263	.6800	.7157	.6362	.6513	.5548	.8006	.6198	.4932	.3992
100	.8193	.7731	.9212	.8939	.9406	.9071	.8617	.7965	.9590	.8769	-	-

Table 5(a) Empirical power of tests under Exponential (λ =1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.4226	.3059	1.000	.9867	1.000	1.000	1.000	.9966	.3609	.2662	.5376	.4097
20	.7065	.5899	1.000	1.000	1.000	1.000	1.000	1.000	.6055	.5075	.8695	.7839
25	.8036	.7040	1.000	1.000	1.000	1.000	1.000	1.000	.6945	.6017	.9369	.8840
30	.8788	.7967	1.000	1.000	1.000	1.000	1.000	1.000	.7569	.6791	.9709	.9457
50	.9828	.9608	1.000	1.000	1.000	1.000	1.000	1.000	.9169	.8747	.9998	.9985
100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.9960	.9911	1.000	1.000

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

Table 5(b) Empirical power of tests under Exponential (λ =1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.1945	.1437	.3551	.3187	.1830	.1046	.4592	.3238	.3842	.2478	.5640	.4316
20	.5600	.4761	.6452	.5939	.2747	.2018	.8280	.7218	.5168	.3763	.9028	.8366
25	.7011	.6106	.7457	.6967	.3139	.2391	.9153	.8489	.5718	.4262	.9623	.9218
30	.8109	.7245	.8281	.7828	.3380	.2652	.9605	.9207	.6243	.4757	.9864	.9677
50	.9830	.9534	.9695	.9476	.4474	.3744	.9989	.9960	.7580	.6239	.9998	.9996
100	1.000	.9999	.9999	.9998	.6461	.5785	1.000	1.000	.9225	.8490	-	-

Table 6(a) Empirical power of tests under Lognormal (0,1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.5858	.4798	1.000	1.000	1.000	1.000	1.000	1.000	.5347	.4486	.6835	.5825
20	.8640	.7985	1.000	1.000	1.000	1.000	1.000	1.000	.8166	.7581	.9461	.9090
25	.9318	.8858	1.000	1.000	1.000	1.000	1.000	1.000	.8903	.8433	.9801	.9627
30	.9651	.9345	1.000	1.000	1.000	1.000	1.000	1.000	.9331	.9046	.9924	.9850
50	.9986	.9955	1.000	1.000	1.000	1.000	1.000	1.000	.9895	.9837	.9998	.9998
100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.9999	1.000	1.000

Table 6(b) Empirical power of tests under Lognormal (0,1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.3472	.2894	.5219	.4897	.2685	.2047	.6165	.4983	.5503	.4681	.6981	.5979
20	.7808	.7240	.8354	.8074	.4865	.4215	.9257	.8815	.7660	.6931	.9602	.9295
25	.8871	.8422	.9079	.8856	.5692	.5112	.9739	.9515	.8262	.7653	.9860	.9747
30	.9475	.9122	.9531	.9369	.6401	.5837	.9902	.9806	.8750	.8262	.9954	.9908
50	.9982	.9955	.9971	.9944	.8178	.7817	.9997	.9997	.9596	.9404	.9999	.9998
100	1.000	1.000	1.000	1.000	.9685	.9574	1.000	1.000	.9989	.9967	-	-

Table 7(a) Empirical power of tests under Uniform (-1,1) Distribution

Sample size(n)	Test Statistics											
	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.1321	.0692	.0566	.0180	.2071	.0839	.1769	.0401	.0911	.0425	.1084	.0427
20	.1886	.1031	.1863	.0564	1.000	.9600	.9935	.6146	.2009	.0978	.1755	.0695
25	.2196	.1222	.3211	.1020	1.000	1.000	1.000	.9909	.2406	.1303	.2210	.0928
30	.2581	.1510	.5078	.1812	1.000	1.000	1.000	1.000	.3544	.2216	.2731	.1372
50	.4189	.2760	.9969	.8157	1.000	1.000	1.000	1.000	.6955	.5541	.6505	.4493
100	.7605	.6070	1.000	1.000	1.000	1.000	1.000	1.000	.9786	.9599	.9845	.9486

Table 7(b) Empirical power of tests under Uniform (-1, 1) Distribution

Sample size(n)	Test Statistics											
	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.0036	.0019	.0255	.0186	.2074	.0935	.1174	.0596	.0496	.0221	.1721	.0784
20	.0014	.0008	.0061	.0040	.3698	.2127	.1933	.0915	.0394	.0031	.3588	.2045
25	.0010	.0001	.0031	.0019	.4521	.2832	.2577	.1336	.0915	.0089	.4729	.2935
30	.0010	.0000	.0022	.0013	.5269	.3523	.3245	.1768	.1402	.0166	.6182	.4189
50	.0271	.0002	.0008	.0003	.7703	.6222	.6586	.4509	.4415	.1567	.9473	.8592
100	.9126	.5626	.5002	.0430	.9761	.9374	.9874	.9530	.9015	.7040	-	-

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

Table 8(a) Empirical power of tests under Contaminated Normal Distribution

Sample size(n)	K-S		AD		Z_A		Z_C		D		r	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.7477	.6788	.3271	.2037	.4578	.3410	.5825	.5157	.3039	.2184	.3588	.2633
20	.9564	.9323	.4005	.2599	.6483	.5250	.7546	.6953	.5225	.4273	.5700	.4727
25	.9825	.9708	.4453	.3003	.7054	.6040	.8127	.7564	.5930	.5048	.6417	.5472
30	.9935	.9890	.4789	.3270	.7760	.6783	.8608	.8151	.6677	.5827	.7023	.6232
50	.9999	.9996	.6393	.4720	.9196	.8628	.9526	.9327	.8341	.7737	.8536	.8023
100	1.000	1.000	.8878	.7740	.9954	.9908	.9979	.9966	.9759	.9608	.9730	.9577

Table 8(b) Empirical power of tests under Contaminated Normal Distribution

Sample size(n)	JB		RJB		BS(T_w)		DH		GMG(R_{SJ})		W	
	α =.10	.05	.10	.05	.10	.05	.10	.05	.10	.05	.10	.05
10	.1610	.1176	.3313	.2952	.2284	.1654	.3573	.2367	.4233	.3170	.1072	.2227
20	.4401	.3852	.5732	.5350	.4521	.3795	.5706	.4705	.5995	.4883	.2620	.4136
25	.5272	.4708	.6439	.6051	.5280	.4565	.6409	.5432	.6467	.5387	.3101	.4693
30	.6168	.5617	.7149	.6791	.6000	.5294	.7114	.6230	.7047	.6061	.3512	.5079
50	.8030	.7597	.8597	.8302	.7795	.7257	.8518	.7887	.8377	.7575	.4590	.6094
100	.9670	.9546	.9769	.9691	.9576	.9358	.9748	.9586	.9666	.9390	-	-

5. DISCUSSIONS

Table 1(a) and Table 1(b) show the empirical level of twelve tests for six different sample sizes. It is seen that all the test statistics almost satisfy its nominal levels. However, the K-S, D and R_{sJ} are found to be anticonservative in few situations and JB and RJB are found to be conservative in all sample sizes. But as the sample sizes increases, empirical levels of the entire test statistics come closer to nominal levels.

Table 2(a) and Table 2(b) above show the empirical power of twelve tests under the alternative of the Cauchy distribution. It is seen that power of all the tests increases as the sample sizes increases. Out of the twelve tests, power of Z_A, Z_C and R_{sJ} tests seems to be more than the other tests. But as the sample sizes increases power of all the tests come closer to each other and finally come to exactly equal to one.

Table 3(a) and Table 3(b) above display empirical power of all tests mention in section 2. Under the alternative of Logistic distribution. In this case also empirical power of Z_A and Z_C seem to be higher than the other tests. Power of Anderson-Darling test (AD) is also quite good and higher than the other tests except Z_A and Z_C and closed to Z_A and Z_C in large sample cases. It is seen that Empirical power of JB test is the lowest of all. However, empirical power of all the tests increases as the sample sizes increases.

Table 4(a) and 4(b) depict the empirical power of tests under the alternative of Double Exponential distribution. Here, along with the Z_A and Z_C tests power of R_{sJ} are more than the other tests. Power of AD, r, RJB, DH and D are less than above three tests but found to be more than remaining tests. Here also, empirical power of all the tests increases as the sample sizes increases.

Table 5(a) and 5(b) show the empirical power of tests under the alternative of Exponential distribution. Here we have seen that empirical power of Z_A, Z_C and AD are exactly equal and even in small sample sizes. Power of K-S, D, r, RJB, DH, Rs_J and W are more or less similar but less than above three tests and higher than BS and JB tests.

Table 6(a) and 6(b) show the power of tests under the alternative of Lognormal distribution. It is observed that power of all the tests is similar as the exponential distribution. That is empirical power of Z_A,Z_C and AD are higher than others and BS and JB are exhibit lowest power and other remaining tests lies in between these two groups. But in large sample cases empirical power of all the tests are found to be almost equal.

Table 7(a) and 7(b) show the empirical power of tests under the alternative of Uniform distribution. Here we have seen that empirical power of Z_A, $Z_{C,}$ BS and W are higher than the other tests. The empirical power of the JB test seems to be the lowest of all the tests. Here also, empirical power of all the tests increases as the sample sizes increases.

Table 8(a) and 8(b) show the empirical power of tests under the alternative of Modified Normal distribution. Here we have seen that empirical power of the K-S test seems to be higher than the other tests.

The empirical power of r, RJB and DH are more or less similar but less than Z_A, Z_C and R_{SJ}. The empirical power of W test seems to be the lowest of all the tests. Here also, the power of the tests increases as the sample sizes increases.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

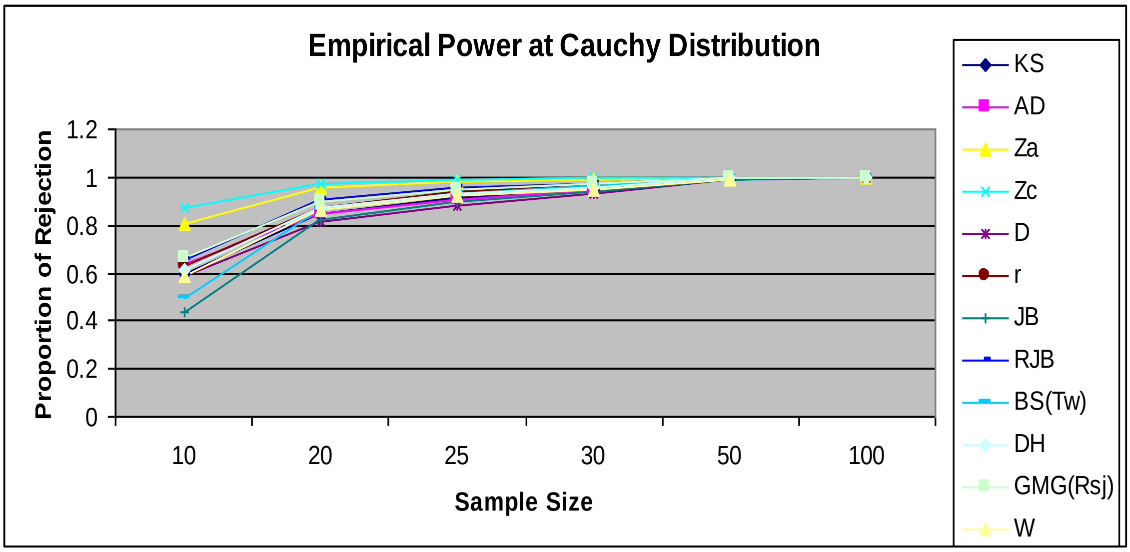

Fig.1 Empirical Power of test under Cauchy (0, 1) Distribution (for α=.05)

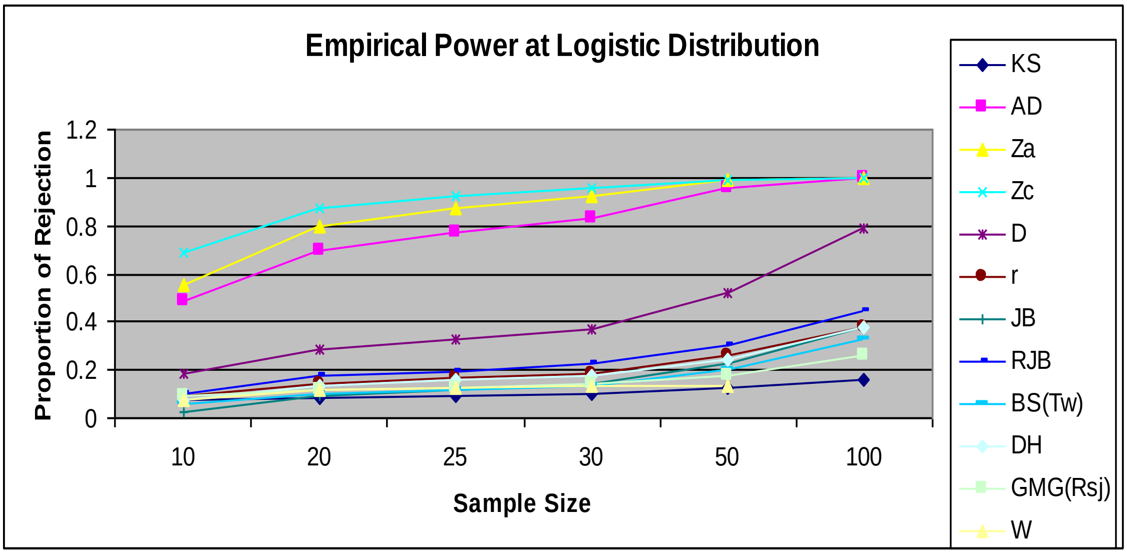

Fig.2 Empirical Power of test under Logistic (0,1)Distribution (for α=.05)

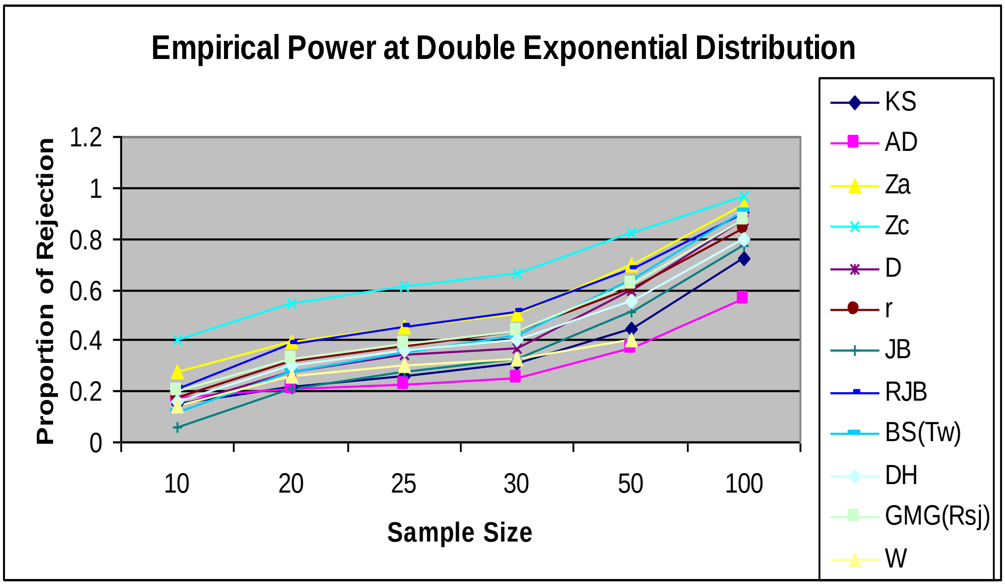

Fig.3 Empirical Power of test under Double Exponential (0,1)Distribution (for α=.05)

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

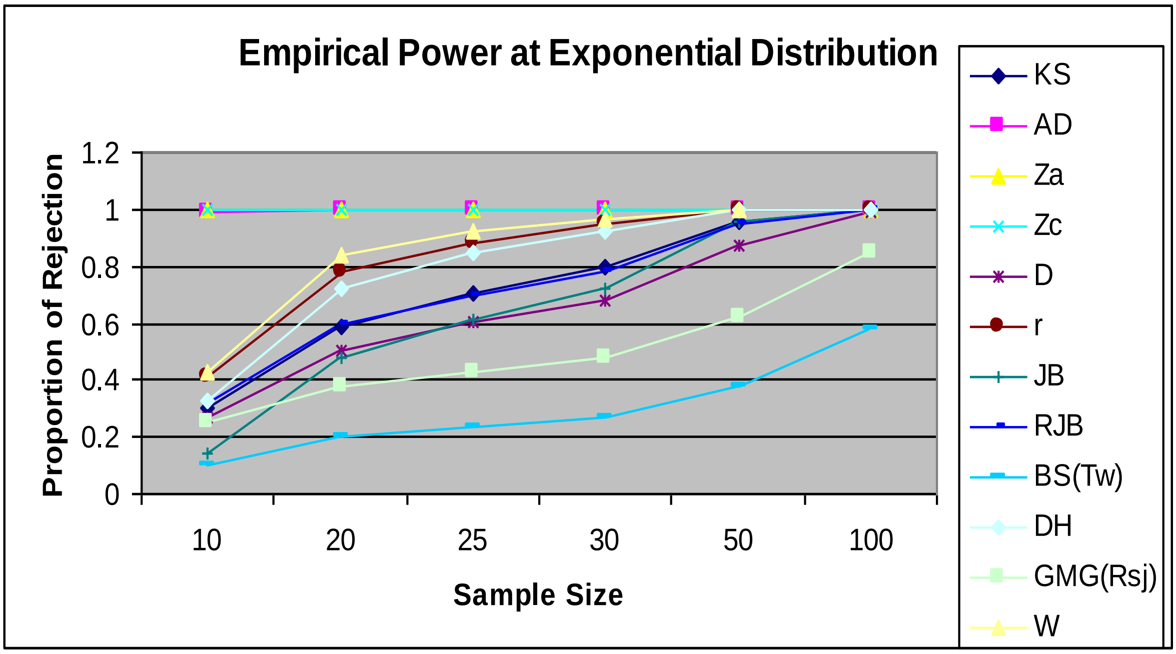

Fig.4 Empirical Power of test under Exponential (λ=1) Distribution (for α=.05)

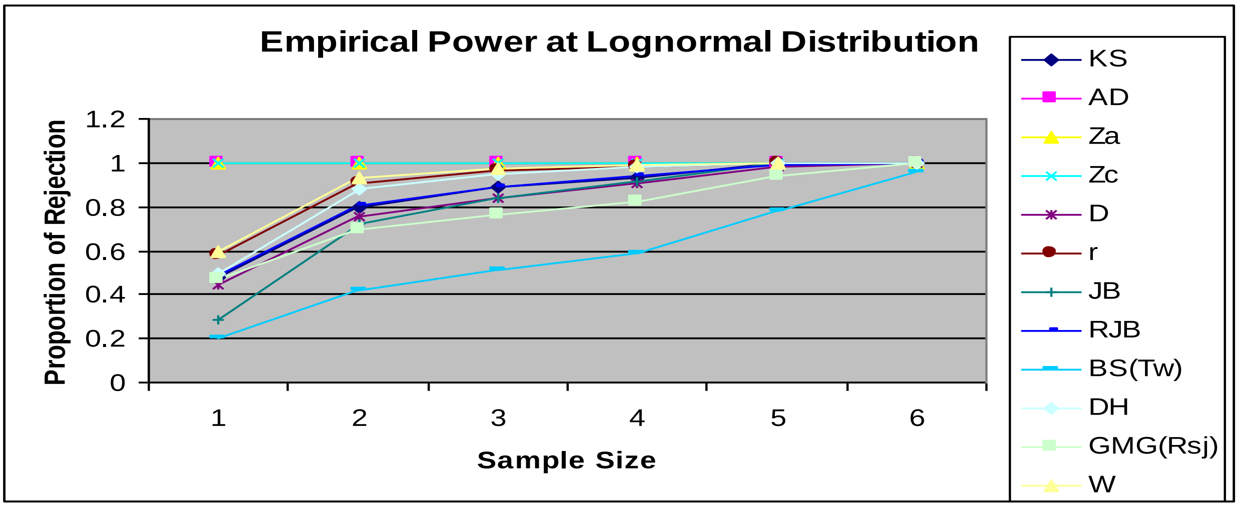

Fig.5 Empirical Power of test under Lognormal (0,1) Distribution (for α=.05)

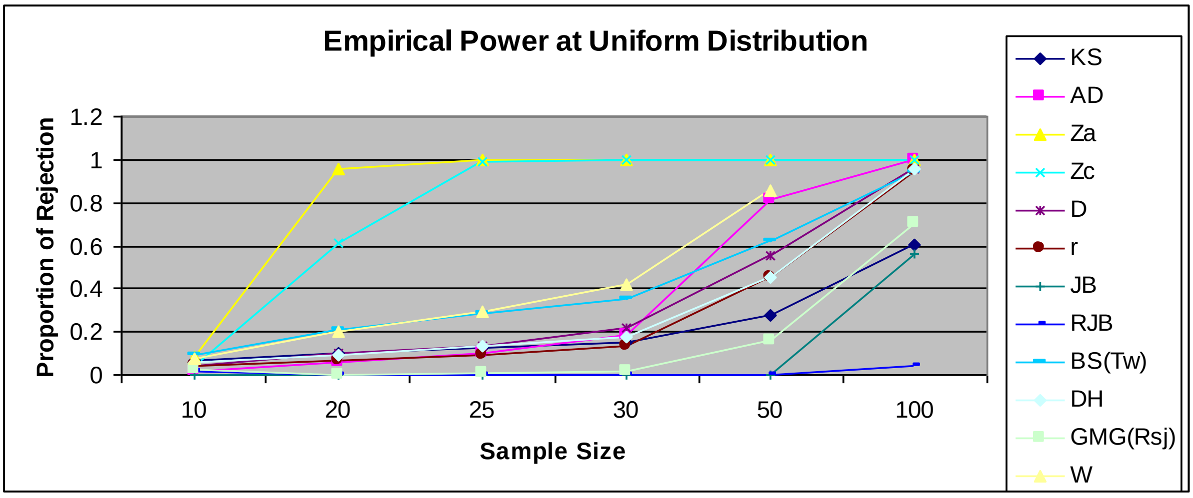

Fig.6 Empirical Power of Uniform (-1,1) Distribution (for α=.05)

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 6, June 2016

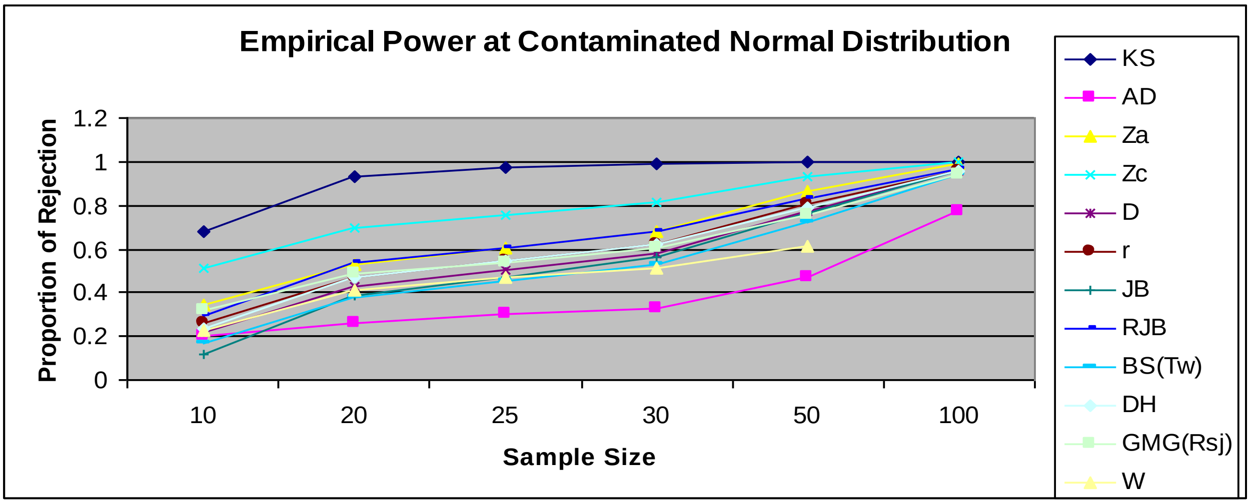

Fig.7 Empirical Power of Contaminated Normal Distribution (for α=.05)

6. CONCLUSION

Power of Z_A and Z_C are found to be more than the other tests for both symmetrical and asymmetrical alternative. Gel-Miao-Gastwirth (R_{sJ}) test also exhibits almost equal power as the Z_A and Z_C tests in large sample cases under the alternative of Cauchy distribution. Anderson-Darling test is same as Z_A and Z_C in case of asymmetrical distribution but shows lower power in case of alternative of symmetrical distribution. Over all power of Filliben correlation test r is good, although its power is less than Z_A and Z_C. Power of Kolmogorov-Sminov test is found to be higher than the other tests only for Contaminated Normal distribution. Tests JB and BS (except for Uniform distribution) show less power in all situations discussed here. Finally, we arrive at the conclusion that Z_A and Z_C may be recommended for all situations. We may give second preference to Filliben correlation test(r) and Anderson–Darling test (AD) for test of normality.

REFERENCES

[1] Anderson,T.W. Darling, D.A. (1952): Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes, Ann. Math. Stat. 23(2), 193-212.
[2] Bonett,D.G. and Seier,E.(2002):A Test of Normality with high uniform power,Comp.Stat.,Data Anal.,40(3),435-445.
[3] Box,G.E.P. and Muller, M.E.A. (1958): A Note on Generation of Normal Deviates, Ann.Math. Stat. 29,610-611.
[4] D’Agostino,R.B.(1971):An omnibus test of normality for moderate and large size samples, Biometrika,58(2)341-348.
[5] Doornik, J.A. and Hansen,H.(1994): A omnibus test for univariate and multivariate normality, Nuffield College,Oxford.
[6] Farrel, P.J. and Stewart, K.R.(2006):Comprehensive Study of Tests for Normality and Symmetry: Extending the Spiegelhalter Test, Journal of Computation and Simulation,76(9),803-816.
[7] Filliben,J.J.(1975):The probability plot correlation coefficient test for normality, Technometrics 17(1), 111-117.
[8] Gel,Y.R. and Gastwirth,J.L.(2008): A robust modification of Jarque-Bera test of normality, Econom. Lett. 99(1), 30-32.
[9] Jarque,C. and Bera, A. (1980): Efficient test for normality ,homoscedasticity and serial independence of regression residuals, Econom. Litt.,6(3),255-259.
[10] Min, I. (2007): A nonparametric test of the conditional normality of housing demand, Appl.Econ. Lett 14(2),105-109.
[11] Shapiro, S.S. and Wilk, M.B. (1965): An analysis of variance test for normality (complete samples), Biometrika, 52(3/4), 591-611.
[12] Stephens,M.A.(1969):Test for Normality, Technical Report No. 152,Dept. of Statics, Stanford University.
[13] Stephens, M.A.(1974): EDF statistics for goodness of fit and some comparisons, Jour.Amer. Stat. Assoc. 69(347), 730-737.
[14] Yazici,B. and Yolacan,S.(2007):A Comparison of Various Tests of Normality, Journal of Statistical Computation and Simulation, 7, 2,1 75-183.