Comparison of Some Normality Tests Under Different Conditions

By
Hamman Kenas Labanbi
Benham Zangaluka Reuben
Danjuma Jibasen
&
Emmanuel Torsen
Department of Statistics
Modibbo Adama University, Yola, Nigeria

Abstract

This study compared normality tests in terms of power of a test and type 1 error rate. Ten normality tests were chosen for the comparison; Shapiro-Wilk (SW), Shapiro Francia (SF), D'Agostino D(DA), Robust Jarque-Bera (RJB), Modified Jarque-Bera (MJB), 1st Hosking (HK1), 1st Cabana Cabana (CC1), 2nd Zhang Wu (ZW2), Chen Shapiro (CS)and Gel-Miao-Gaswirth (GMG). In terms of power of a test, data were simulated from R version 4.3.1 The distributions were categorized into symmetric distributions (Laplace, Uniform and Logistic distribution) and asymmetric distributions (Beta, Gamma and Weibull distribution). The results for the two categories of distributions were computed, for symmetric distributions, the most powerful tests were Gel-Miao-Gaswirth test followed by Modified Jarque-Bera tets and Rubost Jarque-Bera test: For asymmetric distributions, the most powerful tests were 2nd Zhang-Wu test followed by 1st Cabana Cabana test and Shapiro Wilk test. In terms of type 1 error rate, data was simulated from Standard Normal Distribution, results shows that RJB has highest type 1error rate at $n \leq 100$ *which is followed closely by HK1 at sample sizes* $n \leq 40$*, most of the normality tests do not show any consistent pattern with changes in the sample size. It is therefore recommended that for symmetric distributions GMG, followed by MJB and RJB tests should be used while for asymmetric distributions ZW2 followed by CC1 and SW should be used while considering the sample sizes. Software engineers should include GMG, MJB, RJB, ZW2 and CC1 tests in stastistical software where the tests are not available.*

Keywords: *Normality test, Power of a test, Type 1 error.*

Introduction

Most statistical procedures require assessing the assumption of normality. Many of the statistical procedures including correlation, regression, one or two sample means test of hypothesis and analysis of variance (parametric tests), rely on the assumption that the data follow a normal distribution, that is, it is assumed that the population from which the samples are taken are normally distributed (Das & Imon, 2016). When carrying out statistical analysis using parametric methods, the assumption of normality is a fundamental concern for the analyst, therefore testing the normality of the data should be the initial step in any analysis that relies on the assumption of normality (Demir,2022).

When normality assumptions are not fulfilled, researchers use data transformation (such as the log or square root) or outlier's detection procedures to normalize distributions. When such techniques are ineffective, it is recommended that non-parametric test be employed (Khatun, 2021). Ignoring normality and other assumptions can compromise the accuracy and reliability of conclusions drawn about reality (Das & Imon, 2016).

According to Abiodun et al. (2022), there are three fundamental methods for assessing the normality assumption. The most straight forward approach is employing graphical method. The normal quantile-quantile plot (QQ plot) stands out as widely preferred and efficient diagnostic tool for assessing the normality of the data. Other common graphical methods that can be used to assess the normality assumption include histogram, box-plot and stem-and-leaf plot. While graphical methods can be valuable for assessing normality with n independent observations, they alone are insufficient to definitively confirm the normality assumption, subjectivity issues arise as its challenging to gauge the strength and extend of normality solely from these plots. What appears 'normal' to one analyst may not be perceived as such for others. Experience and good statistical knowledge are needed in order to interpret the graph. To support the graphical methods, it's essential to conduct a more formal method which is the numerical methods and formal normality tests before making any conclusion about the normality of the data. Judgement on the normality of the data will be much improved by combining the graphical methods, numerical methods and formal normality tests. The numerical methods include the skewness and kurtosis coefficients whereas for the normality tests, there are various procedures available for testing the assumption of normality (Ukponmwan&Ajibade,2017). Multiple studies have compared various normality tests to assess their performance under different conditions. With the emergence of additional normality tests, there is continual requirement to assess and compare these new tests against the established once to determine their effectiveness. This study compares the outcomes of various comparisons of normality, the tests that are previously identified as the most powerful along with a recently proposed test, to ascertain the most powerful among them.

Literature Review

Many attempts have been made in order to carry out comparison of normality test; Ukponnmwa and Ajibade (2017) compared nine normality tests; W/S, Jaque Bera, Adjusted Jaque Bera, D'Agostino, Shapiro Wilk, Shapiro Francia, RyanJoiner, Lilliefors'and Anderson Darling test statistics, with a view to determining the effectiveness of the techniques to accurately determine whether a set of data is from normal distribution or not. Simulated data of sizes 5, 10, …, 100 was used for the study, it is concluded that D'Agostino D is uniformly most powerful and SW should be replaced by SF because the latter is more powerful. Hernandez (2021) examined several power comparisons between normality tests published in the last 3 decades which was used to rank 55 of the most common methods. The overall winner of this analysis was the regression-based Shapiro-Wilk (SW) normality test. Rana et al. (2021) proposed a new test Modified Jarque-Bera test, simulation experiment was used to compare the test with 6 other tests; Jarque-Bera (JB) test, Robust Jarque-Bera (RJB) test, Shapiro Wilk (SW) test, Kolmogorov Smirnov (KS) tests, Lilliefors (LF) and Anderson Darling (AD), using different sample sizes, it was concluded that MJB is the most powerful. Uyanto (2022) compared the power of 50 normality tests using simulation procedure at different sample sizes and concluded that, for symmetric distribution RJB and GMG are the most powerful while for asymmetric distribution, CC1 and ZW2 are the most powerful. Arnastausk aite, et al. (2021) used simulation procedure to examined the performance of 31 normality tests, at different sample sizes and concluded that the most powerful test for the groups of symmetric, asymmetric and modified normal distributions was 1st Hosking test (for smaller sample size) and N-metric test (for larger sample sizes).

Methodology

The data in this research work was simulated from standard normal, Laplace, Uniform, Logistics, Beta, Gamma and Weibull distribution at different sample sizes using R-version 4.3.1.Ten normality tests; Shapiro-Wilk, Shapiro-Francia, Robust Jarque-Bera, D'Agostino D, Modified Jarque-Bera, 1st Hosking, 1st Cabanna-Cabanna, 2nd Zhang Wu, Chen-Shapiro and Gel-Miao-Gaswirth test were studied. Simulation procedure was used to evaluate the power of SW, SF, RJB, DA, MJB, HK1, CC1, ZW2, CS and GMG test statistics in testing if a random variable of n independent observations come from a population with normal distribution. The distributions were categorised into two; symmetric distributions uniform$(0,1)$, Laplace$(0,1)$ and Logistic$(3,0)$, and asymmetric distributions Weibull$(2,1)$, Gamma$(3,2)$ and Beta$(3,6)$ at different sample sizes $n = 10,20,30,40,50,60,80,100,200,500 \,\&1000.$ Data was also generated from standard normal distribution to evaluate the Type 1 error rate of the tests. The number of simulations $N = 5000.$

Shapiro-Wilk test

The Shapiro-Wilk test was developed by Shapiro and Wilk (1965) as cited in uyanto (2022). Given an ordered random sample, $y_1 < y_2 < \cdots < y_n$, the original Shapiro-wilk test statistics is define as

$$W = \frac{(\sum_{i=1}^{n} a_i \, y_i \,)^2}{\sum_{i=1}^{n}(y_i \, - \bar{y})^2} \tag{1}$$

Where y_i is the i^{th}order statistics

$\bar{y}$ Is the sample mean

$$a_i \; = (a_{1,} \dots , a_n \,) = \frac{M^T V^{-1}}{(M^T V^{-1} V^{-1})^{1/2}}$$

And $M = (m_{1,} \dots , m_n)^T$ are the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution and V is the covariance matrix of those order statistics the value of w lies between 0 and 1. Small values of w lead the rejection of normality where as a value of 1 indicates normality of the data.

Shapiro-Francia test

The Shapiro-Francia test, a simplified version of Shapiro-Wilk test, was developed in 1972 Shapiro and Francia (1972) as cited in Lee et.al,(2016). The test statistic is defined as

$$T_{SF} = \frac{(\sum_{i=1}^{n} b_i y_i)}{\sum_{i=1}^{n}(y_i \, - \bar{y})^2} \tag{2}$$

where the constants b_i are given by $(b_i, \dots , b_n) = \frac{\mu^T}{(\mu^T \mu)^{1/2}}$ and μ is the vector of expected values of standard normal order statistics.

Chen-Shapiro test

(Chen and Shapiro,1995) is based on the normalized spacing and defined as

$$CS = \frac{1}{(n-1)s}\sum_{i=1}^{n-1}\frac{X_{(i+1)} - X_{(i)}}{M_{(i+1)} - M_{(i)}} \qquad (3)$$

In which the $M_{(i)}$ is the i^{th} quantile of a standard normal distribution obtained by $\Phi^{-1}\frac{(i-0.375)}{(n+0.025)}$ the normality hypothesis of the data is rejected for small values of CS.

D'Agostino D Test

According to D'Agostino and Stephen (1986) as cited in Ukponmwan & Ajibade(2017) the test statistic can be used to detect departure of data set from normality. It requires that the observations be ordered in ascending order and the mean deviations of the ordered data, used to compute $D_{calculated}$ which was compared with the range/interval of values from table values of D($D_{tabulated}$)The null hypothesis is rejected if the$D_{calculated}$value falls outsid e the range of values from the D'Agostino critical values, otherwise, it is accepted. The test statistic is;

$$D = \frac{T}{\sqrt{n^3 SS}} \qquad (4)$$

$$T = \sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)x_i \qquad (5)$$

$SS = \sum_{i=1}^{n}(x_i - x)^2$, x_i is the i^{th} observationand $n = sample\ size$. The critical value of the test is obtained from the Critical Value for D'Agostino D normality test.

Modified Jarque-Bera Test

Bowman and Shenton (1975) as cited in Rana et al. (2021) proposed the Jarqua-Bera test which was subsequently derived by Jarque and Bera (1987) as cited in Rana et al. (2021). According to Jarque and Bera Jarque and Bera (1987) as cited in Rana et al. (2021) the Jarqua-Bera test has optimum asymptotic power properties and good finite sample performance. Based on the sample skewness and kurtosis the Jarqua-Bera statistic is given by

$$JB = n\left[\frac{\sqrt{(b_1)}^2}{6} + \frac{(b_2-3)^2}{24}\right] \qquad (6)$$

Under the normality, the Jarqua-Bera test statistic follows a chi-squared distribution with two degrees of freedom. A significantly large value of Jarqua-Bera leads to the rejection of the normality assumption. Following the idea of Gel and Gastwirth (2008) as cited in Rana et al. (2021), the robust version of the second central moment is used in the proposed modified Jarque–Bera test.Normalised Median Absolute Deviation(MAD),is defined as

$$\varphi^2 = MAD = 1.4826^2[median\{|X_i - median(X_i)|\}]^2 \qquad (7)$$

The JB statistic is based on sample skewness $\frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}}$ and kurtosis $\frac{\hat{\mu}_4}{\hat{\mu}_2^2}$. However, the modified sample estimates of skewness $\frac{\hat{\mu}_3}{\phi^3}$ and kurtosis $\frac{\hat{\mu}_4}{\phi^4}$ respectively, which leads to the Modified Jarque-Bera (MJB) test statistic as given below.

$$MJB = \frac{n}{6}\left(\frac{\hat{\mu}_3}{\varphi^3}\right)^2 + \frac{n}{216}\left(\frac{\hat{\mu}_4}{\varphi^4}\right)^2 \quad \text{(Rana et al.,2021).} \qquad (8)$$

Robust Jarque–Bera test

In 2007, Gastwirth et al.modified the Jarque–Bera test and got a more powerful Jarque–Bera test.RJB test statistic is defined as:

$$RJB = \frac{n}{6}\left(\frac{m_3}{J_n{}^3}\right)^2 + \frac{n}{64}\left(\frac{m_4}{J_n{}^4} - 3\right)^2 \quad (9)$$

RJB where m_3 , m_4 are the third and fourth moments, respectively, and J_n is the ratio of the standard deviation.

Cabana-Cabana test(CC1-CC2)

Cabana and Cabana (1994) as cited in Arnastauskaite et al. (2021) proposed the CC1 and CC2 tests. The$CC1(T_{S,l)}$ and $CC2(T_{K,l)}$, respectively are defined as;

$T_{S,l} = \max\left|W_{S,l}(x)\right|$, $T_{K,l} = max\left|W_{k,l}(x)\right|$ where $W_{S,l}(x)$ and $W_{K,l}(x)$ approximate the transform estimated empirical sensitive to change in skewness and kurtosis and define as;

$$W_{S,l}(x) = \Phi(x).\overline{H_3} - \phi(x).\sum_{J=1}^{l}\frac{1}{\sqrt{j}}H_{j-1}(x).\overline{H_{J+3}}, \quad (10)$$

$W_{K,l} = -\ \Phi(x).\overline{H_3} + [\Phi(x) - x.\phi(x)].$

$$\overline{H_4} - \ \phi(x).\sum_{J=2}^{l}\left(\sqrt{\frac{j}{j-1}}H_{j-2}(x).H_i(x)\right).H_{j+2} \quad (11)$$

Where L is a dimensionality parameter, $\Phi(x)$ is the probability density function of the standard normal distribution, $H_{j(.)}$ is the j^{th} order normalized Hermite polynomial, and $\overline{H_J}$ is the j^{th} order normalized mean of the Hermite polynomial defined as:

$$\overline{H_J} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}H_j(x_i)$$

The 1st Hosking test

In 1990, Hosking has shown the r_{th} order sample L- moment can be estimated by

$$l_r = \sum\nolimits_{k=0}^{r-1} p^*_{r-1,k}\, b_k \quad (12)$$

$$p^*_{r-1,k} = (-1)^{r-k}(r\mathrm{C}k)$$

$$b_k = \frac{1}{n}\sum\nolimits_{i=1}^{n}\frac{(i-1)(i-2)\ldots(i-k)}{(n-1)(n-2)\ldots(n-k)}x_{(i)}$$

Based on the second, third and fourth sample L-moments, which have similarities with the corresponding central moments, Hosking (Hosking ,1990) also defines new measures of skewness and kurtosis, termed L-skewness τ_3 and L-kurtosis τ_4 as follows

$\tau_3 = \frac{l_3}{l_2},\ \tau_4 = \frac{l_4}{l_2},$

The value of τ_3 is bounded between $-1\ to\ 1$ for all distributions and is close to zero for the normal distribution, while the value of τ_4 is ≤ 1 for all distributions and is close to 0.1226 for the normal distribution. Hosking has suggested that normality could be tested based on τ_3 and τ_4 according to the following statistic T_{Lmom}

$$T_{Lmom} = \frac{\tau_3 - \mu_{\tau_3}}{var(\tau_3)} + \frac{\tau_4 - \mu_{\tau_4}}{var(\tau_4)}$$

where μ_{τ_3} and μ_{τ_4} are the mean of τ_3 and τ_4, and $var(\tau_3)$ and $var(\tau_4)$ are their corresponding variances. Nonetheless, μ_{τ_3} and μ_{τ_4} are expected to be close to 0 and 0.1226.The normality hypothesis of the data is rejected for large values of T_{Lmom}.

Zhang–Wu ($ZW1$–Z_C, $ZW2$–Z_A)

In 2005, Zhang and Wu presented the $ZW1\ and\ ZW2$ goodness-of-fit tests. The Z_C and Z_A statistics are similar to the Cramer–von Mises and Anderson–Darling tests statistics based on the empirical distribution function. The statistic of the test is defined as:

$$Z_C = \sum_{i=1}^{n} \left[In \frac{\Phi(z_i)^{-1} - 1}{\frac{n - 0.5}{i - 0.75} - 1} \right]^2 \tag{13}$$

$$Z_A = -\sum_{i=1}^{n} \left[\frac{In\Phi(z_{(i)})}{n - i + 0.5} + \frac{In\left[1 - \Phi(z_{(i)})\right]}{i - 0.5} \right] \tag{14}$$

Where $\Phi(z_{(i)}) = {i - 0.5}/{n}$

Gel–Miao–Gastwirth test (GMG)

In 2007, Gastwirth, Miao & Gel proposed the GMG test. The statistic of the test is defined as:

$$R_{sJ} = \frac{s}{J_n} \tag{15}$$

Where J_n is the ratio of the standard deviation and the robust measure of dispersion is defined as:

$$J_n = \frac{\sqrt{\pi/2}}{n} \sum_{i=1}^{n} |x_i - M|, \tag{16}$$

Where M is the median of the sample.

Results and Discussion

Normality Test on Simulated Data (power of a test)

The power comparison of different normality tests on simulated data under symmetric and asymmetric distributions at different sample sizes are presented in Tables 1 and 2 and the plot of power of different normality tests in Figures 1 and 2

Table 1: Average power of normality test obtained for a group of symmetric distributions $at\ \alpha = 0.05$

	Simulated power									
n	SW	SF	RJB	DA	MJB	HK1	CC1	ZW2	CS	GMG
10	0.1044	0.1089	0.1016	0.0963	0.1554	0.1481	0.0773	0.1640	0.1546	0.1791
20	0.2002	0.1752	0.1741	0.1461	0.2728	0.2605	0.1337	0.2781	0.3066	0.2591
30	0.2961	0.2647	0.2440	0.1573	0.3530	0.4177	0.1969	0.4214	0.4211	0.3533
40	0.3945	0.3569	0.2675	0.1725	0.4598	0.5158	0.2564	0.4877	0.4881	0.4201
50	0.4903	0.4426	0.3074	0.1850	0.5415	0.5850	0.3172	0.5456	0.5410	0.4869
60	0.5675	0.5319	0.3727	0.1918	0.6148	0.6388	0.4072	0.5748	0.5611	0.5414
80	0.6370	0.6416	0.4200	0.2013	0.6809	0.6800	0.4405	0.6136	0.6051	0.5916
100	0.7007	0.7347	0.4637	0.2306	0.7431	0.7307	0.5512	0.6550	0.6870	0.7555
200	0.8030	0.8357	0.8730	0.3014	0.8551	0.8004	0.7673	0.7844	0.7330	0.9632
500	0.8842	0.9585	0.9703	0.3657	0.9870	0.9073	0.9475	0.9241	0.8929	1.0000

Table 1 shows that for sample sizes $(10,20,30,40,50,60,80\ \&100)$ the power of most tests was relatively low. At $n = 200$, GMG is the most powerful tests achieving a power of 0.9632, it was followed closely by RJB(0.8730),MJB(0.8551),SF(0.8357) and SW (0.8030), at $n = 500$ MJB, SF, RJB, MJB, CC1, ZW2 and HK1 have the highest power with power approaching 1 this is followed by CS(0.8929) and SW(0.8842), at $n = 1000$,GMG has the highest power reaching a power of 1, all tests have power approaching 1 except DA which only achieved a power of 0.4350. DA is the least powerful tests.

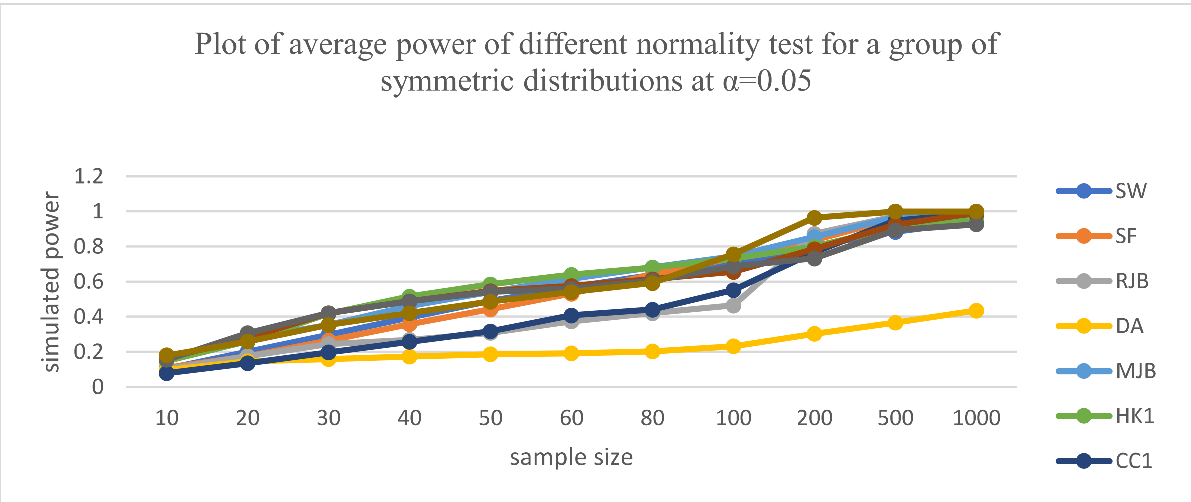

Fig 1: Plot of average power of different normality test for a group of symmetric distributions at α=0.05

Fig 1 shows clearly that GMG is the most powerful test followed closely by MJB, the least powerful is DA.

Table 2: Average Power of normality test obtained for a group of asymmetric distributions $at\ \alpha = 0.05$

Table 2 shows that for sample sizes $(10,20,30,\&40)$ the power of most tests was relatively

	Simulated power									
n	SW	SF	RJB	DA	MJB	HK1	CC1	ZW2	CS	GMG
10	0.1469	0.1439	0.1191	0.1357	0.1266	0.1192	0.1470	0.1709	0.1160	0.1615
20	0.3078	0.2938	0.2179	0.2797	0.2428	0.2785	0.2589	0.3831	0.3391	0.2127
30	0.4796	0.4299	0.3147	0.4157	0.3445	0.4642	0.4969	0.6183	0.5524	0.3070
40	0.6169	0.5518	0.3813	0.5331	0.4541	0.6398	0.5975	0.7669	0.6995	0.3606
50	0.7347	0.6656	0.4709	0.6189	0.5377	0.7200	0.7245	0.8417	0.7917	0.4160
60	0.8104	0.7519	0.5421	0.7145	0.6302	0.8079	0.8148	0.9111	0.8796	0.4649
80	0.9115	0.8693	0.6585	0.8208	0.6758	0.8697	0.8616	0.9631	0.9439	0.5036
100	0.9609	0.9335	0.7625	0.8839	0.8127	0.9362	0.9241	0.9854	0.9703	0.5378
200	0.9996	0.9994	0.9823	0.9903	0.9656	0.9811	0.9997	0.9999	0.9937	0.6610
500	1.0000	1.0000	1.0000	1.0000	0.9991	1.0000	1.0000	1.0000	1.0000	0.6711
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.8653

low.

At $n = 50$, ZW2 has the highest power with power value of 0.8417, at $n = 60$, ZW2 has the highest power with power value of 0.9111, followed by CC1,CS,SW and HK1. At $n = 80$,ZW2 (0.9631) has the highest power followed by CS (0.9439), SW (0.9115), SF (0.8693), HK1(0.8697), CC1(0.8616) and DA (0.8208). At $n = 100$, ZW2 has the highest power this is followed closely by CS, SW, SF, HK1 and CC1 has the highest power with power approaching 1 followed by DA (0.8839) and MJB(0.8127). At $n = 200$ the power of all tests approaches 1 except GMG with relatively lower power of 0.6610, at $n \geq 500$ all tests power reaches 1 except GMG which has power of 0.8683 even at $n = 1000$. GMG is the least powerful test.

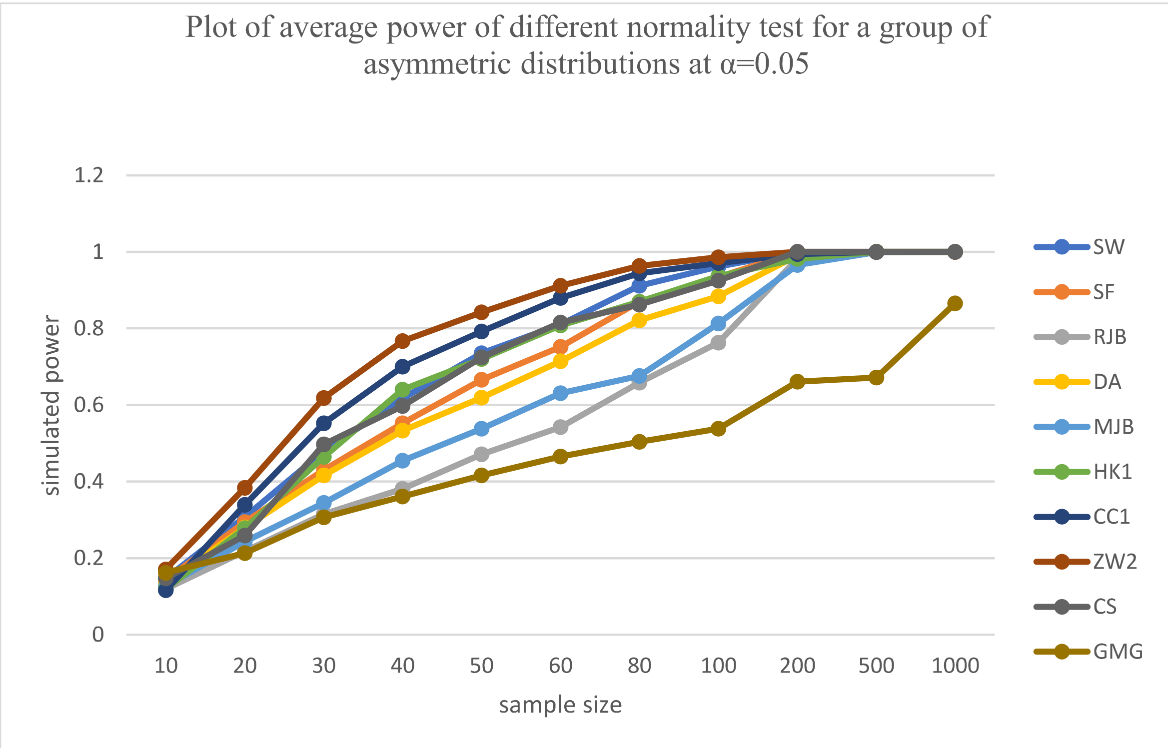

Fig 2: Plot of average power of different normality test for asymmetric distributions at α=0.05

Fig 2 shows clearly that ZW2 is the most powerful test followed closely by CC1, the least powerful is GMG. At $n \geq 500$, all tests achieved maximum power except GMG.

4.2 Type 1 Error Rate of Different Normality Tests for Standard Normal Distribution

Results of the type 1 error of different normality tests on simulated data from standard normal distribution are presented in Table 3 and the plot in Figure 3

Table 3: Type 1 error of different normality tests for standard normal distribution at $\alpha = 0.05$

	Simulated type 1 error									
n	SW	SF	RJB	DA	MJB	HK1	CC1	ZW2	CS	GMG
10	0.049	0.052	0.060	0.051	0.052	0.057	0.054	0.047	0.050	0.044
20	0.054	0.049	0.059	0.050	0.053	0.058	0.056	0.041	0.052	0.047
30	0.052	0.054	0.066	0.048	0.045	0.058	0.049	0.048	0.048	0.052
40	0.047	0.054	0.066	0.047	0.050	0.060	0.053	0.052	0.049	0.053
50	0.047	0.050	0.056	0.049	0.051	0.050	0.055	0.048	0.050	0.051
60	0.051	0.054	0.056	0.054	0.046	0.054	0.053	0.050	0.049	0.052
80	0.054	0.056	0.059	0.054	0.053	0.054	0.057	0.054	0.052	0.054
100	0.048	0.048	0.058	0.051	0.054	0.049	0.049	0.052	0.055	0.053
200	0.046	0.052	0.046	0.046	0.050	0.050	0.059	0.049	0.048	0.056
500	0.052	0.053	0.045	0.057	0.053	0.046	0.054	0.052	0.050	0.054
1000	0.051	0.051	0.042	0.055	0.051	0.055	0.054	0.053	0.050	0.057

Table 3 shows that the test with the highest type 1 error is RJB which has type 1 error > 0.05 for sample sizes $10 - 100$, and < 0.05 for sample sizes < 200, GMG has type 1 error < 0.05 for $n = 10$ & 20 and > 0.05 for $n \geq 30$, HK1 has type 1 error > 0.05 for sample size $n \leq 40$, most of the normality tests do not show any consistent pattern with changes in the sample size.

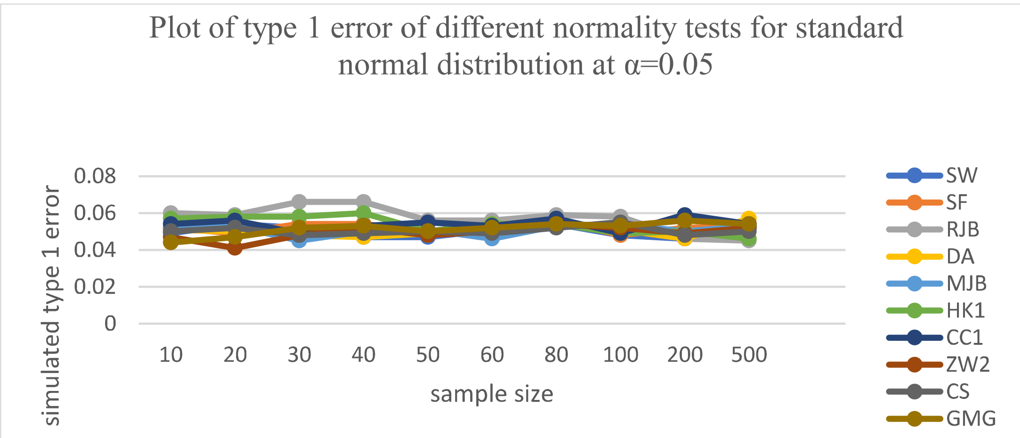

Fig 3: Plot of type 1 error of different normality tests for standard normal distribution at α=0.05
Fig 3 shows that RJB test have the highest type 1 error rate at sample sizes $10 - 100$, which is followed closely by HK1 at sample sizes $10 - 40$, most of the tests have type 1 error close to 0.05 which shows their ability to control type 1 error rate.

Conclusion

In general, it can be concluded that among the 10 tests considered GMG is the most powerful tests for group of symmetric distributions followed by MJB and RJB, while ZW2 is the most powerful tests for group of asymmetric distribution followed by CC1 and SW. The findings of this paper are in good agreement with Uyanto (2022), but MJB test was not included in the paper and the MJB test turned out to be one of the best tests for symmetric distributions. DA test has very low power for group of symmetric distributions, these findings disagree with the work of Ukponnmwa and Ajibade (2017) that classified DA test as uniformly most powerful test, DA test has a good power but only for the group of asymmetric distributions. No test is considered good for smaller sample size $n \leq 30$ for all distributions considered. There is no test which is most powerful for all distributions. In practice, it's crucial to choose the right normality test depending on the sample size and the particular characteristic of the data distribution under examination.

Recommendations

It is therefore recommended that: Researchers should examine the distributions and then choose an appropriate test based on the expected distribution and size of the sample. Researchers should use GMG, MJB and RJB for symmetric distribution and ZW2, CS and SW for asymmetric distribution while considering the sample size. Software engineers should include GMG, MJB, RJB, ZW2 and CC1 test in statistical software where the tests are not available.

References

Abiodun,N.L., Matovu,M.S. & Olanrewaju, R.O.(2022). Statistical Powers of Univariate Normality Tests: Comparative Analysis of 2016 Election Process in Uganda, *European Journal of Statistics,*6(2),2-14.

Arnastauskaite, J., Ruzgas,T. & Braženas,M.(2021).An Exhaustive Power Comparison of Normality Tests.*Mathematics,9*,788. https://doi.org/10.3390/math9070788

Cabana A, Cabana EM (1994). Goodness-of-Fit and Comparison Tests of the Kolmogorov Smirnov Type for Bivariate Populations. *The Annals of Statistics*, 22(3), 1447–1459.

Chen L, Shapiro SS (1995). An Alternative Test for Normality Based on Normalized Spacings. *Journal of Statistical Computation and Simulation*, 53(3-4), 269–287.

Das K.R. & Imon A.H.M.R. (2016). A Brief Review of Test of Normality. *American Journal of Theoretical and Applied Statistics,*5(1),5-12.

Demir, D. (2022) Comparison of Normality Tests in Terms of Sample Sizes under Different Skewness and Kurtosis Coefficients. *International Journal of Assessment Tools in Education* 9(2).,397–409.https://doi.org/10.21449/ijate.1101295

Gastwirth J.L., Gel Y.R., Hui W.L.W., Lyubchich V., Miao W., & Noguchi K (2007). *lawstat: Tools for Biostatistics, Public Policy, and Law*. R package version 3.4, URL https://CRAN.R-project.org/package=lawstat.

Hernandez,H. (2021). Testing for Normality:what is the best method? *Forschem research*,6,5. doi:10.13140/RG.2.2.13926.14406

Hosking, J.R.M.(1990). L-Moments:Analysis and Estimation of Distribution Using Linear Combinations order Statistics. *Journal of Royal Statistical Society.Series B(methodological),51(1),*105-124

Khatun,N.(2021). Application of Normality test in Statistical Anaysis.*open journal of statistics,*11(1), 113-122.doi:10.4236/ojs.2021.111006

Lee,C.,Park,S.&Jeong,J.(2016).Comprehensive comparison of normality tests:Empirical study using many different types of data. *Journal of the Korean Data & InformationScienceSociety,***27**(5),13991412.http://dx.doi.org/10.7465/jkdi.2016.27.5.1399

Rana, S., Eshita N.N., & Al Mamun A.S. (2021). Robust Normality Test in the Presence of Outliers.*Journal of Physic.: Conference Seri*es. Doi:10.1088/1742-6596/ 1863 /012009

Ukponmwan H. N. &, Ajibade F. B. (2017). Evaluation of Techniques for Univariate Normality Test Using Monte Carlo Simulation. *American Journal of Theoretical and Applied Statistics. Special Issue: Statistical Distributions and Modelling in Applied Mathematics,* 6. (5-1) 51-61. doi: 10.11648/j.ajtas.s.2017060501.18

Uyanto S.S. (2022). An Extensive Comparison of 50 Univariate Goodness of fit test of Normality. *Australian journal of statistics*,51, 45-97.

Zhang J. & Wu Y. (2005). Likelihood-ratio tests of normality, 49(3),709-21. Doi:10.1016/j.csda.2004.05.034.