Kütahya Dumlupınar University
Institute of Graduate Studies

Journal of Scientific
Reports-A
E-ISSN: 2687-6167
Number 52, March 2023

RESEARCH ARTICLE

INVESTIGATION OF SOME UNIVARIATE NORMALITY TESTS IN TERMS OF TYPE-I ERRORS AND TEST POWER

Sevda KORKMAZ[1], Yıldırım DEMİR[2*]

[1] Van Yüzüncü Yıl University, Institute of Science, Department of Statistics, Van, tekirdagsevda@gmail.com,
ORCID: 0000-0001-8247-4645
[2] Van Yüzüncü Yıl University, Faculty of Economics and Administrative Sciences, Department of Statistics, Van,
ydemir@yyu.edu.tr, ORCID: 0000-0002-6350-8122

Receive Date:22.12.2022 *Accepted Date: 27.03.2023*

ABSTRACT

In this study, Shapiro-Wilk, Kolmogorov-Smirnov, Skewness, Kurtosis, Lilliefors, Jargue-Bera and D'Agostino -Pearson tests, which are univariate normality tests, were compared in point of type-I error and power performances. For comparisons, samples were created in various distributions and sample volumes by simulation technique, and the probability of type-I error was taken as 0.05 in comparisons. Thus, it is aimed to determine the best test to check whether the normality condition is met in univariate data. As a result of the comparison, it was determined that the Jargue-Bera test gave better results than the other tests in point of type-I error probability. In addition, when the normality tests examined in all distributions were taken into account and compared, it was concluded that the Shapiro-Wilk gives better results than other tests in general for normal and non-normal distributions, and that D'Agostino -Pearson, Skewness and Jargue-Bera tests were also stronger than the other tests. In addition, it was determined that the increase in sample sizes increased power of the test. In conclusion, it can be said that in addition to the distribution pattern, type-I error probability and sample size are also very important factors for test power.

Keywords: *Univariate normality tests, Type-I error, Power of test*

1. INTRODUCTION

In order to apply statistical analysis correctly, it is necessary to know the dataset distribution. In cases where the assumptions of the normal distribution are met, one of the assumptions of the parametric statistical tests is also met. For this reason, it should be determined whether the data of the sample studied show a normal distribution. Because the expected benefits from the tests to be used are related to whether the normality assumption is fulfilled. For this reason, the assumption of normality should be checked before the data are analyzed in order for the results obtained by statistical methods to reflect the real situation [1]. Normality tests are basically divided into two as univariate and multivariate.

To ascertain whether the data exhibit normal distribution, numerous univariate normality tests have been created and extensively studied. To ascertain whether the data has a normal distribution, numerous normality tests are employed. Researchers use these tests to determine whether the data come from a normal distribution [2]. However, it has been an important issue to know which of these tests can be used as the most effective method for the control of the normality prerequisite. For this reason, univariate normality tests should be compared in terms of type-I error and test power under different experimental conditions, and the performances of these tests should be measured under the same conditions. Which of these tests used has the higher power and whichever gives the same level of type-I error at the end of the experiment, the test in question should be preferred in testing whether the assumption of normality is met [3].

In this study, Shapiro-Wilk (SW), Kolmogorov-Smirnov (KS), Skewness, Kurtosis, D'Agostino-Pearson (DP), Lilliefors and Jargue-Bera (JB) tests, which are univariate normality tests, are compared in terms of power performance and type-I error. In order to make the said comparisons, normal, symmetrical and different curvature sample distributions were created in different sample sizes with the simulation technique, and the tests were compared using these data sets. In the comparisons made, the probability of type-I error was taken as 0.05 and the tests were compared according to 10000 simulation results using the Python programming language. Comparisons were made by taking sample sizes of 10, 20, 25, 30, 40, 50, 70, 100, 150. Thus, the effects of different sample sizes on the tests are also observed. As a result, before the statistical calculations of the univariate data, the normality test was selected by taking into account the sample size of the available data to check whether the normality prerequisite was met.

The aim of this study is to compare some univariate normality tests in terms of type I error and test power in data sets with different sample sizes and distributions.

2. MATERIAL AND METHOD

It is known that the Monte Carlo technique provides excellent dimensional control and has good power in generating data [4]. For this reason, this technique is used in the simulation study for the comparison of the tests. Since the simulation results for the sample and alternative distributions vary at different levels, the data are derived according to different sample sizes and distributions [2]. Thus, the effects of sample size and data distribution on the tests are observed.

Kurtosis, Skewness, SW, KS, DP, Lilliefors and JB tests, which are univariate normality tests, were compared in terms of type-I error and power performances. In the comparisons performed, the type-I error value was taken as 0.05. Data were derived from a normally distributed population with 50 mean and 5 standard deviations, taking into account different sample sizes. Each normality test statistic was calculated for all 10000 samples created by simulation, and the rejection numbers of the basic hypothesis were counted for each test and then these numbers were converted to %. In addition, different skewness (α_3) and kurtosis (α_4) values and type-I error rates were examined.

In order to compare the power of the specified univariate normality tests, non-normal Chi-Square, Uniform, t and Beta distributions were also used in addition to the normal distribution.

2.1. Type-I Error, Type-II Error and Power of Test

Basic idea of statistical hypothesis testing is to decide whether a sample of data comes from a population, assuming true a formulated hypothesis about the population.

When a sample is not typical of the population, generalizing from it to the whole population might result in inaccurate conclusions and interpretations. Rejecting an actually true hypothesis leads to a type-I error, and not rejecting an actually false hypothesis leads to a type-II error. Type-I (α) and type-II (β) errors can never be completely prevented, however by boosting sample numbers, error levels can be decreased. Because the larger the sample size, the better the population is represented and the less the difference between the population and the sample [5]. The α and β values are not independent of each other, and there is an inverse proportion between them. Since the increase in the sample size reduces the sampling error, both errors are minimized [6].

If the H_0 hypothesis is not true, H_0 rejecting is a correct decision. The rejection of the false H_0 hypothesis is defined as the power of the test (1–β), and the power of a test increases as the sample size increases [7].

2.2. Distribution Models

A probability distribution describes how the values of a random variable are distributed. Random variables in distributions are divided into two as continuous and discrete distributions in terms of the values they take. If the random variables have a probability of being in a certain interval and being continuous, the distribution is called continuous distribution. There are many continuous statistical distributions. Normal, chi-square, beta, t and uniform distribution are among these distributions [8].

2.2.1. Normal distribution

Of all distributions, the normal distribution is the most well-known and often utilized. The normal distribution, a hypothetical symmetrical distribution used for comparison, is actually a family of distributions and has the following properties.

In the standard normal distribution, the mode, mean and median are equal and curve symmetrical.

The total area under the curve is equal to 1.

The curve is bell-shaped.

The data are largely close to the mean. The data decreases with distance from the mean in both directions.

Almost all data (99.7%) fall within a range of ±3 standard deviations from the mean [9].

The normal distribution depends on the mean and standard deviation, and the normal distribution is calculated by Eq. 1. For a sample of size *n*, with standard deviation σ and the population mean μ.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1)$$

The standard deviation serves as a measure for dispersion; the higher this value, the more spread out the curve. The mean, otherwise, is a measure of central tendency and determines the position of the normal distribution curve that it divides into two [9].

2.2.2. Chi-Square distribution

A chi-square random variable is defined as the sum of the squares of independently distributed standard normal random variables, describing the additive property of independent chi-square random variables [10].

Where x is a random variable and n is the degrees-of-freedom of the distribution, the Chi-Square Distribution's Probability Density Function is defined by Eq. 2 and the Cumulative Probability Density Function by Eq. 3 [11].

$$f(x) = \begin{cases} \frac{1}{\Gamma(\frac{n}{2})2^{n/2}} x^{\frac{n}{2}-1} e^{-x/2} & 0 < x < \infty \\ 0 & x \leq 0 \end{cases} \quad (2)$$

$$F(x) = \int_0^x \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = 1 - \sum_{k=0}^{n-1} \frac{e^{-2x}2x^k}{k!}, x > 0 \quad (3)$$

The sum of the squares of the values of the n-unit sample, independently selected from any main population with a normal distribution with a variance of 1 and an arithmetic mean of 0, expresses the chi-square distribution. Chi-square indicates a continuous value greater than or equal to zero.

For small degrees of freedom, the distribution is skewed to the right, and for large degrees of freedom, the distribution approaches a normal curve.

The chi-square tests whether the observed frequencies according to H_0 approach the calculated expected frequencies, and in general, the test statistic is calculated with Eq. 4.

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - B_i)^2}{B_i} \quad (4)$$

Here O_i, is the observed frequencies and B_i, is the expected frequencies [12].

2.2.3. Beta distribution

The beta density function is a two parameter continuous density function defined in the range $0 \leq x \leq 1$. If x, which is a random variable, has a beta probability density distribution with the parameters α and β, the density function of x is given by Eq. 5.

$$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, & 0 \leq x \leq 1 \\ 0, & Other\ cases \end{cases} \quad (5)$$

Here, the value of $B(\alpha, \beta)$ represents the gamma function (Γ) and is calculated with Eq. 6 [13].

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad (6)$$

2.2.4. t distribution

One of the key characteristics of the t distribution, which has many uses in mathematical statistics, is that the distribution approaches the normal distribution as sample sizes approach infinity [14]. For a given sample size or degree-of-freedom, the set of all t values measured from each possible random sample gives the t-distribution. The parameters μ and σ are meaningful descriptive measures that locate the center and describe the spread associated with a random variable x.

The test statistic for the *n*-size t distribution with sample mean $\bar{x}$, standard deviation s, population mean μ is given with Eq. 7 [15].

$$t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} \quad (7)$$

2.2.5. Uniform distribution

The simplest form of continuous probability distribution is uniform probability distribution. If the random variable x takes every value with the same probability in a given interval, the distribution is uniform. The uniform distribution, which takes values in a certain range, is defined by the parameters a and b. It can be said that a random variable x has a uniform probability distribution in the range [a, b], with a minimum and b maximum value [16]. Thus, the probability density function of x is represented by Eq. 8 and the cumulative distribution function by Eq. 9.

$$f(x) = \begin{cases} \frac{1}{(b-a)} & a \leq x \leq b \\ 0 & Other\ cases \end{cases} \quad (8)$$

$$(x) = \int_{-\infty}^{x} \frac{1}{b-a} dx = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1, & x \geq b \end{cases} \quad (9)$$

The expected values for the mean and variance of a random uniform variable are calculated with Eq. 10 and Eq. 11.

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a}\left(\frac{x^2}{2}\Big|_a^b\right) => \quad \mu = \frac{b+a}{2} \quad (10)$$

$$Var(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{3}(b^2 + ab + a^2) - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12} \quad (11)$$

In addition to being used to create random numbers from other distributions, the uniform distribution can also be used as a "first guess" if just the random variable x's range between a and b is known. Additionally, a uniform distribution can be used to describe the probability behavior of such a phenomenon in real-world problems with smooth behavior within a particular range [15].

2.3. Univariate Normality Tests

The majority of frequently employed statistical techniques are predicated on the idea that the data are normally distributed. In order to obtain more reliable and accurate results from the studies, it is important to determine whether the normalcy assumption is provided [17]. Normality assumption can be checked in two ways, graphical and analytical methods. While graphical methods for normality test show the distribution of data with various graphs, analytical methods include many different statistical test solution methods using mathematical models [18].

As a result of all statistical tests, a probability (p) value of the test statistics is calculated. The p value in statistical analysis determines whether a test becomes significant. A p-value over the significance level in one study implies that the data set is normally distributed, and the higher the value, the more sensitive the statistic utilized in comparison to the other studies [18]. In general, hypotheses regarding tests are established as follows.

H_0: The sample is drawn from a population with a normal distribution.
H_1: The sample is not drawn from a normally distributed population.

2.3.1. Shapiro-Wilk test

The test was proposed by Shapiro and Wilk (1965) for testing the assumption of normal distribution. By dividing the square of an appropriate linear component of the sample rank statistics by the sum of squares, the SW test statistic is calculated [19]. SW test statistic is calculated with Eq. 12.

$$SW = \frac{(a'x)^2}{(n-1)S^2} = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (12)$$

Here, it is denoted by $a' = (a_1, a_2, \dots a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}}$. While $V = V_{ij}$ represents the nxn dimensional covariance matrix, $m' = (m_1, m_2, \dots, m_n)$ represents the vector of expected values of n rank statistics in the standard normal distribution $N(0,1)$. In addition, $x_i = (x_1, x_2, \dots, x_n)$ gives the arithmetic mean of the observations, $\bar{x}$ gives the values of the observations ordered from the smallest to the largest of an n dimensional random sample with a normal distribution with a mean of 0 and a variance of 1 [19].

2.3.2. Skewness test

Skewness is used to determine the goodness of fit of the data by referring to a particular type of distribution. The skewness test is a measure that reflects the degree to which a distribution is asymmetrical. Distributions with long tails to the right are called positive curves and distributions with long tails to the left are called negative curves. The third moment of a distribution about the mean gives the measure of the skewness [20]. Accordingly, the skewness measure is calculated with Eq. 13.

$$m_3 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^3}{n} \tag{13}$$

Since m_3 is a third order value, the unitless statistics g_1 is generally used to indicate the curvature. g_1 is an estimator of the population parameter γ_1 and is calculated with Eq. 14. S is the sample standard deviation.

$$g_1 = \frac{m_3}{S^3} \tag{14}$$

When the skewness value is between -3 and +3, the distribution is symmetrical, and when this value is 0, the symmetrical distribution approaches a perfect distribution. If the value of g_1 is greater than 0, the distribution is skewed to the right, and less than 0 causes the distribution to be skewed to the left. Although the normal distribution is symmetrical, not all symmetrical distributions show normal distribution [20].

2.3.3. Kurtosis test

Kurtosis is often defined as a measure that reflects the degree to which a distribution reaches its peak. In other words, kurtosis gives information about the height of a distribution according to the standard deviation value, and the most important reason for measuring kurtosis is to determine whether the data are derived from a population with a normal distribution. Kurtosis is generally considered within the framework of three general categories: mesokurtic, leptokurtic, and platikurtic, all of which have representative frequency distributions. For normal distributions, the Kurtosis is called mesokurtic, if the Kurtosis is very high, it is called leptokurtic, and if it tends to be much broader, it is called platikurtic [20]. The kurtosis is calculated using certain quantitative values in the distribution, and the most effective kurtosis measure is obtained with the fourth moment relative to the mean. Thus, the kurtosis value, which is the fourth moment with respect to the mean, is calculated by Eq. 15.

$$m_4 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^4}{n} \tag{15}$$

Since the calculated value for m_4 is of the fourth order, the unitless statistics g_2, which is an estimator of the population parameter γ_2, is used to indicate the kurtosis and is calculated with Eq.16.

$$g_2 = \frac{m_4}{S^4} \tag{16}$$

When a distribution has a mesokurtic (normal) distribution, the value of g_2 equals 0. If the value of g_2 is significantly greater than 0, it creates a leptokurtic distribution and a small value creates a platikurtic distribution [20].

2.3.4. D'Agostino-Pearson test

D'Agostino and Pearson [21], D'Agostino [22] and Zar [23] demonstrated that the D'Agostino-Pearson test, proposed by D'Agostino and Pearson in 1973, is a very effective method for assessing fit to normal distribution. The D'Agostino-Pearson test statistic is based on the sum of the squares of the kurtosis and Skewness test statistics values and is calculated with Eq. 17 [20]. Here, $Z(\sqrt{b_1})$ shows the skewness test statistical value, while $Z(b_2)$ shows the kurtosis test statistical value. The hypotheses regarding the D'Agostino-Pearson test are established as follows.

H_0: Sample distribution conforms to normal distribution
H_1: Sample distribution is not suitable for normal distribution

$$\chi^2 = Z\left(\sqrt{b_1}\right)^2 + Z(b_2)^2 \tag{17}$$

The D'Agostino-Pearson test is based on the χ^2 statistic, as it has a two-degrees-of-freedom chi-square $(\chi^2_{(2)})$ distribution. Thus, when $\chi^2 > \chi^2_{\alpha,2}$, the H_0 hypothesis is rejected [20].

2.3.5. Jarque-Bera test

Jargue-Bera test, which was proposed by Jargue and Bera in 1980 for testing normality in univariate data, is based on skewness and kurtosis tests. The Jargue-Bera test statistic is calculated using the skewness and kurtosis coefficients obtained from a $(x_1, x_2, \dots, x_n)$ dataset assumed to be normally distribution [24]. If $\{x_1, x_2, \dots, x_n\}$ is assumed to be a randomly drawn sample from an independent population, the jth trigonometric moment with respect to the mean is calculated by Eq.18.

$$m_j = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^j \qquad j = 2, 3 \tag{18}$$

Here; Since $\bar{x}$ is the sample mean, $\widehat{\sigma}^2$ can be written as m_2. Thus, the skewness and kurtosis statistics are calculated with the JB test statistic Eq. 19, to be $\sqrt{b_1} = \frac{m_3}{\widehat{\sigma}^3}$ and $b_2 = \frac{m_4}{\widehat{\sigma}^4}$ respectively.

$$JB = n\left[\frac{\left(\sqrt{b_1}\right)^2}{6} + \frac{(b_2-3)^2}{24}\right] \tag{19}$$

The JB test statistic has an asymptotic chi-square distribution with two degrees-of-freedom $(\chi^2_{(2)})$ under the assumption of normality [24].

2.3.6. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test was first proposed by Kolmogorov in 1933 as an alternative to the chi-square fit test and was revised by Smirnov in 1939. This test is based on the largest absolute difference between the expected and cumulative distribution functions [25]. In order to calculate the

JSR A
Journal of Scientific Reports

Korkmaz, S. and Demir, Y., Journal of Scientific Reports-A, Number 52, 376-395, March 2023.

values of the cumulative distribution function, the observation values of the sample are ordered from the smallest to the largest, as $x_1 < x_2 < \cdots < x_n$. Then, Z values and their cumulative probabilities are calculated for these observation values. Thus, Kolmogorov-Smirnov (KS) test statistic for n ordered data is calculated with Eq. 20.

$$D = max|F_o(X) - F_e(X)| \tag{20}$$

Here, $F_o(X)$ shows the observed cumulative frequency value of the variable X, while $F_e(X)$ shows the expected cumulative frequency value of the variable X.

2.3.7. Lilliefors test

Lilliefors test, which is a modification of the Kolomogorov-Smirnov goodness-of-fit test, is one of the best tests developed by Lilliefors (1967) and Van Soest (1967) and used to test the assumption of normality [26; 27]. In cases where the sample mean and variance are unknown and must be estimated from the population, Lilliefors table values obtained by Monte Carlo calculations are used to test whether this sample is drawn from a normal population [28].

This test of normality is stronger than other tests under many non-normal circumstances, according to Lilliefors (1967). In addition, according to Dagnelie (1968), an analytical formula can be used to estimate the critical values provided by Lilliefors [29]. This formula eliminates the risk of error when entering table values and facilitates the software of computer programs [26]. The observed distribution function of the n dimensional $x_1, x_2, \dots, x_n$ random sample is $S_n(x)$ for each value of x and is defined by Eq. 21.

$$S(x) = \begin{cases} 0 & x < x_1 \\ \frac{i}{n} & x_{(i+1)} \qquad i = 1, 2, \dots, (n-1) \\ 0 & x \geq x_n \end{cases} \tag{21}$$

The vertical distance between the distribution function $S(x)$ and the theoretical distribution function $F(x)$ is denoted by D. The size of the differences in D results, which represents sampling errors, is used as a measure of conformity to the normal distribution [30].

$$D = mak.|S(x_i) - F(x_i)| \tag{22}$$

3. RESULTS

3.1. Comparison of Tests in terms of Type-I Error

In this study, 7 of the univariate normal distribution tests were examined and the performances of these tests were determined with data sets owner different sample sizes obtained by simulation. Analysis results for Type-I error values are given in Table 1.

In Table 1, when error values are taken as 0.05, it is seen that type-I error probabilities of test statistics are generally close to 5%. For all sample sizes, JB is the test with the lowest probability of type-I

JSR A
Journal of Scientific Reports

Korkmaz, S. and Demir, Y., Journal of Scientific Reports-A, Number 52, 376-395, March 2023.

error, that is, the probability of rejecting the H_0 hypothesis, regarding that these samples are drawn from a normally distributed population.

Table 1. Type-I error values for different sample sizes in normal distribution.

α	n	DP	Kurtosis	Skewness	SW	KS	Lilliefors	JB
	10	5.77	4.38	5.04	5.03	5.14	5.06	0.91
	20	5.58	4.48	4.99	5.23	4.99	5.09	2.43
	25	5.69	4.72	5.06	4.87	5.28	5.12	2.78
	30	5.76	5.13	5.08	5.07	4.56	5.08	3.10
0.05	40	5.99	5.17	5.05	5.24	5.70	5.45	3.61
	50	5.80	5.37	5.00	5.10	5.08	4.99	3.71
	70	5.02	5.08	4.92	4.89	5.25	3.74	3.60
	100	5.44	5.29	5.17	4.93	4.83	4.96	4.23
	150	5.41	5.58	4.81	4.95	4.81	4.48	4.44

In all sample sizes, it is determined that the type-I error probabilities of the JB test are less than 5% and has lower values than the other six tests. As the sample size increases, the probability of type-I error in the JB test also increases. When the sample size is $n \leq 25$, the JB test is followed by the kurtosis test, the KS test when it is 30 and 100, the skewness test when it is 40, and the Lilliefors test when it is 50, 70 and 150. In general, it can be said that type-I error values in the skewness test do not change much with the sample size and are generally around 5%.

Type-I error probabilities of DP test range from 5% to 6%, and highest type-I error probability is found to be 5.99% when $n = 40$. In addition, among these 7 tests, the highest probability of type-I error is obtained from the DP test for all sample sizes except $n = 70$ and 150. Type-I error probabilities in the KS, SW, and Lilliefors tests are generally around 5%, and the highest probability of type-I error in these three tests is obtained when the sample size is 40, and is found to be 5.70%, 5.24%, and 5.45%, respectively. In addition, the lowest probability value among these three tests, when $n = 70$, belongs to the Lilliefors test with a value of 3.74%, and this value is very close to the type-I error probability of the JB test for the same sample size.

The tests are compared in terms of type-I error rates the different skewness, kurtosis and in the sample sizes and the results are given in Table 2.

Table 2. Type-I error rate values for different skewness, kurtosis and sample sizes.

α	n	DP	Kurtosis	Skewness	SW	KS	Lilliefors	JB
	10	5.72	3.66	5.24	4.94	4.72	4.67	1.01
	20	5.44	4.32	4.88	4.73	4.87	4.56	2.26
$\alpha_3 = 0$ $\alpha_4 = 1$	25	5.49	4.48	4.76	4.76	4.68	4.62	2.81
	30	6.05	5.25	5.39	5.32	5.32	5.57	3.20
	40	5.71	4.94	4.93	5.00	4.72	5.04	3.32

JSR A
Journal of Scientific Reports

Korkmaz, S. and Demir, Y., Journal of Scientific Reports-A, Number 52, 376-395, March 2023.

	50	5.96	5.34	5.16	5.25	4.81	5.07	4.18
	70	5.58	5.35	4.97	4.94	5.20	3.22	3.75
	100	5.36	5.20	4.67	4.87	5.07	4.42	4.01
	150	5.53	5.36	5.16	5.02	5.34	2.96	4.65
$\alpha_3 = 0.05$ $\alpha_4 = 2.8$	10	5.68	3.82	5.07	4.81	4.82	4.62	0.86
	20	5.96	5.08	5.24	5.12	4.81	5.15	2.48
	25	6.32	5.45	5.48	5.57	5.21	4.74	3.23
	30	5.69	5.10	5.04	5.12	4.99	5.12	3.13
	40	5.54	5.29	4.77	4.99	4.93	4.95	3.17
	50	5.83	5.35	5.10	4.98	5.10	5.10	3.82
	70	6.37	5.58	5.49	5.08	4.80	3.52	4.34
	100	5.46	5.30	5.18	4.83	5.11	4.95	4.02
	150	5.45	4.92	5.19	5.07	5.16	3.13	4.42
$\alpha_3 = 0.05$ $\alpha_4 = 3.2$	10	5.90	3.91	5.09	4.85	5.11	5.04	0.83
	20	5.70	4.78	5.29	5.17	5.44	5.27	2.49
	25	5.40	4.74	5.02	4.93	5.10	5.08	2.65
	30	5.80	4.91	4.97	5.23	4.98	5.38	3.19
	40	5.71	5.18	5.04	5.24	5.01	5.24	3.44
	50	5.67	5.28	4.69	5.15	4.90	4.98	3.57
	70	5.57	5.60	4.75	5.31	4.95	3.30	4.00
	100	5.38	5.11	4.86	5.05	4.73	4.94	4.13
	150	5.48	5.19	5.11	5.24	5.69	3.46	4.33
$\alpha_3 = -0.05$ $\alpha_4 = 2.8$	10	6.04	4.21	5.28	4.96	4.96	4.84	0.83
	20	5.67	4.64	5.25	4.99	4.83	5.01	2.53
	25	5.59	4.67	4.57	4.78	5.31	4.93	2.43
	30	6.06	5.35	5.02	4.96	4.75	4.96	3.24
	40	6.05	5.43	5.15	5.25	5.30	4.96	3.71
	50	5.61	5.42	5.01	5.05	5.46	4.58	3.62
	70	5.92	5.75	4.85	5.11	5.23	3.24	4.04
	100	5.58	5.10	4.99	4.71	4.73	4.97	4.10
	150	5.26	5.23	4.87	5.14	4.76	3.25	4.44
$\alpha_3 = -0.05$ $\alpha_4 = 3.2$	10	6.42	4.32	5.75	5.05	4.73	5.45	0.75
	20	5.37	4.27	4.96	5.08	5.20	5.29	2.32
	25	5.40	4.40	4.98	4.8	4.88	5.26	2.63
	30	5.82	5.26	5.13	5.26	5.20	5.05	3.12
	40	5.50	5.35	5.06	5.04	5.07	5.14	3.47
	50	5.42	5.19	4.95	4.64	4.74	4.75	3.84
	70	5.93	5.78	5.22	5.27	4.90	3.30	4.15
	100	5.43	5.40	4.76	5.00	5.00	5.19	4.21
	150	5.31	5.35	4.76	4.75	4.92	3.46	4.08

When Table 2 is examined, it is observed that although the skewness and kurtosis values are changed, the normal distribution tests show similar behaviors to those in Table 1 in terms of type-I error rates and do not show significant differences from Table 1.

JSR A
Journal of Scientific Reports

3.2. Comparison of Methods in terms of Power of Test

In order to compare the performance of the 7 tests in terms of power, data sets suitable for chi-square, beta, t and uniform distribution are generated and comparisons are made on these data sets.

Using chi-square distributions with various degrees of freedom, the comparison of univariate normality tests is looked at in terms of test power. Because the chi-square distribution approaches the normal as the degree of freedom grows [31]. Thus, the power values obtain from data sets with univariate 1 and 5 degrees of freedom (df) chi-square distributions in different sample sizes are given in Table 3.

When Table 3 is examined, it is seen that the test powers approach 100% with the increase in sample size in all samples drawn from a population with 1 degree of freedom chi-square distribution. It is seen that the SW test is stronger than the other tests in all sample sizes, the JB test is weaker when $n = 10$ and the kurtosis test is weaker when $n > 10$ than the other tests. The highest power value (99.99%) is reached when $n = 40$ with the SW test. When the sample size is $n \geq 25$, it was determined that the power performances of all the tests except the kurtosis were higher than 84%.

Table 3. Power values of the tests in the chi-square distribution (%).

α: 0.05	**n**	**DP**	**Kurtosis**	**Skewness**	**SW**	**KS**	**Lilliefors**	**JB**
	10	49.63	34.43	56.32	73.36	54.47	53.6	27.34
	20	80.78	55.26	89.49	98.29	86.29	88.27	71.89
	25	89.97	63.66	95.38	99.74	94.65	95.51	84.36
	30	94.14	70.23	98.07	99.93	99.88	98.12	92.15
df=1	40	98.80	80.14	99.73	99.99	99.93	99.85	98.27
	50	99.88	86.89	99.97	99.99	99.99	99.99	99.79
	70	99.99	94.54	99.99	99.99	99.99	99.99	99.99
	100	99.99	98.56	99.99	99.99	99.99	99.99	99.99
	150	99.99	99.87	99.99	99.99	99.99	99.99	99.99
	10	18.93	12.69	19.92	20.19	14.34	14.68	6.32
	20	35.59	21.15	41.32	45.15	24.19	28.11	24.99
	25	42.17	23.64	50.08	54.46	35.67	32.47	35.52
	30	48.46	26.19	58.78	64.30	37.91	32.21	39.99
df=5	40	61.33	31.65	73.61	79.35	48.88	48.43	55.00
	50	72.77	36.67	83.75	89.26	59.82	59.73	68.43
	70	88.03	45.42	94.68	97.63	65.52	69.72	86.37
	100	97.73	56.13	98.93	98.93	89.87	88.83	96.97
	150	99.96	69.02	99.96	99.99	97.79	96.57	99.95

According to the results of the 5-degrees-of-freedom Chi-square distribution, it is observed that the test powers increased with the increase in sample size. In addition, it is determined that the SW test is more powerful than the other tests in all sample sizes. In general, the most powerful test after the SW test is the skewness test, followed by the DP test. The JB test is the weakest test when the sample size is 10, and the kurtosis test is the weakest for other sample sizes. In the same sample size, it can be said

JSR A
Journal of Scientific Reports

that the SW test and the skewness test give closer values to each other. In addition, a parallelism is observed between the power values of the KS test and the Lilliefors test.

In Table 4, the power values of the normality tests used in data sets with different sample sizes (2, 5) and (1, 1) parameter Beta distributions are given.

Beta distribution with 1 and 1 parameters has a flat and symmetrical distribution with Kurtosis and skewness values, while beta distribution with 2 and 5 parameters has an asymmetric and vertical distribution [31]. In this study, the power performances of univariate normality tests are also examined in beta distributions at the specified parameter values. According to the results of beta distribution with 2 and 5 parameters given in Table 4, the SW test is more powerful than other tests in all sample sizes. When $n = 10$, KS is the strongest test after the SW test, while the SW test and SW are followed by the skewness test in other sample sizes. It is observed that the weakest test is the JB test when $n < 40$, and the kurtosis test for larger sample sizes. Also, after $n \geq 50$, the sample size hardly affect the power of the kurtosis test.

According to the results of 1 and 1 parameter beta distribution, it is seen that $n = 10$ the first two strongest tests are SW and kurtosis, respectively, and the strongest test is kurtosis when there is $n > 10$. The second strongest test following the kurtosis test is SW when $n < 30$ and DP when $n \geq 30$. At small sample sizes ($n \leq 50$) JB is the weakest test, while at larger sample sizes, especially when is $n \geq 100$, a huge improvement in the power of JB is seen and so, skewness becomes the weakest test. Also, when $n = 150$ the power performances of the kurtosis, DP, JB and SW tests approach 100%.

Table 4. Power values of normality tests in beta distributions (%).

α: 0.05	**n**	**DP**	**Kurtosis**	**Skewness**	**SW**	**KS**	**Lilliefors**	**JB**
Beta (2.5)	10	7.99	5.47	7.91	8.98	8.16	7.64	1.54
	20	11.51	7.84	13.25	16.93	12.21	11.92	5.17
	25	13.51	8.86	16.31	21.46	13.45	13.66	6.27
	30	15.71	9.20	20.59	27.30	14.67	16.01	8.31
	40	19.29	9.84	26.48	38.08	19.18	20.60	11.32
	50	24.64	11.19	34.90	50.24	23.25	26.01	16.05
	70	37.24	11.63	50.75	71.23	43.58	30.14	27.15
	100	61.14	11.71	70.15	89.22	50.73	49.45	50.83
	150	89.99	11.15	89.82	98.93	67.09	62.37	85.40
Beta (1.1)	10	2.78	6.92	2.12	7.92	2.03	5.89	0.25
	20	14.78	29.36	0.72	19.83	5.02	9.64	0.10
	25	27.23	44.35	0.50	28.61	29.60	11.79	0.03
	30	39.36	56.86	0.32	37.99	33.31	14.41	0.01
	40	63.18	76.88	0.28	57.58	46.95	19.09	0.01
	50	80.76	88.91	0.25	75.52	57.70	26.39	0.01
	70	95.87	97.88	0.18	93.66	70.88	30.86	3.64
	100	99.78	99.91	0.23	99.65	85.14	58.65	56.60

	150	99.99	99.99	0.15	99.99	93.94	75.08	98.64

As the degrees of freedom increase in the t distribution, the type-I errors get very large values and the distribution approaches the standard normal distribution [32]. In this study, power performances of univariate normality tests are also investigated with two different degrees of freedom t distributions. Power performances of normality tests for various sample sizes in t distributions with 10 and 5 degrees of freedom are given in Table 5.

According to the results of the 10-degrees-of-freedom t distribution given in Table 5, normality tests for all sample sizes perform quite poorly because the $t_{\alpha;10}$ distribution has a symmetrical and vertical distribution. When $n < 150$, the DP test gives the best results in terms of power performance. When the sample size is $n < 40$, the skewness test gives the best results after DP. The JB test, which performed poorly at first and even showed the worst performance when $n = 10$, generally gives the best results after DP when $n \geq 40$, and becomes the best test when $n \geq 150$. In addition, when $n > 10$, the worst performance belongs to Lillefors and KS tests, respectively, and the power performances of these two tests were found to be very close to each other. Thus, it can be said that the kurtosis, DP and JB tests, which are the moment tests that can test the steepness, generally have better performances than the other tests.

Considering the inclination and kurtosis values of the 5-degrees-of-freedom t distribution, its $t_{\alpha;5}$ distribution is symmetrical and vertical. When the 5-degrees-of-freedom t distribution is examined, it is seen that all test statistics have weak power performances in all sample sizes. However, as the sample sizes increased, significant improvements occurred in the JB, kurtosis, DP and SW tests, albeit insufficiently, compared to the other tests, but even at high sample sizes ($n < 150$), the estimated test statistical power values remained below 70%. When $n \geq 150$, the most powerful tests are JB, kurtosis, DP and SW, respectively, the powers of these tests exceeded 71%. When $n < 70$, DP is the strongest test, and when $n \geq 70$, the most powerful test is the JB test. Although the weakest test varies according to the sample size, it is seen that the JB test is the weakest test when $n = 10$, and the KS test is generally the weakest in other sample sizes.

Table 5. Power values of the tests in t distribution (%).

α: 0.05	n	DP	Kurtosis	Skewness	SW	KS	Lilliefors	JB
df=10	10	9.72	6.37	8.74	7.26	6.12	6.61	1.94
	20	12.58	9.65	11.20	9.70	7.32	6.86	7.66
	25	13.86	10.71	12.42	10.68	7.63	7.90	9.60
	30	15.57	12.07	13.64	11.72	7.98	7.68	11.83
	40	16.99	13.69	14.29	13.54	8.73	8.15	14.66
	50	19.41	16.14	15.74	15.63	8.81	8.60	17.73
	70	23.29	20.06	17.79	18.98	9.57	6.91	22.89
	100	26.91	24.73	18.80	22.63	10.81	10.20	20.21
	150	34.08	33.70	21.30	22.91	11.63	9.52	37.59
df=5	10	14.67	10.23	13.47	11.03	9.23	9.43	4.43
	20	23.18	18.88	20.79	19.08	10.46	13.46	16.63

25	26.73	22.01	23.69	22.21	13.11	14.44	21.27
30	29.68	25.01	25.15	24.97	13.86	16.23	25.35
40	35.34	31.01	28.99	30.55	16.98	18.69	32.90
50	39.24	36.46	30.80	34.80	20.66	21.25	38.70
70	49.02	46.76	36.38	45.37	23.57	21.75	50.30
100	60.35	60.12	40.52	56.33	32.77	33.37	63.27
150	73.74	75.37	44.58	71.23	43.26	38.15	77.69

The distribution of the data drawn from the uniform distribution has a symmetrical feature. The power performances of the tests at different sample sizes in uniform distribution are given in Table 6.

When Table 6 is examined, it is seen that the strongest test is skewness except for the case of $n = 10$, the second strongest test is SW in small sample sizes $(n < 30)$, DP in larger sample sizes $(n \geq 30)$. When $n = 150$, it is seen that the powers of the kurtosis, DP and SW tests approach 100%. In addition, the power of the JB test, which has the weakest power when $n < 70$, approached 100% in large sample sizes $(n = 150)$. Again, although the skewness test performed very poorly and there is a slight improvement, not with the increase in the sample volumes, there is a decline in the test performance.

Table 6. Powers of normality tests in uniform distribution (%).

α	n	DP	Kurtosis	Skewness	SW	KS	Lilliefors	JB
0.05	10	2.83	7.47	1.97	8.67	5.65	6.36	0.21
	20	15.51	29.98	0.65	20.24	6.98	9.78	0.06
	25	27.22	43.90	0.36	28.78	9.67	12.2	0.02
	30	40.28	57.18	0.41	38.53	12.81	14.79	0.05
	40	62.82	76.78	0.39	57.80	19.09	20.96	0.02
	50	79.74	88.46	0.23	74.95	21.23	26.29	0.03
	70	95.81	98.13	0.19	94.00	39.08	30.95	3.60
	100	99.70	99.86	0.08	94.31	55.24	59.06	56.5
	150	99.99	99.99	0.09	99.99	80.57	75.78	98.65

4. DISCUSSION AND CONCLUSION

In this study, it is determined that the best results in terms of type-I error in different sample sizes are given by the JB test. The same results are obtained for normal distributions in the studies conducted by [33] and [3]. It is determined from the simulation results that the most powerful test for normal distributions is the skewness test by [2]. [34] is determined that the JB test is not a strong test in the chi-square distribution, similar results are obtained for n=10 in this study, and the kurtosis test is the weakest test for larger sample sizes. In this study, the lowest test power is obtained from the t distributions, especially the 10 degrees of freedom t distribution. [35] is stated that all tests have low power in the beta (2, 5) distribution and [33] in the t distribution. In this study, it is observed that SW is a strong test in general in non-normal distributions and the test power increased as the sample size increased. Similarly, [1] and [2] suggest SW for non-normally distributed distributions; [25], on the other hand, states that although SW has low power in small sample sizes in normal and non-normal distributions, it is generally a powerful test.

[36] are compared the results of nine statistical procedures for normality assessment. They are stated that the SW test is more sensitive than many alternative tests used to test normality in small samples ($n < 20$). [37] compars some tests with the SW test at different distributions and sample sizes and tested their power performance for normality. it is stated that the Anderson-Darling (AD) test has a power close to the SW test and can be used as an alternative. [38] compared univariate normality tests using 20, 50, and 100 sample sizes in different distributions in terms of type-I error and test power, and found that the AD test is more powerful than the KS test. [39] compared the KS, SW, AD, and Lilliefors tests for test power. It is observed that the SW test is the strongest test of normality in normal and non-normal distributions, followed by the AD, Lilliefors and KS tests, respectively. In addition, it has been determined that these four tests have low power performance in small-sized samples.

With the Monte Carlo simulation study, univariate normality tests are examined in terms of type-I error value in samples taken from a normally distributed population. When the test results according to $\alpha = 0.05$ are compared, the JB test type-I error value remains well below 5%, and the type-I error value increases with the sample size, reaching a maximum of 4.44%. Type-I error probabilities of skewness, kurtosis, KS, SW, and Lilliefors tests ranged from 4.5-5.5% with small deviations, and generally close to 5%. However, although the DP test has high power values in general, type-I error rates could not maintain the 5% level and all values of type-I error rates are found to be between 5% and 6%, usually close to 6%. Type-I error probabilities of the kurtosis test, although they are slightly low at the beginning, increased slightly with the increase in the sample size, reaching over 5%. On the contrary, in the Lilliefors test, while type-I error probabilities are above 5% at the beginning, type-I error probabilities decreases to around 4.5% as the sample size increased.

In a 1-degree-of-freedom chi-square distribution, the power performances of the other tests except the kurtosis test exceed 80% when the sample size is at least 25, and this performance approaches 100% when the sample sizes are enlarged. Power performances above 80% for the other tests except the kurtosis test in the chi-square distribution with 5 degrees of freedom are obtained only when the sample size is 100. In chi-square distributions with 1 and 5 degrees of freedom, increasing the sample size increases the power performances of the normality tests, while increasing the degrees of freedom decreases the power performances of the normality tests.

It is seen that the power performances of the normality tests are low in small sample sizes in the 2 and 5 parameter beta distribution. So much so that when $n = 100$, only the SW test is over 80%, and when $n = 150$, the power performances of the DP, skewness, SW and JB tests are over 80%. It has been observed that the power performances of the kurtosis test do not change much with the sample sizes and perform very poorly. In the 1 and 1 parameter beta distribution, as the sample size increases, the test powers also increase. However, when $n = 50$, it is seen that the powers of only the DP and kurtosis tests exceed 80% and the skewness test performs very poorly in general. In general, the power performances of the normality tests in the 2 and 5 parameter beta distribution are much better than the power performances obtained from the 1 and 1 parameter beta distribution. In addition, it is observed that the kurtosis test in the 2 and 5 parameter beta distribution and the skewness test in the 1 and 1 parameter beta distribution gave very bad results.

It is determined that the test powers in the 5-degrees-of-freedom t distribution are weak in all sample sizes and there is a certain improvement in the power values of all tests with the increase in sample sizes. The power values of the tests in the t-distribution with 10 degrees of freedom are quite weak in all sample sizes. The highest power values are obtained with the JB test, such that the JB reached a maximum with value of 77.69 in the 5-degrees-of-freedom distribution and 37.59 in the 10 degrees-of-freedom distribution.

When n≤70 in the uniform distribution, the power performance of the JB test and the skewness test in other sample sizes are quite low. When n≤50, only the power of the DP and kurtosis tests are exceeded 80%. The JB test, which performs very poorly in small sample sizes, reaches a very high power of 98.65% when n=150.

When the DP, kurtosis, skewness and JB tests from the moment tests are compared, it is seen that the skewness and SW tests are stronger than the other moment tests in non-normal distributions such as $\chi^2_{(1)}, \chi^2_{(5)}$ and beta (2, 5). In addition, it is observed that the DP test is more powerful than the JB test, and the JB, which gives bad results in small sample sizes, gives very good results in large sample sizes. In beta distributions, it is observed that the kurtosis and skewness tests are highly affected by the distribution parameters.

When the normality tests examined in this study are taken into account and compared in all distributions, it is determined that SW gave better results than other tests for normal and non-normal distributions, and DP, skewness and JB tests are also strong. In addition, it can be said that the increase in sample size increases the test power, and as a result, the sample size is an important parameter for the power of the test as well as the distribution pattern. Moreover, it was determined that SW test gave good results in all distributions, JB could be used as an alternative to SW in large sample sizes, and normality tests gave poor results in 5 and 10 degrees of freedom t distributions.

Thus, the study is capable of helping the selection of appropriate univariate normality tests to obtain more reliable results by considering the sample distribution and size in future studies.

ACKNOWLEDGEMENT

The author declares that there are no conflict of interests.

REFERENCES

[1] Noughabi, H. A., and Arghami, N. R. (2009). Monte carlo comparison of seven normality tests. Journal of Statistical Computation and Simulation, 81, 965-972.

[2] Adefisoye, J., Golam Kibria, B., and George, F. (2016). Performances of several univariate tests of normality: An empirical study. Journal of Biometrics & Biostatistics, 7, 1-8.

[3] Özer, A. (2007). Comparison of normality tests. Unpublished Master's Thesis, University of Ankara, Ankara.

[4] Dufour, J. M., Farhat, A., Gardiol, L., and Khalaf, L. (1998). Simulation-based finite sample normality tests in linear regressions. Econometrics Journal, 1, 154-173.

[5] Banerjee, A., Chitnis, U. B., Jadhav, S. L., Bhawalkar, J. S., and Chaudhury, S. (2009). Hypothesis testing, type I and type II errors. Industrial Psychiatry Journal, 18, 127-131.

[6] Eygü, H. (2020). Çözümlü güncel örneklerle olasılık ve istatistik. Ankara: Nobel Yayınları. ss.251-252.

[7] Elsayir, H. A. (2018). Factors determining the power of a statistical test for the difference between means and proportions. American Journal of Mathematics and Statistics, 8, 171-178.

[8] Ugar, H. (1999). Functions of random variables and applications to probability distributions. Unpublished Master's Thesis, University of Istanbul, Istanbul.

[9] Gordon, S. (2006). The normal distribution. Produced by UPS, Sydney.

[10] Lancaster, H. O., and Seneta, E. (2005). Chi-square distribution, p. armitage and t. Colton in Encyclopedia of Biostatistics, Wiley, Chichester.

[11] Karagöz, Y. (2003). Demonstration of relationship between exponential and chi-square distributions with random numbers produced by simulation. C.Ü. İktisadi ve İdari Bilimler Dergisi, 4(1), 197-209 .

[12] Dündar, D. (1987). Hipotez testi ve ki-kare ile bir uygulama. İstanbul Üniversitesi İktisat Fakültesi Mecmuası, 45, 196-212.

[13] Wackerly, D. D., Mendenhall, W., and Scheaffer, R. L. (2008). Mathematical statistics with applications. Cengage Learning, USA.

[14] Koepf, W., and Masjed-Jamei, M. (2006). A generalization of student's t-distribution from the viewpoint of special functions. Integral Transforms and Special Functions, 17, 863-875.

[15] Ramachandran, K. M., and Tsokos, C. P. (2009). Mathematical statistics with applications. Academic Press, USA.

[16] Saraçoğlu, Ö. (2002). Düşük akım hidrolojisi ve akdeniz bölgesi'nde uygulanması. Unpublished Master's Thesis, University of Istanbul, Istanbul.

[17] Cain, M. K., Zhang, Z., and Yuan, K. (2016). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. Behavior Research Methods, 49, 1716-1735.

[18] Nosakhare, U. H., and Bright, A. F. (2017). Statistical analysis of strength of w/s test of normality against non-normal distribution using monte carlo simulation. American Journal of Theoretical and Applied Statistics, 6, 62-65.

[19] Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrica, 52, 591-611.

[20] Sheskin, D. J. (2000). Handbook of parametric and nonparametric statistical procedures. Chapman & Hall/CRC, USA.

[21] D'Agostino, R. B., and Pearson, E.,S. (1973). Tests for departure from normality. empirical results for the distributions of b_2 and $\sqrt{b_1}$. Biometria, 60, 613-622.

[22] D'Agostino, R. B. (1986). In tests of the normal distribution, rb d'agostino and ma stephens, goodness of fit techniques. Marcel Dekker, Inc, New York.

[23] Zar, J.H. (1999). Biostatistical analysis. Prentice Hall, New Jersey.

[24] Dong, L. B., and Giles, D. E. A. (2004). An empirical likelihood ratio test for normality. Communications in Statistics-Simulation and Computation, 36, 197-215.

[25] Razali, N. M., and Yap, B. W. (2010). Power comparisons of some selected normality tests. Proceedings of the Regional Conference on Statistical Sciences, July 2010, Selangor, Malaysia. 126-138.

[26] Abdi, H., and Molin, P. (2007). Lilliefors test of normality, N.J. Salkind, Encyclopedia of measurement and statistics. SAGE Publications, Inc, California.

[27] Soest, J. V. (1967). Some experimental results concerning tests of normality. Statistica Neerlandica, 21, 91-97.

[28] Lilliefors, H. W. (1967). On the kolmogorov-smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association, 62, 399-402.

[29] Dagnelie, P. (1968). A propos de l'emploi du test de Kolmogorov-Smirnov comme test de normalité. Biométrie-Praximétrie, 9, 3-13.

[30] Genceli, M. (2007). Kolmogorov-smirnov, lilliefors and shaphiro-wilk tests for normality. Sigma Journal of Engineering and Natural Sciences, 25, 306-328.

JSR A
Journal of Scientific Reports

Korkmaz, S. and Demir, Y., Journal of Scientific Reports-A, Number 52, 376-395, March 2023.

[31] Akçadağ, H. İ. (2013). Tek değişkenli ve çok değişkenli bazı normallik testlerinin karşılaştırılması, Yayımlanmamış Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.

[32] Keskin, S. (2002). Varyansların homojenliğini test etmede kullanılan bazı yöntemlerin ı. tip hata ve testin gücü bakımından irdelenmesi, Yayımlanmamış Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

[33] Öztuna, D., Elhan, A. H., and Tuccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. Tübitak Academic Journals, 36, 171-176.

[34] Thadewald, T., and Büning, H. (2007). Jarque-bera test and its competitors for testing normality-a power comparison. Journal of Applied Statistics, 34, 87-105.

[35] Mendeş, M., and Pala A. (2003). Type I error rate and power of three normality tests. Pakistan Journal of Information and Technology, 2, 135-139.

[36] Shapiro, S. S., Wilk, M. B., and Chen, H. J. (1968). A comparative study of various tests for normality. Journal of the American Statistical Association, 63, 1343-1372.

[37] Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. Journal of the America Statistical Association, 69, 730-737.

[38] Seier, E. (2002). Comparison of tests for univariate normality. site file:///C:/Users/ASUS/Downloads/Comparison_of_tests_of_univariate_normality.pdf. Access Date: 20.12.2021.

[39] Razali, N. M., and Yap, B. W. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. Journal of Statistical Modeling and Analytics, 2, 21-33.