

**Statistical Normality Tests- A Guide for Non-Professional Statisticians**Okoroiwu, H.U.^{1*} and Okoroiwu, E.N.².Haematology Unit, Department of Medical Laboratory Science, University of Calabar, Nigeria¹,
Department of Computer Science, University of Nigeria, Nsukka².

Author for Correspondence*: +234- 803-883-3901/okoroiwuhenshaw@gmail.com

Summary

Statistical accuracy is essential in biomedical research. While choosing statistical tests, certain assumptions are made. Among the assumptions for choice for parametric statistics include assumption of normality. The confirmation of normality of data under treatment is essential for the validity of the test procedure. The aim of this review is to highlight the common tests of normality and their suitability.

Keywords: Normality tests, Tests of normality, Assumption of normality, Biostatistics, Statistical errors, Parametric statistics, Nonparametric statistics, Biometry.

Introduction

Statistical literacy is an essential skill for biomedical and clinical scientist as it helps them to draw conclusions from uncertain facts (Okoroiwu and Akwiwu, 2019; Nahm, 2016). Statistical errors are common in scientific literatures and about 50% of published articles have at least 1 error (Ghasemi and Zahediasl, 2012; Curan and Benos, 2004). In our earlier review, we discussed choice of parametric and nonparametric statistical procedures (Okoroiwu and Akwiwu, 2019). However, parametric procedures make certain assumptions which includes the assumption that the data under consideration follows Gaussian (normal) distribution (named after the German mathematician Karl Friedrich Gauss in 1809; 1777-1855); that is, the population from which the samples are drawn from are normally distributed (Gbasemi and Zahediasl, 2012; Altman and Bland 1995). From the visual perspective, the normal distribution, often referred to as the bell-shaped curve, is a graph of

frequency distribution can be described mathematically and observed empirically. The shape of the normal distribution is such that the closer the score is to the mean value, the more frequent it occurs. As score deviate more and more from the mean (that is becoming higher or lower), the lower the frequency with which that score occurs (Sheskin, 2011). When this assumption of normality is not valid, then it becomes difficult to draw an accurate and reliable conclusion about the reality (Gbasemi and Zahediasl, 2012; Oztuna *et al.*, 2006). Although the central limit theory states that when sample size has 100 or more observations, violation of the normality is not a major issue (Gbeseme and Zahediasl, 2012; Altman and Bland, 1995), however, for meaningful conclusions, assumption of the normality should be followed irrespective of the sample size. Hence the need to validate the assumption before using parametric tests. This review is aimed at discussing the test used in assessing the normality of data with non-statisticians in perspective.

Normality test

Assumption of normality is required by most statistical procedures (and they fall into the parametric procedures). Tests of normality are statistical inference procedures designed to test the underlying distribution of a random variable if it is normally distributed (D'Agustino *et al.*, 1990). Many normality tests have been developed by different authors. Since the invention of the chi-squared goodness-of-fit test for normality by Pearson in 1900, considerable attention has been given to the problem of testing normality and a fair number of tests can be found in the literature. Cramer von Mises and

Kolmogorov–Smirnov proposed their tests in 1931 and 1933, respectively. Following this, almost two decades after, Anderson and Darling suggested their test. Afterwards, Kuiper's and Shapiro and Wilk's tests of normality were introduced. Another new normality test 'Vasicek's test of normality' was suggested in 1976. Jarque and Bera proposed their test in 1987 (Noughabi and Arghami, 2011). Normality tests can be visual/graphical method, numerical method and formal statistical hypothesis normality tests. The visual /graphical method uses graphs or represent data for interpretation of normality and includes quantile -quantile plot (Q – Q plot), histogram, box-and-whisker plot, stem-and-leaf plot. The down side of graphical method is that it could be subjective in interpretation and require vast experience and good statistical knowledge for interpretation. Hence, they are not sufficient to provide conclusive evidence that the normal assumption holds, and are expected to be coupled with other tests for normality while interpreting. The numerical method includes the skewness and Kurtosis coefficients. The third category are formal statistical hypothesis testing normality tests that are constructed to be applied under certain conditions for assumptions. The common ones included in most statistical software are: Shapiro-Wilk test, Kolmogorov-Smirnov test, Anderson-Darling test, Lilliefors tests, Pearson's chi squared goodness of fit test, Cramer-Von Mises test, Shapiro-Francia test, Ryan-Joiner test, Jarque-Bera test and D'Agostino-Pearson K^2 test (Ghasemi and Zahediabl, 2012; Razali and Wah, 2011; Yap and Sim, 2011). Normality tests available in common statistical software are listed in table 1.

1. Visual/ Graphical method

As stated earlier, this method uses graphs and plot to represent normality/ deviation from normality and has the downside of objectivity. However, the good side is that when data is presented, usually, readers of an article can judge the distribution assumption themselves (Altman and Bland, 1996). Some examples of graphical method are presented below.

i. Histogram

Histogram is the easiest and simplest graphical plot (Das and Imon, 2016). It plots the observed values against their frequency, providing both

visual and judgment about whether the distribution is bell-shaped and insight about gaps in the data and outliers if present (Ghasemi and Zahediabl, 2012; Peat and Barton, 2005). Also, it gives insight about skewness or symmetry (Das and Imon, 2016). An example of representation of normality using histogram is show in figure 1

ii. Stem- and-leave plot

Stem-and-leave plots were described by John Tukey (Tukey, 1966) and was popularized by Edward Tufte in the “Visual display of quantitative information” (Tufte, 1983). Stem-and-leave plot provides an alternative to line plot or histogram for obtaining a picture of the data distribution when data can be represented as an integer. The information portrayed is identical as that of histogram, but it has an advantage that the information presented do not lose any information about the original data. Like the histogram, they show frequency of observation along with the median value, lowest and highest values of the distribution, other sample percentiles from the display of data (Das and Imon, 2016). The benefits of stem-and-leaf plot include: more information is retained than histogram; reveals fine structure while showing the distribution; allows easy hand calculation of measures based on ordered values (e.g. median, quantile). However, some of the limitations include: problem with large data set; whether extra digits are useful to the task and; comparison can be awkward (Brath and Banissi, 2017). Stem-and-leaf plot showing skewed and normally distributed data is represented in figure 2.

iii. Box – and – Whisker plot

Box-and-whisker plot were introduced by the eminent statistician Tukey (Tukey, 1977) for the purpose of giving rapid visualization of batches of data (Larsen, 1985). Box-and Whiskey plot has another name as five number summaries where it needs first quartile, second quartile or median third, third quartile, minimum and maximum values to display (Razali and Wah, 2011). Box-and-Whisker plot differ from ordinary plot in several ways which includes showing only values that are necessary to convey relative values of the data entries (Larsen, 1985). Figure 3 shows Box-and-Whiskers plot differentiating normally distributed and non-normally distributed data.

iv. Probability-probability plot (P-P plot).

The P-P plot is a graph of the cumulative probability (percentile) of a variable against cumulative probability (percentile) of a particular distribution (normal distribution in this case). It assesses how closely related two data sets agree. The plot gives an idea about outlier, skewness, kurtosis (Das and Imon, 2016; Holmgren, 2015; Gbasemi *et al.*, 2012). Here, if the sample is normally distributed, the result would be a straight diagonal line (Field, 2009). P-P plot of both normally and non-normally distributed data is shown in figure 4.

v. Quantile-quantile plot (Q-Q plot)

The Q-Q plot is a common tool to analyze the goodness of fit of a sample data to a theoretical distribution. In actual sense, it compares the quantiles of a data distribution with the quantile of a standardized theoretical distribution in this case, normal distribution (Das and Imon 2016; Velez and Morales, 2015). Basically, it allows the user to compare a theoretical model represented by a 45° slope line with an empirical quantile function represented by all sample points (Valez and Morales, 2015). The deviation from the diagonal line in the Q-Q plot imply that, in the extremes, the data extend farther out than expected under normality. However, there are several drawbacks to Q-Q plot as data exploratory tools which includes; sample variation makes it difficult to interpret the plot especially the behavior in the tails; the computational difficulties when the sample size is large. Kurtosis less than 3 indicates thicker tail and lower peak while kurtosis greater than 3 indicates higher peak and thin tail. Example of Q-Q plot of normally and non-normally distributed data is represented in figure 5.

2. Numerical method

The numerical method includes skewness and kurtosis coefficient.

i. Skewness

Skewness is a measure of the distribution asymmetry (von Hippel, 2010). It characterizes the degree of asymmetry of a given distribution around its mean (Cisar and Cisar, 2010). Skewness has also been defined mathematically

with respect to the third moment about the mean: which is simply the expected value of the distribution of cubed Z score. If the distribution of the data is symmetric, then skewness will be close to 0 as the skewness of a normal distribution is zero. A positive skewness indicates a distribution with an asymmetric tail extending toward more negative values. Conversely, negative skewness indicates a distribution with an asymmetric tail more negative value (Cisar and Cisar, 2010). In plain terms, skewness greater than zero, indicates skewness to the right and vice versa.

ii. Kurtosis

Pearson in 1905 introduced kurtosis as a measure of how flat the top of a symmetric distribution is when compared to a normal distribution of the same variance (Decarlo, 1997). Kurtosis is a measure of whether the data are heavily-tailed or light-tailed relative to a normal distribution. Mathematically, kurtosis is defined as the standardized population moment about the mean (DeCarlo, 1997). Kurtosis characterizes the relative peakedness or flatness of a distribution compared with the normal distribution. Positive kurtosis indicates a relatively peaked distribution while a negative kurtosis indicates a relatively flat distribution. On the other hand, data sets with high kurtosis tend to have heavy tails or outliers whereas data sets with low kurtosis have light tails or lack outliers (Cisar and Cirsar, 2010).

3. Formal statistical hypothesis normality tests

Aside the graphical and numerical methods that are used as complement, there are formal statistical hypothesis normality tests. These tests are based on: empirical distribution function (EDF); and descriptive measures (Regression and correlation), moments tests and the Pearson Chi-square statistics. Below are some of the common ones (Das and Imon, 2016). An example of SPSS output of normality test is shown in figure 6. When the p-value of the normality test is less than 0.05, it indicates that the data are non-normally distributed.

a. Correlation and regression tests

Regression tests are based on the fact that the

distribution function $F(x)$ of a random variable $X \sim N(\mu, \sigma)$ is a straight line when represented on a normal probability plot (Q-Q plot). Therefore, given an ordered sample of values $x(1) \cdot \cdot \cdot x(n)$ with $F(x(i)), \cdot \cdot \cdot F(x(n))$ values of their distribution function, points $(x(i), F(x(i)))$ should align approximately according to a straight line in a Q-Q plot, and any departure from that alignment indicates the data's lack of normality. The common ones are Shapiro-Wilk test and Shapiro-Francia (SF) test.

i. Shapiro-Wilk test (SW)

The Shapiro-Wilk test is one of the most popular tests of normality that has good power properties and it is based on correlation within given observations and associated normal scores. The test statistics was derived by Shapiro and Wilk in 1965 (Das and Imon, 2016). Shapiro-Wilk test is the ratio of the best estimator of the variance to the usual corrected sum of squares estimator of the variance. The statistics is positive and less than or equal to 1. Being closer to 1 indicates normality (Park, 2006). The original Shapiro-Wilk test was restricted to sample size ≤ 50 . However, the limitations were modified by Royston (1982) to broaden the restriction of the sample size to 2000.

ii. Shapiro-Francia test (SF)

The Shapiro-Francia test (Shapiro, 1972) can be visualized as the ratio of the sum of the square of a linear combination of the order statistics to the corrected sum of squares estimator of the variance. The criteria of Shapiro and Wilk (1965) and Shapiro and Francia (1972) have the same form, differing only in the definition of the coefficient. More so, Shapiro and Francia's modification of the original Shapiro-Wilk test model cater for sample size up to 5000 in the Shapiro-Francia test. However, it is pertinent to note that popular statistical programs such as SAS and SPSS do not support this statistic (Park, 2006; Sarkadi, 1975).

b. Empirical distribution (EDF) tests

Empirical distribution function tests were designed to compare the cumulative distribution function (CDF) of the normal distribution with the EDF estimated from the data to determine if there is agreement between them (Mbah and

Paothong, 2015; Razali and Wah, 2011). Dufour and Colleagues (1998) defined EDF tests as those based on a measure of discrepancy between the empirical and hypothesized distributions. The EDF tests can further be divided into those belonging into *supremum* and *square* class of the discrepancies (Razali and Wah, 2011). The common ones are Kolmogorov-Smirnov test, Anderson Darling test, Cramer-Von Mises test, and Lilliefors test (LF).

i. Kolmogorov – Smirnov test (KS)

The Kolmogorov-Smirnov test was first derived by Kolmogorov (1933) and later modified and proposed as a test by Smirnov (1945). It belongs to the *supremum* clan of empirical distribution function (EDF) tests and this clan of statistics is estimated based on the largest vertical difference between the hypothesized and empirical distribution. If the values are large, it indicates that the data is not normal. If the results are significant, then, the null hypothesis of no difference between the observed data distribution and a normal distribution is rejected. In plain terms, a value less than 0.05 indicates that the data are non-normally distributed (Yap and Sim, 2016; Razali and Wah, 2011). SPSS reports Kolmogorov-Smirnov test with Lilliefors correction.

ii. Anderson – Darling tests (AD)

The Anderson-Darling test was proposed in 1974 by Stephens who proposed the test based on empirical distribution by extending the work of Anderson and Darling (1952). The test is a modification of the Cramer-Von Mises test. It only differs from Cramer-Von Mises test in the sense that it gives more weight to the tails of the distribution than does Cramer-von Mises test. It is also distribution-free as the critical values do not depend on the distribution being tested (Mbah and Paothong, 2015; Razali and Wah, 2011; Yap and Sim, 2010).

iii. Cramer-Von Mises Test (CVM)

The Cramer-Von Mises test was developed by Cramer, Von-Mises and Smirnov (Yap and Sim 2010; Conover, 1999). The test judges the goodness of fit of a hypothesized distribution $F^x(x)$ compared with EDF $F_n(x)$ based on some defined statistic that is distribution free. It is an

alternative to Kolmogorov-Smirnov test (Yap and Sim, 2010).

iv. Lilliefors test (LF)

The Lilliefors test is a modification of the Kolmogorov-Smirnov test, and was developed by Lilliefors (Lilliefors, 1967). The test is best suited when the parameters of the hypothesized distribution are completely known. Nevertheless, sometimes it is difficult to initially or completely specify the parameters as the distribution is unknown. In such case, the parameters need to be estimated based on the sample data. Lilliefors compares the empirical distribution x with a normal distribution where its unknown μ and σ are estimated from the given sample data (Razalli and Wah, 2011; Lilliefors, 1967).

c. Moment tests

Moment tests are those derived from the recognition that the departure of normality may be detected based on the sample moments which are the skewness and kurtosis. The two common ones are D'Agostino-Pearson and Jarque-Bera tests.

i. D'Agostino-Pearson Omnibus test (DA)

D'Agostino in 1973 proposed the D'Agostino – Pearson K^2 test standing on the basis of skewness and kurtosis tests and these are also assessed through moments. The statistics follows a chi-squared distribution with two degrees of freedom if the population is from a normal distribution. A large value of K^2 leads to the rejection of the normality assumption (Das and Imon, 2016; Yap and Sim, 2010). It is often referred to as omnibus test where “omnibus” means it is able to detect deviations from normality due to either skewness or kurtosis.

ii. Jarque-Bera test (JB)

The Jarque-Bera test was originally proposed by Bowman and Shenton (1975). The test procedure combined squares of normalized skewness and kurtosis in a single distribution with two degrees of freedom. Jarque-Bera statistics equals zero when the distribution has zero skewness and kurtosis is 3. The null hypothesis is that the skewness is zero and kurtosis is 3. Large values of skewness and kurtosis far from 3 lead to

rejection of null hypothesis of normality (Das and Imon, 2016; Yap and Sim, 2011). When normality is rejected using the JB test, one also obtains information on the process: the distribution may be skewed to the right (or to the left) and/or exhibit long (or short) tails.

d. Chi square goodness of fit test (CSQ)

It is the oldest and well know goodness of fit test. It was first presented by Pearson (Conover, 1999; Pearson, 1900). However, it is not highly recommended for continuous distribution since it uses only the counts of observations in each cell rather than the observations themselves in computing the test statistics (Yap and Sim, 2011).

Power comparison among the normality test

A test normality test is said to be powerful when it has a high probability of rejecting the null hypothesis of normality when the sample under study is taken from a non-normal distribution. In making comparison, all tests should have the same probability of rejecting the null hypothesis when the distribution is truly normal (that is, having the same type 1 error) (Yap and Sim (2011).

In a simulated power comparison between 8 normality tests, *Yap and Sim (2011)* showed that for asymmetric (i.e skewed) distribution, Shapiro-Wilk test had the highest power followed closely by Anderson-Darling test. However, for symmetric distribution with short-tailed distribution (low kurtosis), D'Agostino Pearson Omnibus test and Shapiro-Wilk tests have good power, whereas in symmetric long-tailed distribution (high kurtosis), Jarqua-Bera test, D'Agostino Pearson Omnibus test and Shapiro-Wilk test had comparable power. Their result also showed that Kolmogorov-Smirnov test and Chi-square goodness of fit performed poorly. Similarly, D'Agostino et al. (1990) had previously reported Kolmogorov-Smirnov test and Chi-square goodness of fit test as having poor power. Overall, Shapiro Wilk test has been reported as the most powerful normality test for non-normal distribution (Razali and Wah, 2011; Yap and Sim, 2011; Oztuna *et al.*, 2006).

Normality testing and statistical packages

The advent of the statistical programmes and packages has made it possible for non-professional statisticians to perform statistical

analysis with few lines of click and “OK”. Listed in table 2 below are some popular statistical packages and test of normality available in them

Table 1: Some common statistical software and normality tests present in them.

Statistical Software	Tests										
	SW	SF	KS	LF	CVM	AD	JB	CSQ	RJ	SKKU	DA
SAS	+		+		+	+					
SPSS	+		+	+							
SPLUS	+	+	+					+			
STATISTICA	+		+	+				+			
STATA	+	+								+	+
STATGRAPHICS	+		+	+	+	+		+		+	
MIINITAB			+			+	+		+		
MATLAB		+	+	+	+	+	+	+			
R	+	+	+	+	+	+	+	+			+
Python	+					+					+
IMSL Library	+			+				+			
Epi Info											

+: Present

SW: Shapiro-Wilk; SF: Shapiro-Francia; KS: Kolmogorov – Smirnov test; LF: Lilliefors test; CVM: Cramer-Von Mises Test; AD: Anderson – Darling tests; JB: Jarque-Bera test; CSQ: Chi square goodness of fit test (CSQ); RJ: Ryan-Joiner; SKKU: Skewness–Kurtosis test; DA: D'Agostino -Pearson Omnibus test.

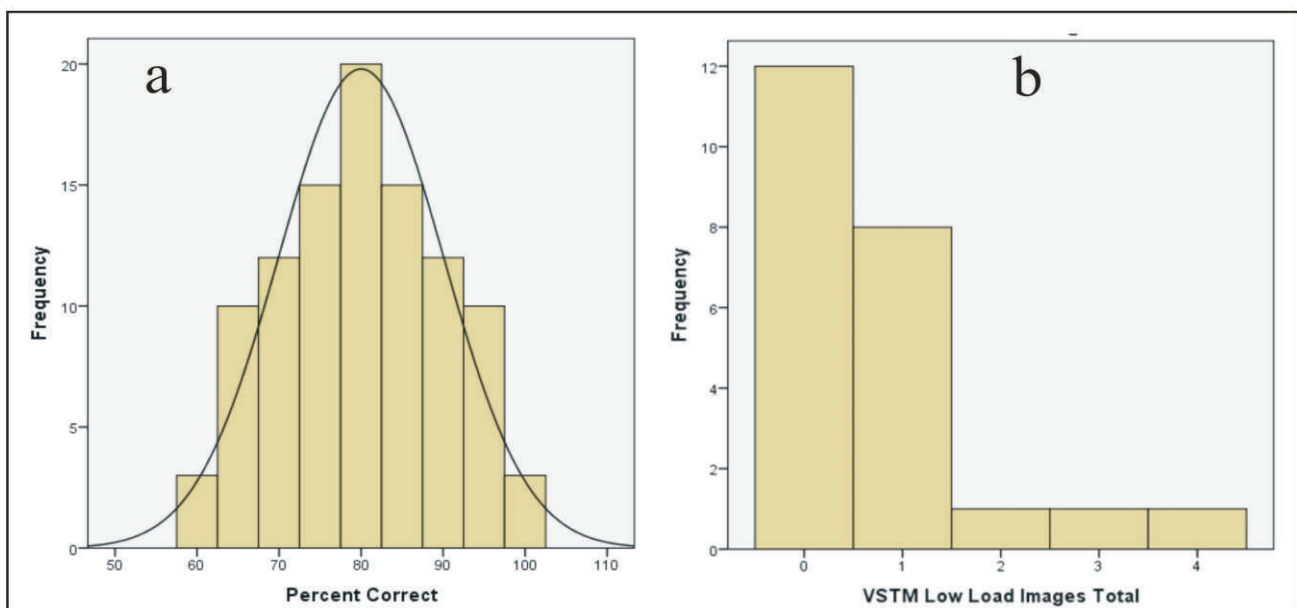


Figure 1: Histogram showing normally and non-normally distributed data.

(a): Normally distributed data.

(b): Non-normally distributed data

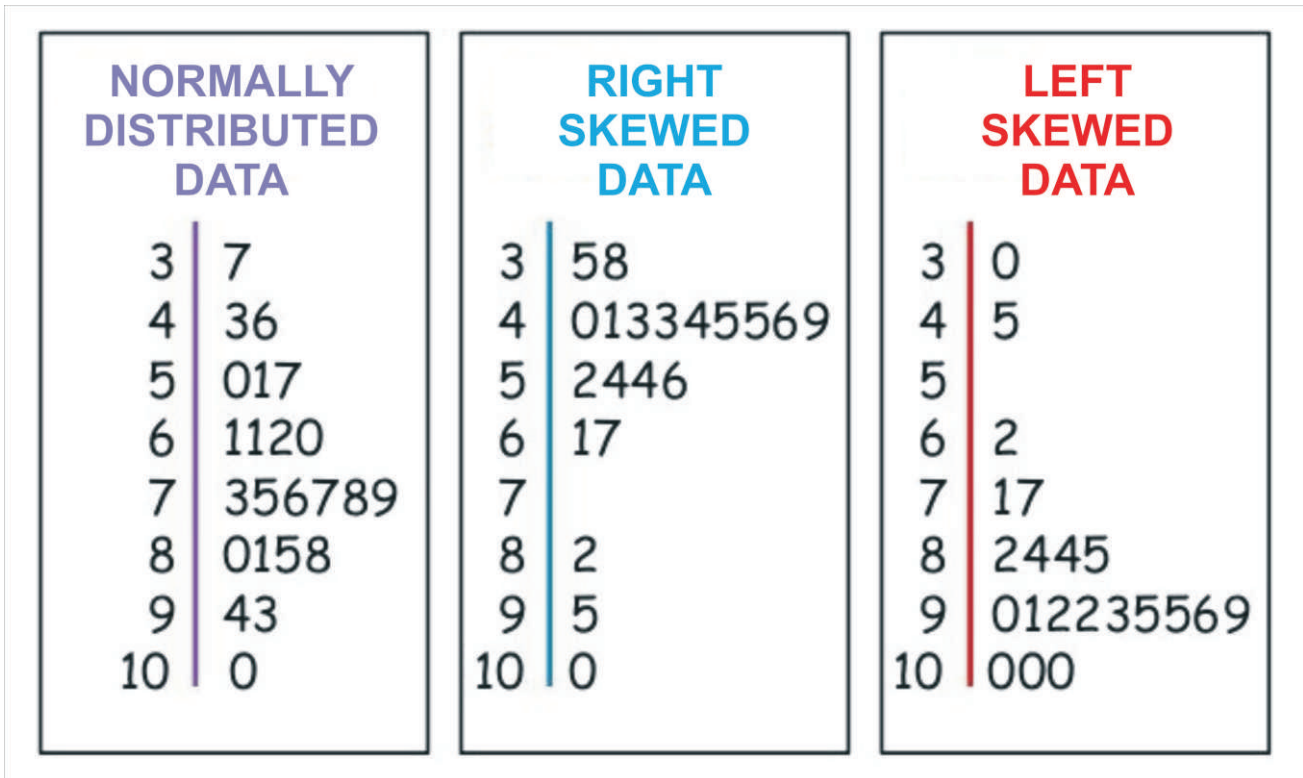


Figure 2: Stem-and-leaf plot showing normally and non-normally distributed data.

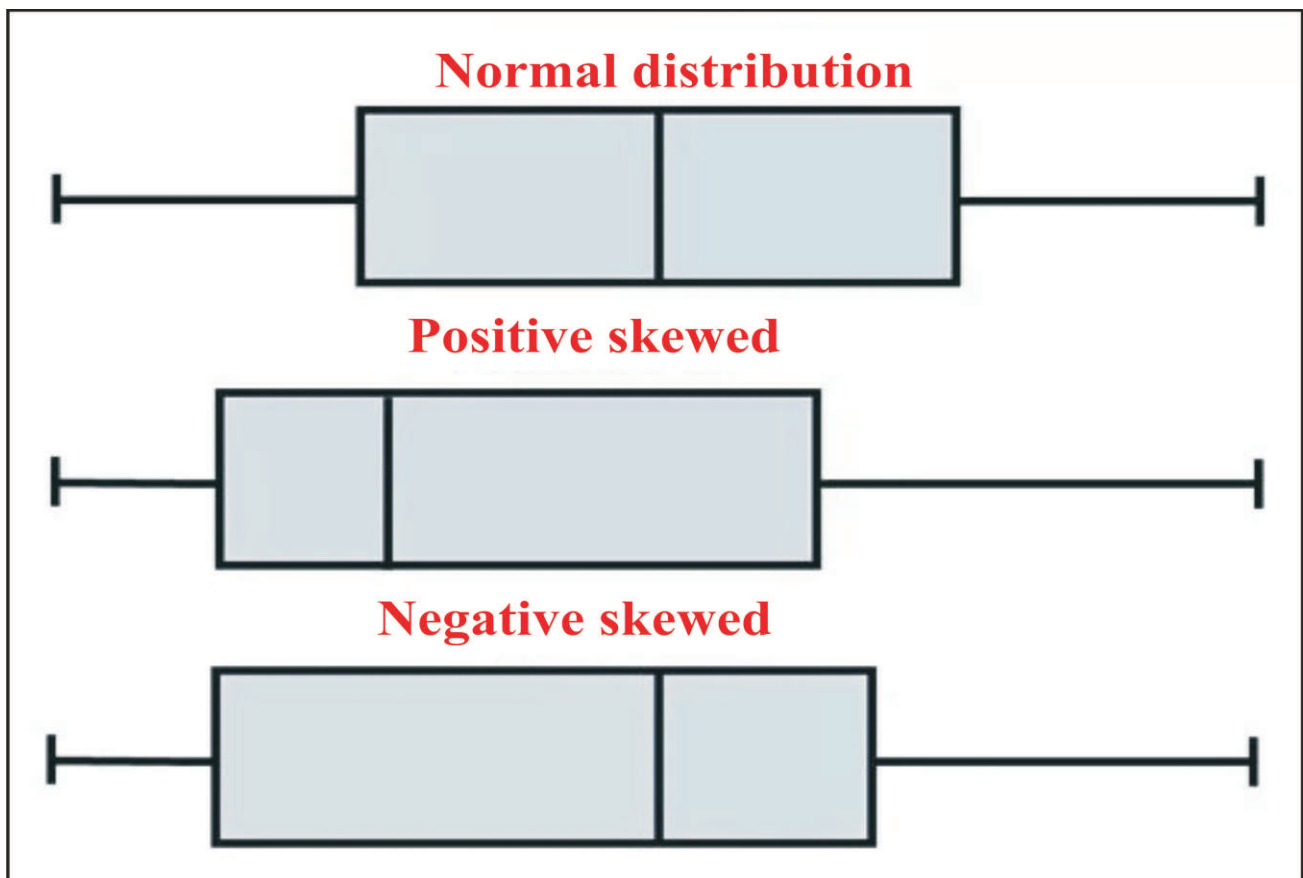


Figure 3: Box-and-whisker showing normally and non-normally distributed data.

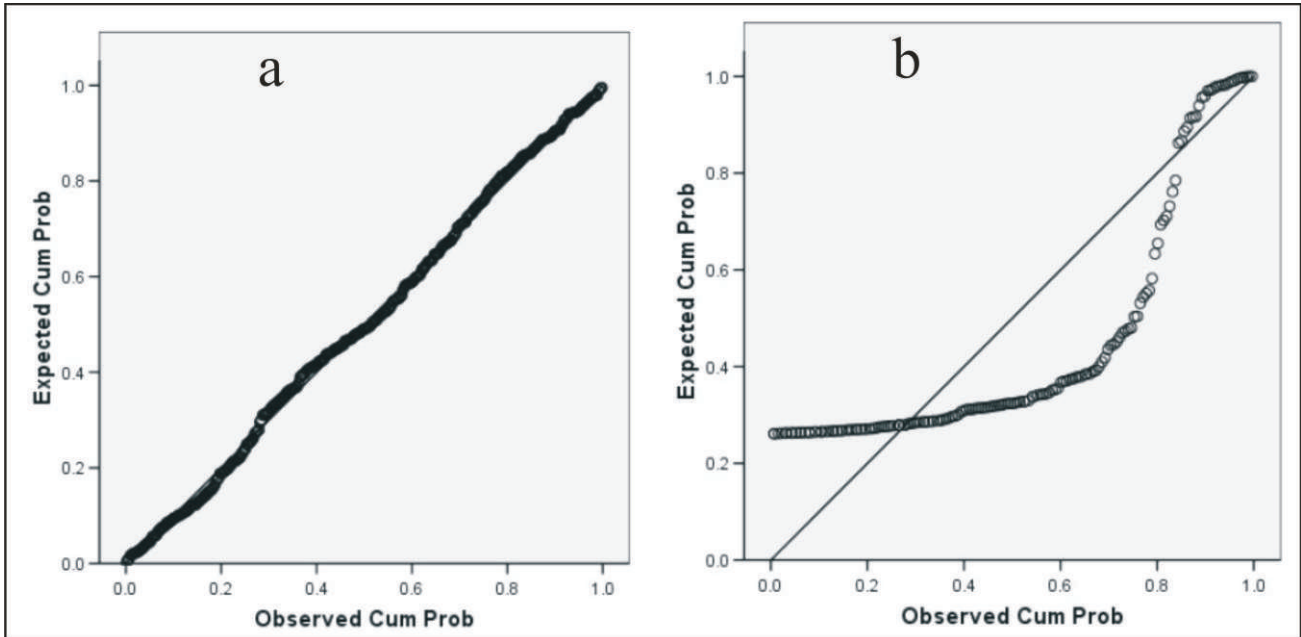


Figure 4: P-P plot showing normally and non-normally distributed data.
(a): Normally distributed data.
(b): Non-normally distributed data.

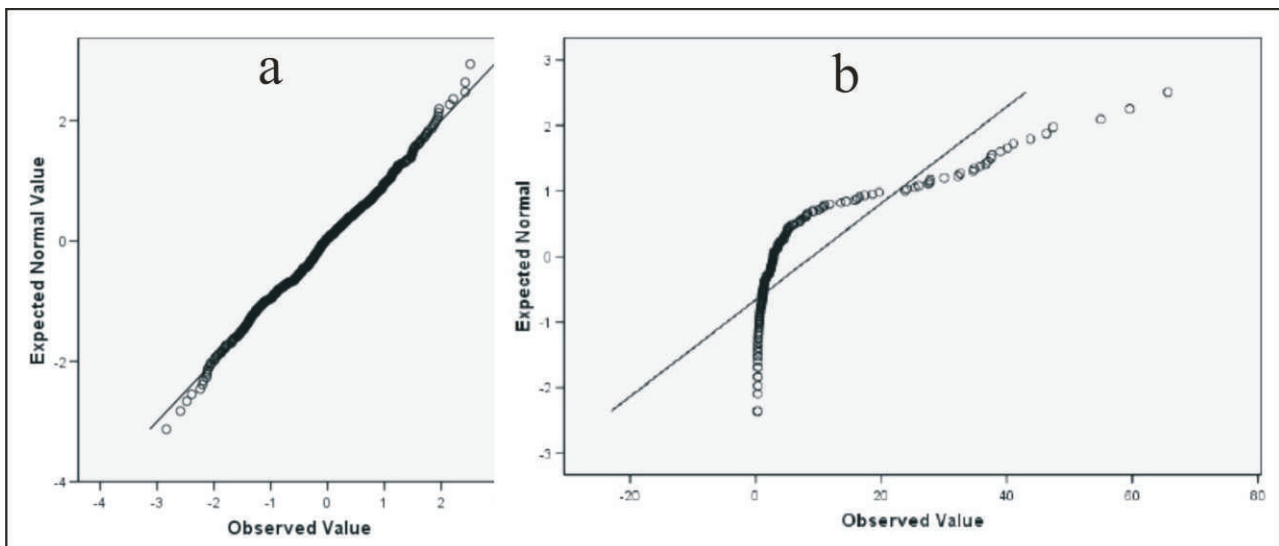


Figure 5: Q-Q plot showing normally and non-normally distributed data.
(a): Normally distributed data.
(b): Non-normally distributed data

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
GSH	.210	180	.000	.823	180	.000

a. Lilliefors Significance Correction

Figure 6: SPSS output showing normality test output.



References

- Altman, D.G. and Bland, J.M. (1996). Detecting skewness from summary information. *British Medical Journal*; 313 (7066): 1200.
- Altman, D.G., Bland, J.M. (1995). Statistical notes; the normal distribution. *British Medical Journal*; 310 (6975): 298.
- Bowman, K.O. and Shenton, B.R. (1975). "Omibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2 ". *Biometrics*; 64: 243–250.
- Brath, R. and Banissi, E. (2017). Stem and leaf plot extended for text visualization. 14th International Conference on Computer Graphics, Imaging and Visualization. Marrakesh, Morocco 23 - 25 May 2017 London South Bank University. pp. 99-104.
- Cisar, P. and Cisar, S.M. (2010). Skewness and kurtosis in function of selection of network traffic distribution. *Acta Polytechnica Hungarica*; 7(2):95-106.
- Conover, W. J. (1999). Practical nonparametric statistics. 3rd ed. New York: John Wiley and Sons: 126-216.
- Currans-Everett, D., Benos, D. J. (2004). Guidelines for reporting statistics in journal published by American Physiological society. *American Journal of Physiology, Endocrinology and Metabolism*; 2004: E 189–91.
- D'Agustino, R.B., Belanger, A. and D'Agustino Jnr, R.B. (1990). A suggestion for using powerful and informative test of normality. *The American Statistician*; 44(4): 316–321.
- DeCarlo, L.T. (1997). On the meaning and uses of kurtosis. *Psychological Methods*; 2(3): 292–307.
- Desgagne, A. and de Micheaux, L. (2017). A powerful and interpretable alternative to Jarque-Bera test of normality on second power skewness and kurtosis, using Rao's score test on the APD family. *Journal of Applied Statistics*; 45(13): 2307-2327.
- Field, A. (2009). Discovering statistics using SPP. 3 ed. London: Sage publications Ltd.: 822.
- Gbasemi, A., Zahediasl, S. (2012). Normality test for statistical analysis: a guide to non-statisticians. *International Journal of Endocrinology and Metabolism*; 10(2): 486–489.
- Holmgren, E.B. (1995). The P.P plot as a method for comparing treatment effects. *Journal of the American Statistical Association*; 90(429): 359–365.
- Larsen, R.D. (1985). Box – and – whisker plots. *Journal of Chemical Education*; 62(4): 302–305.
- Lilliefors, H.W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of American Statistical Association*; 62: 534–544.
- Mbah, A.K. and Paothong, A. (2015). Shapiro-Francia test compared to other normality test using expected p-value. *Journal of Statistical Computation and Simulation*; 85(15): 3002-3016.
- Naham, F.S. (2016). Non parametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*; 69(1): 8–14.
- Noughabi, H.A. and Arghami, N. R. (2011). Monte Carlo comparison of seven normality tests *Journal of Statistical Computation and Simulation*; 81(8):965-972.
- Okoroiwu, H.U., Akwiwu, E.C. (2019). Choice of parametric and non-parametric statistical procedures in clinical and biomedical research. *Sokoto Journal of Medical Laboratory Science*; 4(2): 5–15.
- Oztuna, D., Elhan, A. H., Tuccar, E. (2006). Investigation of four different normality test in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences*; 36(3); 171-176.
- Park, H.M. (2006). Univariate analysis and normality test using SAS, STATA, and SPSS. The trustees of Indian University.
- Peat, J. and Barton, B. (2005). Medical Statistics: a guide to data analysis and critical appraisal. Blackwell Publishing.
- Razali, N.M. and Wali, Y.B. (2011). Power comparisons of Shapiro – Wilk, Kolmogorov – Smirnov, Lilliefors and Anderson – Darling tests. *Journal of Statistical Modelling and Analysis*; 2(1):21–33.
- Royston, J.P. (1982). An extension of Shapiro-Wilk test for normality to large samples. *Applied Statistics*; 31:115-124.
- Royston, J.P. (1995). Remark AS R94: A remark

- on algorithm AS 181: The W-test for normality. *Applied Statistics*; 44:547-555.
- Sanchez-Epigares, J.A., Grima, P. and Marco-Almagro, L. (2018). Graphical comparison of normality tests for unimodal distribution of data. *Journal of Statistical Computation and Simulation*; DOI: 10.1080/00949655.2018.1539085.
- Sarkadi, B.K. (1975). The consistency of Shapiro-Francia test. *Biometrika*; 62(2):445-450.
- Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*; 52:591 – 611.
- Shapiro, S.S. (1972). An approximate analysis of variance test for normality. *Journal of American Statistical Association*; 67:215-216.
- Stine, R.A. (2016). Explaining normal quantile – quantile plots through animation; The water falling analogy. Department of Statistics. The Wharton School of the University of Pennsylvania, Philadelphia.
- Tufte, E. (1983). Visual display of quantitative information. Graphic Press.
- Turkey, J. (1966). Some graphic and semigraphic displays. “Statistical papers in honour of George W. Snedecor: 5.
- Turkey, J.W. (1977). “Exploratory Data Analysis”, Addison – Welsely, Reading, MA, 1977.
- Velez, J.I. and Morales, J.C.C (2015). A modified Q – Q plot for large sample sizes. *Comunicaciones en Estadística*; 8(2): 163 – 172.
- Von Hippel, P. (2010). Skewness. In: Lovric, M. (ed). International Encyclopedia of Statistical Science. New York: Springer.
- Yap, B.W. and Sim, C.H. (2011). Comparison of various types of normality tests. *Journal of Statistical Computation and Simulation*; 81 (12): 2141 – 2155.

Citation: Okoroiwu, H.U. and Okoroiwu, E.N. Statistical Normality Tests- A Guide for Non-Professional Statisticians. *Sokoto Journal of Medical Laboratory Science*; 4(3): 90 - 99.

Copyright. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.