Research article

Ultra-fine transformation of data for normality

Mohammad M. Hamasha a, **, Haneen Ali b, *, Sa’d Hamasha c, Abdulaziz Ahmed d

** Department of Industrial Engineering, Faculty of Engineering, The Hashemite University, B.O.Box 330127, Zarqa 13313, Jordan
* Health Services Administration Program, Auburn University, Auburn, AL, USA
a Department of Industrial Engineering, Faculty of Engineering, The Hashemite University, B.O.Box 330127, Zarqa 13313, Jordan
b Department of Industrial Engineering, Auburn University, Auburn, AL, USA
c Department of Health Services Administration, University of Alabama at Birmingham, Birmingham, AL, USA

d Department of Industrial Engineering, Faculty of Engineering, The Hashemite University, B.O.Box 330127, Zarqa 13313, Jordan

ARTICLE INFO

Keywords:
Normal distribution
Data normalization
Transform for normality
Ultra-fine transformation

ABSTRACT

Normally distributed data is crucial for the application of large-scale statistical analysis. To statisticians, the most important assumptions of statistical users are the adequacy of the data and the normal distribution of the data. However, users are constantly forced to deal with unusual data. This includes changing the method used to be less sensitive to non-normal data or transforming that data to normal data. In addition, common mathematical transformation methods (for example, Box-Cox) do not work on complex distributions, and each method works on limited data shapes. In this paper, a novel approach is presented to transform any data into normally distributed data. We refer to our approach as the Ultra-fine transformation method. The article’s novelty is that the proposed approach is powerful enough to accurately transform any data with any distribution to the standard normal distribution. Besides this approach’s usefulness, it is simple in both theory and in application, and users can easily retrieve the original data from its transformed state. Therefore, we recommend using this method for the data used in the statistical method, even if the data are normal.

1. Introduction

Many authors claim that the normal distribution is the most important probability distribution in statistics [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. This is due to three reasons: 1) Many phenomena in nature are distributed naturally [18]; 2) Many statistical analyses require the application of normally distributed data; and 3) Lastly, many statistical applications, such as statistical quality control, need normally distributed data to be applied.

The normality is required to perform a t-test, F-test, and χ2-test. The researchers reported and detailed the impact of non-normal data on different statistical analysis methods. In the next two paragraphs, we quickly discuss the effect of non-normal data on two statistical methods identified as examples, analysis of variance (ANOVA) and statistical process control (SPC).

Normally distributed data are required for all types of ANOVA [19]. However, the literature on one-way fixed-effect ANOVA is selected as an example. Further, a failure to obtain normalized data leads to disregard of this method and use of alternative, less sensitive methods for normality [19], such as the Brown–Forsythe test [20], the non-parametric Kruskal Wallis [21], or the Welch test [22]. Alexander & Govern [23] developed an approximation to a one-way fixed-effects ANOVA to reduce the effect of non-normal data. Guo & Luh [24] extended the Alexander and Govern model to minimize the non-normality impact further. They recommended the use of trimmed means instead of means.

The assumption of normality is very important for the application of SPC [25, 26]. For example, a potentially serious error is made if quality characteristics (data) are plotted on a Shewhart control chart [25]. Yourstone & Zimmer [26] analyzed the abnormal quality characteristics and showed that they lead to severe effects of the measures of the control chart and associated probabilities.

To get around the problem of non-normal data, statisticians recommend using sample means rather than individual units [27, 28, 29, 30, 31]. According to the central limit theorem, the distribution of the sample means with sufficient sample size is approximately normal even the original data is non-normal. However, there is not enough data to make samples to use the central limit theorem in many cases, especially in service. Therefore, statisticians have worked to develop some control charts that are less sensitive to non-normal data, such as EWMA [27], multivariate EWMA [28], median EWMA [29], and CUSUM [30, 31].

Finally, we can find reports of non-normal data problem everywhere and in all areas [32, 33, 34, 35, 36, 37]. For example, the effect of

* Corresponding author.
** Corresponding author.
E-mail addresses: mhamasha@hu.edu.jo (M.M. Hamasha), hba0007@auburn.edu (H. Ali).
non-normality has been reported in the following statistical methods: variables sampling plans [32, 33], the cutscore operating function [34], structural equation models [35, 36], economical design [37].

Evaluating the normality of data is a prerequisite for many statistical tests [38]. Methods for assessing normality can be divided into two main groups: graphic and numerical. Graphical methods provide subjective judgment on distributions, while numerical methods provide objective judgment [39]. However, graphical methods require significant experience in interpreting the data and determining the level of normality. Furthermore, numerical methods have a significant drawback: the lack of sensitivity for small sample sizes and excessive sensitivity for large sample sizes. In contrast, graphical methods do not suffer from hyposensitivity or hypersensitivity. The graphical methods include histogram, steam-and-leaf plot, boxplot, detrended probability plot, quantile-quantile (Q-Q) plot, and probability-probability (P-P) plot [40, 41, 42]. The numerical methods include Kolmogorov-Smirnov (K-S) test, Shapiro-Wilk test, Lilliefors corrected K-S test, Anderson-Darling test, D’Agostino skewness test, Cramer-von Mises test, D’Agostino-Pearson omnibus test, Ancombe-Glynn kurtosis test, and the Jarque-Bera test [42, 43].

A different statistical test requires an extra level of normality. For example, the data should be normal for a t-test, while a slight violation of normality is acceptable for the EWMA control chart. However, significantly non-normal data is invalid for the EWMA control chart. As discussed before, the statisticians attempted to handle each normal requirement of the test alone by minimizing the effect of the non-normal data, for example, modifying the method used.

However, one of the solutions used to deal with non-normal data is to transform the data into a data set closer to normal, as discussed in the following paragraphs.

Since the early twentieth century, statisticians have been thinking about transferring non-normally distributed data to normal, and several transformation formulas have been developed. Most of the developed formulas are mathematically based. The following paragraphs will present five of the most popular transformation methods: Log, Box-Cox, Yeo-Johnson, and Square-root transformations methods.

Log Transformation: This method’s ease and efficiency is always the first choice to normalize the data. It works if the data curve is smooth and skewed to the right only. Any time data consist of time, distance or money are more likely to be skewed to the right, and the log transformation is the choice. The idea of this transformation is to add a log to the data with a constant multiplier. After having this replacement done, the normality test must be conducted. If the data is still not normal or normality is not enough, other methods can be considered [44, 45, 46].

Box-Cox Transformation: Box-Cox is another transformation method introduced by (Box and Cox, 1964). They developed their formula based on the assumption that the relationship between the transformed data and the original data are subjected to the following formula [47].

\[ y' = \beta x + \epsilon. \] (1)

Their model is expressed further in the following two equations. Eq. (2) is for the positive value of \( y \), and Eq. (3) is for the negative value of \( y \).

\[
y(\lambda) = \begin{cases} 
\frac{(y^\lambda - 1)}{\lambda}, & \lambda \neq 0 \\
\log y, & \lambda = 0 
\end{cases}
\] (2)

\[
y(\lambda) = \begin{cases} 
\frac{(y + \lambda_0)^\lambda - 1}{\lambda}, & \lambda \neq 0 \\
\log(y + \lambda_0), & \lambda = 0 
\end{cases}
\] (3)

Practically, it was found that this model works fine for the range of \( \lambda \) of (-3, 3). However, it is not the right method for large numbers of \( \lambda \), say \( \lambda = 100 \).

Yeo-Johnson Transformation: This transformation formula is introduced by Yeo & Johnson [48], as an enhanced version of the Box-Cox model. Although the results of this method are better than Box-Cox, the same limitations are noticed. See Eq. (4).

\[
y(\lambda, y) = \begin{cases} 
\frac{[(y + 1)^\lambda - 1]}{\lambda}, & \lambda \neq 0 & y \geq 0 \\
\log(y + 1), & \lambda = 0 & y \geq 0 \\
\frac{[-(y + 1)^\lambda - 1]}{(2 - \lambda)}, & \lambda \neq 0 & y < 0 \\
\log(-y + 1), & \lambda = 0 & y < 0 
\end{cases}
\] (4)

Reciprocal Transformation: This method works if the data is skewed to the right. It maps the data point (x) to (b/x) if positive and to (-b/x) if negative, where b is constant [49].

Square-Root Transformation: This method takes the square root of each data point. In general, we use square root to transform the data if the data point count is small. Therefore, this method cannot handle the negative point directly. However, we may add a constant to the data to make it all positive before applying this method [49].

Note that these methods and formula requires a rigorous selection of optimal parameters. While these parameters reduce the problem of non-normality, they are not necessary to give a complete solution. For example, if two peaks appear in the probability distribution due to the overlap of two affecting variables, no one of these methods can give a fine normally distributed transform data. To the best of the authors’ knowledge, no mathematical solutions are available to the complicated data distribution.

In the next section (Section 2), we will present a transformation model. By using the provided model, it is guaranteed to have precisely normally distributed transformed data. Moreover, the simplicity of the model and the absence of any kind of restrictions is an additional strength.
to it. In fact, this makes it very practical. In Section 3, we present the application of the model using two numerical cases. In Section 4, we present the results and discuss their implications for the use of the model. In Section 5, we present an industrial application of the model. The application is from Industrial and Systems Engineering. Finally, in Section 6, we present the conclusions on the application of the transformation model.

2. Ultra-fine transformation to normality

Suppose we have an ascending set of independent observations, \( X \) (i.e., \( x(1), x(2), \ldots, x(n) \)), and the elements of the set are assigned to the corresponding rank in the set of ranks, \( R \) (i.e., 1, 2, ..., \( n \)), as expressed in Eq. (5). In other words, \( x_1 \) is mapped to 1, \( x(2) \) is mapped to 2, ... \( x(n) \) is mapped to \( n \). The rank array is transformed into a \( T \) array according to Eq. (6). Combining Eqs. (5) and (6), each element in the observation set is mapped to its corresponding value in the transformed data set, \( T \) that follows the standard normal distribution. The elements of the \( T \) set are normally distributed with a mean of 0 and a standard deviation of 1. If the data set is large enough, then the transformed data is an exact standard normal distribution. The first and last numbers of the \( T \) set are negative infinity and positive infinity, respectively.

\[
\begin{align*}
X & \rightarrow R \quad \text{(5)} \\
R & \rightarrow T, \text{ where } T = \Phi^{-1}\left(\frac{R-1}{n-1}\right) \quad \text{(6)}
\end{align*}
\]

After converting the data, it can be easily returned to the original values by reversing Eqs. (5) and (6), as shown in Eqs. (7) and (8) and Figure 1.

\[
R = \Phi(T) \times (n - 1) + 1 \quad \text{(7)}
\]

\[
R \rightarrow X \quad \text{(8)}
\]

Since the model deals with the order rather than the value of the data itself, the model can handle any distribution regardless of the shape of the data distribution.

3. Numerical examples

To explain the model further, two numerical cases will be explained. Numerical Case 1. Assume we have the following 12 observations: 1, 1, 3, 5, 12, 13, 14, 15, 15, 15, 16 and 18. The histogram (Figure 2) shows that

![Figure 2. Histogram of observations for the numerical case 1.](image)

The P-P plot (Figure 3) of the observations is shown in Figure 4. The transformed data is shown in Figure 5.

![Figure 3. P-P plot of observations for the numerical case 1.](image)

![Figure 4. P-P plot of transformed data for the numerical case 1.](image)

![Figure 5. Histogram of observations for the numerical case 2.](image)

### Table 1. Observation, rank, and transformed observation values for numerical case 1.

| Observation | Rank | Relative Rank, \( \frac{R-1}{n-1} \) | Transformed Observations (Z-score) |
|-------------|------|-----------------------------------|-----------------------------------|
| 1           | 1    | 0.00                              | \( \infty \)                     |
| 1           | 2    | 0.17                              | -1.33518                          |
| 3           | 3    | 0.27                              | -0.90846                          |
| 5           | 4    | 0.36                              | -0.60459                          |
| 12          | 5    | 0.45                              | -0.34876                          |
| 13          | 6    | 0.54                              | -0.11419                          |
| 14          | 7    | 0.63                              | 0.114185                          |
| 15          | 8    | 0.72                              | 0.348756                          |
| 15          | 9    | 0.81                              | 0.604585                          |
| 15          | 10   | 0.90                              | 0.908648                          |
| 16          | 11   | 1.00                              | 1.333178                          |
| 18          | 12   | 1.19                              | \( \infty \)                      |

Numerical Case 2. Assume we have the following 12 observations: 2, 2, 3, 4, 7, 8, 9, 9, 10, 11, 12, 15. The histogram (Figure 5) shows that...
the distribution is very far from normal. Moreover, the P–P plot (Figure 3) shows a non-linear relationship between two possibilities, and thus the data is non-normal. Moreover, the level of normality can be indicated by the P-value in the upper right corner. If the p-value is less than 0.05, then the data is not normal. However, the data becomes closer to the theoretical normal distribution as the p-value approaches 1.00. The p-value is 0.012, which confirms the non-normality of the current data. The first step is to arrange the observations in ascending order and label each one in order from 1 to 12; they are already arranged, as shown in Table 1. The ranks in the second column are substituted into Eq. (6) to obtain the transformed data, as shown in the third column. Figure 4 shows the probability plot of the transformed data without infinite quantities (i.e., the first and last transformed data points) as Minitab cannot take it. The figure indicates that the distribution is normally distributed with a p-value of 0.997, which is very close to 1.00. However, the model is limited to continuous distribution and not designed for discrete distribution.

Numerical Case 2. This case contains 501 observations, as shown in Appendix 1. Due to the large data volume of the second case, we preferred to place the spreadsheet in the appendix rather than the text. Figure 5 shows a right-skewed histogram. The distribution is closer to the exponential than normal. P–P Plot (i.e., Figure 6) shows non-normal data.

The data is arranged in ascending order and assigned in its correct rank from 1 for the minimum value to 501 (i.e., n) for the maximum value. Then, the relative rank is calculated from the absolute rank according to the following Equation, (R-1)/(n-1). After that, the obtained values are reversed using the cumulative standard normal distribution equation.

The transformed data distribution is an exact standard normal distribution, as shown in Figure 7 and Figure 8. Figure 7 represents the histogram of the transformed data, and Figure 8 illustrates the P–P plot of the transformed data.

The first and last transformed data points are ignored, and the rest are plotted. The p-value is 1.00 as seen in the upper right corner of Figure 8, confirming the precise normality of the transformed data. Lastly, we can return the transformed data point to its original observation at any time using the inverse of the model we proposed.

4. Results and discussion

Obtaining normal data by transforming non-normal data using previously the mentioned methods is not guaranteed and often requires rigorous parameters to transform appropriately. In general, successfully transforming non-normal data depends on the format of the data and the method used. Most methods shift the data center to the middle and reduce skewness. Although these methods increase the normality of the distribution, rigor is unnecessary, and the results do not consistently yield a p-value less than 0.05 in a normality test. Therefore, this gap necessitated the development of a model that works based on the locations of the observations in their ascending array. The novelty of this model is that it can deal with any data and convert them into a 100% standard normal distribution with no limitation and regardless of the data’s level of complexity. Moreover, the methodology is easy both in theory and in application and can be used manually or with the assistance of an Excel spreadsheet in addition to advanced software such as R and C++. We also recommend including this model in statistical software, such as SPSS and Minitab. However, the following gab was not covered in the current research. Although we recommend using this transformation even if the raw data is nearly normal, the level of raw data normality that makes this model not an added value for different applications has not been taken into account. This needs more research.

The model also gives the flexibility to estimate the original observation from the transform observation. The idea is simply to use Eq. (7) and then Eq. (8). Consider, for example, the transformed observation of 0.23 in Numerical Case 2, the relative rank is estimated using \( \Phi(0.23) \) (i.e., 0.590954). Then, rank \( (R) \) is estimated by applying \( \left( \text{Relative Rank}\right)^*(501-1)+1 \), which leads to the answer of \( R = 296.475 \). The observation corresponding to the \( R = 296.475 \) is 1.63. Linear interpolation can be included in the values corresponding to integers from both sides. However, for this example, the corresponding values for \( R = 296 \) and \( R = 297 \) are the same (i.e., 1.63), and therefore, no interpolation is required.

If necessary, the number of data can be reduced to 50–75, and the accuracy remains very high. Specifically, we recommend reducing the

| Sample size | n > 100 | 100 > n |
|-------------|---------|---------|
| s           | Integer of \((a/50)\) | 1 (No reduction) |

Table 2. Estimate S.
data if the data volume is huge, say a hundred thousand. Handling 100 data points, for example, is much easier than a huge data set. The following discussion will present a procedure to reduce the data to less than 100 data points. Even with the proposed reduction, the accuracy of the transformation data and the re-transformation data remain highly accurate. The proposed reduction can enhance the manual solution. The first step is to estimate the skip step, S, as shown in Table 2. In reduced data set, \( \{S, 2S, 3S, \ldots, S \times \text{integer}(n/S)\} \) is taken (see Table 3).

In the following discussion, we will apply the data reduction procedure to Numerical Case 2, ensuring that the high normality precision is retained after reduction. Of course, it is possible that Numerical Case 2 does not contain big data and does not call for reduction. Still, this example is only illustrative, and the same procedure applies to the case of big data that calls for a reduction.

The number of data points in Numerical Case 2 is 501, so the S must be 10 (i.e., integer (501/50). Therefore, we are going to start skipping the first nine observations and take the observation number 10. Then we are going to skip the number 11–19 and take the observation number 20, and so on until taking the observation number 500. The reduced data, rank, relative rank, and transformed reduced data are addressed in Appendix II. The distribution of the reduced transformed data is the normal distribution and relatively close to perfection as depicted in Figure 9 and Figure 10.

In the following discussion, we will retrieve original observations from transformed observations and reduced transformed observations and estimate the relative deviation between them, \( D(x) \). First, we calculate the rank through Eq. (9), \( R = \Phi(T) \times (n - 1) + 1 \), then finding the corresponding value to the determined rank using the table (Appendix II). The interpolation is associated almost every time, as the calculated rank is usually a fraction. The relative deviation can be estimated by calculating the original data point using Eqs. (8) and (9) for both the reduced and non-reduced data, calculating the difference between them, and dividing the result by the original data range, as shown in Eq. (9).

### Table 3. Estimating the relative deviation of the reduced original data from the original data.

| Transformed Observation | Original observation (Reduced) | Original observation | D(X) |
|-------------------------|---------------------------------|---------------------|------|
|                         | Retrieved Rank                  | Retrieved Rank      |      |
| -3                      | 1.066                           | 1.674949            | 0.7  |
| 2.5                     | 1.304                           | 4.104833            | 0.701|
| -2                      | 2.115                           | 12.37507            | 0.72375|
| -1.5                    | 4.274                           | 34.4036             | 0.78 |
| -1                      | 8.774                           | 80.32763            | 0.9  |
| -0.5                    | 16.118                          | 155.2688            | 1.11 |
| 0                       | 25.500                          | 251                 | 1.41 |
| 0.5                     | 34.882                          | 346.7312            | 1.967312 |
| 1                       | 42.226                          | 421.6724            | 2.4167|
| 1.5                     | 46.726                          | 467.5964            | 3.591749 |
| 2                       | 48.885                          | 489.6249            | 4.431237 |
| 2.5                     | 49.696                          | 497.8952            | 5.195328 |
| 3                       | 49.934                          | 500.3251            | 6.343324 |

Figure 9. Probability plot of transformed observations in numerical case 2 after reduction.

Figure 10. Histogram of transformed observations in question 2 after reduction.

In the following discussion, we will retrieve original observations from transformed observations and reduced transformed observations and estimate the relative deviation between them, \( D(x) \). First, we calculate the rank through Eq. (9), \( R = \Phi(T) \times (n - 1) + 1 \), then finding the corresponding value to the determined rank using the table (Appendix II). The interpolation is associated almost every time, as the calculated rank is usually a fraction. The relative deviation can be estimated by calculating the original data point using Eqs. (8) and (9) for both the reduced and non-reduced data, calculating the difference between them, and dividing the result by the original data range, as shown in Eq. (9).
The original data (resistivity) and the transformed data. Before creating an individual control plot, estimating the mean, upper control limit (UCL), and lower control limit (LCL) is necessary. Since the transformed data follows the standard normal distribution, the mean is 0, and the standard deviation is 1. Therefore, UCL equals the mean plus three times the standard deviation of the data (i.e., UCL = 3), and LCL equals the mean minus three times the standard deviation of the data (i.e., LCL = -3). The individual control chart was created as in Figure 12 and shows two points outside the chart that belong to the positive infinity point and the negative infinity point. Suppose there is a point out of control other than the positive and negative infinity points. In that case, we must remove the highest or lowest point in the original data, which is the same sign as out of control, and repeat the transformation. By continuing to use this chart in phase II, we skip the non-normality problem.

6. Conclusion

Transforming any non-normal data was successfully conducted through the proposed model. The introduced model can handle any distribution without any limitations and despite data type complexity. The results of transforming data through the proposed model are a highly accurate standard normal distribution. The model is highly recommended to be added to software such as Minitab and SPSS due to the accuracy of its transformation and the simplicity of the theory.

Declarations

Author contribution statement

Mohammad M. Hamasha, Haneen Ali: Conceived and designed the analysis; Contributed analysis tools or data; Wrote the paper.
Sa’d Hamasha, Abdulaziz Ahmed: Contributed analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e09370.

References

[1] M.M. Maghami, M. Bahrani, F.A. Sajadi, On bias reduction estimators of skew-normal and skew-t distributions, J. Appl. Stat. 47 (16) (2020) 3030–3052.
[2] A. Yalçınkaya, B. Şenoğlu, U. Yolcu, Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm, Swarm Evol. Comput. 38 (2020) 127–138.
[3] S. Nadarajah, A generalized normal distribution, J. Appl. Stat. 23 (7) (2006) 685–694.
[4] S. Nadarajah, S. Kotz, A product Pearson-type VII density distribution, J. Comput. Appl. Math. 211 (1) (2008) 103–113.
[5] S. Nadarajah, The product t density distribution arising from the product of two Student’s t PDFs, Stat. Pap. 50 (2009) 605–615.
[6] S. Nadarajah, Some posterior distributions for the normal mean, Int. J. Comput. Math. 88 (6) (2011) 1181–1191.
[7] S. Nadarajah, A.K. Gupta, A posterior distribution for the normal mean arising from a ratio, J. Data Sci. 5 (2015) 441–449.
[8] B.S. Rao, C.S. Kumar, R. Roohi, Variable limits and control charts based on the half normal distribution, J. Test. Eval. 44 (5) (2016) 1878–1884.
[9] V.K. Sharma, Topp–Leon normal distribution with application to increasing failure rate data, J. Stat. Comput. Simulat. 90 (18) (2018) 1782–1803.
[10] K. Ignatieva, Z. Landsman, Conditional tail risk measures for the skewed generalised hyperbolic family, Insur. Math. Econ. 86 (2019) 98–114.
[11] N. Ramli, A.M. Noor, T. Sarmidi, F.F. Said, A.M. Azam, Modelling the volatility of rubber prices in Asuan-3, Int. J. Bus. Soc. 20 (1) (2019) 1–18.
[12] B.L.Q. Alruwaili, Clustering and Cluster Inference of Complex Data Structures, PhD thesis, University of Glasgow, 2019.
[13] R. Al-Athamneh, Reliability Modeling of Microelectronic Interconnections in Long-Term Applications, PhD thesis, Auburn University, 2019.
[14] W. Dong, Event Detection in the Terrain Surface, MS Thesis, Virginia Polytechnic Institute, 2016.
[15] M. Eling, Fitting insurance claims to skewed distributions: are the skew-normal and skew-student good models? Insur.: Insur. Math. Econ. 51 (2) (2012) 239–248.
[16] R. Kaziemi, M.A. Noorizadeh, Comparison between skew-logistic and skew-normal distributions, MATEMATIKA 31 (1) (2015) 15–24.
[17] A. SenGupta, A.K. Laba, A likelihood integrated method for exploratory graphical analysis of change point problem with directional data, Commun. Stat. Theor. M. 57 (11) (2008).
[18] C. Noronha, Data analysis and results. The Theory of Culture-specific Total Quality Management. 89–113.
[19] B. Lenz, The impact of sample non-normality on ANOVA and alternative methods, Br. J. Math. Stat. Psychol. 66 (2) (2013) 224–244.
[20] M.B. Brown, A.B. Foryythe, The ANOVA and multiple comparisons for data with heterogeneous variances, Biometrics 30 (1974) 718–724.
[21] W. H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. J. Am. Stat. Assoc. 47(260), 583.
[22] R. Kazemi, M.A. Noorizadeh, Comparison between skew-logicostic and skew-normal distributions, MATEMATIKA 31 (1) (2015) 15–24.
[23] A. SenGupta, A.K. Laba, A likelihood integrated method for exploratory graphical analysis of change point problem with directional data, Commun. Stat. Theor. M. 57 (11) (2008).
[24] C. Noronha, Data analysis and results. The Theory of Culture-specific Total Quality Management. 89–113.
[25] B. Lenz, The impact of sample non-normality on ANOVA and alternative methods, Br. J. Math. Stat. Psychol. 66 (2) (2013) 224–244.
[26] M.B. Brown, A.B. Foryythe, The ANOVA and multiple comparisons for data with heterogeneous variances, Biometrics 30 (1974) 718–724.
[27] W. H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. J. Am. Stat. Assoc. 47(260), 583.
[28] B. Lantz, The impact of sample non-normality on ANOVA and alternative methods, Br. J. Math. Stat. Psychol. 66 (2) (2013) 224–244.
[29] D.C. Montgomery, The effect of nonnormality on variables sampling plans, Nav. Res. Logist. Q. 32 (1) (1985) 27–33.
[30] J.R. Pace, The Effect of Non-normality on the Cutscore Operating Function: Estimation Correctness in Non-Normal Monte Carlo Simulations, Doctoral dissertation, University of Kansas, 2019.
[31] S. Gao, P.L. Mokhtarian, R.A. Johnston, Nonnormality of data in structural equation models, Transport. Res. Rec. 2082 (1) (2008) 116–124.
[32] R.W.G. Jones, The Effect of Error Non-normality on the Power of Parametric and Non-parametric ANOVA Tests, NAVAL POSTGRADUATE SCHOOL MONTEREY CA, 1971.
[33] M.A. Pasha, M.B. Moghadam, M.A. Rahim, Effects of non-normal quality data on the integrated model of imperfect maintenance, early replacement, and economic design of $$\hat{S}$$ (bar X), $$\hat{S}$$-X control charts, Oper. Res. 20 (4) (2020) 2519–2536.
[34] P. Mishra, C.M. Pandey, U. Singh, A. Gupta, C. Sabu, A. Kesri, Descriptive statistics and normality tests for statistical data, Ann. Card Anaesth. 22 (1) (2019) 67–72.
[35] A. Ghasemi, S. Zahediasl, Normality tests for statistical analysis: a guide for non-statisticians, Int. J. Endocrinol. Metabol. 10 (2) (2012) 486–492.
[36] A. Field, Discovering Statistics Using IBM SPSS Statistics, Sage, 2013.
[37] B. Lantz, The impact of sample non-normality on ANOVA and alternative methods, Br. J. Math. Stat. Psychol. 66 (2) (2013) 224–244.
[38] M.B. Brown, A.B. Foryythe, The ANOVA and multiple comparisons for data with heterogeneous variances, Biometrics 30 (1974) 718–724.
[39] W. H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. J. Am. Stat. Assoc. 47(260), 583–621.
[40] B.L. Welch, On the comparison of several mean values: an alternative approach, Biometrika 38 (1951) 330–336.
[41] A.A. Alexander, D.M. Govern, A new and simpler approximation and ANOVA under variance heterogeneity, J. Educ. Stat. 19 (1994) 91–101.
[42] J.H. Guo, W.M. Luh, Testing methods for the one-way integrated model of imperfect maintenance, early replacement, and economic design of $$\hat{S}$$ (bar X), $$\hat{S}$$-X control charts, Oper. Res. 20 (4) (2020) 2519–2536.
[43] A.C. Elliott, W.A. Woodward, Statistical Analysis Quick Reference Guidebook: with SPSS Examples, Sage, 2007.
[44] L. Fang, Y. Hong, Uncertain revised regression analysis with responses of logarithmic, square root and reciprocal transformations, Soft Comput. 24 (2020) 1069–1079.
[45] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. B 26 (1964) 171–176.
[46] G. Sykes, R. Macarthur, Log transformation of proficiency testing data on the content of genetically modified organisms in food and feed samples: is it justified? Anal. Bioanal. Chem. 412 (2020) 3947–3948.
[47] M.A. Pasha, M.B. Moghadam, M.A. Rahim, Effects of non-normal quality data on the integrated model of imperfect maintenance, early replacement, and economic design of $$\hat{S}$$ (bar X), $$\hat{S}$$-X control charts, Oper. Res. 20 (4) (2020) 2519–2536.
[48] J.H. Guo, W.M. Luh, Testing methods for the one-way integrated model of imperfect maintenance, early replacement, and economic design of $$\hat{S}$$ (bar X), $$\hat{S}$$-X control charts, Oper. Res. 20 (4) (2020) 2519–2536.
[49] A.H. Elian, E. Tüccar, Investigation of four different normality tests in terms of type I error rate and power under different distributions, Turk. J. Med. Sci. 36 (3) (2006) 171–176.
[50] A.C. Elliott, W.A. Woodward, Statistical Analysis Quick Reference Guidebook: with SPSS Examples, Sage, 2007.