JOURNAL OF FUNDAMENTAL MATHEMATICS
AND APPLICATIONS (JFMA) VOL. 8 NO. 1 (2025)
Available online at www.jfma.math.fsm.undip.ac.id

AN ALGORITHM FOR GENERALIZED CONVERSION TO NORMAL DISTRIBUTION FOR INDEPENDENT AND IDENTICALLY DISTRIBUTED RANDOM VARIABLES

Louie Resti S. Rellon

Department of Mathematics, University of Mindanao, Davao City, Philippines
Email : louie_rellon@umindanao.edu.ph

Abstract. The paper analyzes an efficient alternative to the Box-Cox and Johnson's transformation to normality methods which operates under fairly general settings. The method hinges on two results in mathematical statistics: the fact that the cumulative distribution function $F(x)$ of a random variable x always has a $U(0,1)$ distribution and the Box-Mueller transformation of uniform random variables to standard normal random variables. Bounds for the Kolmogorov-Smirnov statistic between the distribution of the transformed observations and the normal distribution are provided by numerical simulation and by appealing to the Dvoretzky-Kiefer- Wolfowitz inequality.
Keywords: transformation to normality, Box-Cox method, Johnson method, inequalities.

I. INTRODUCTION

Most parametric statistical tests of hypotheses in statistical inference rely on the assumption that the data on hand are normally distributed. In fact, for some of these statistical tests, departure from this assumption can lead to serious consequences in terms of either the power of the tests or the level of significance of the tests while others can be quite robust to departures from the normality assumption [5] . Since such statistical tests are often used, it is a good practice to transform the data to one which obeys the normal distribution prior to their use in data analysis.

The most popular method used to transform data to normality is the Box-Cox transformation technique [6]. Thus, if X is any non-negative random variable whose distribution is not normal, then the Box-Cox technique finds an exponent a such that:

$$Y=\frac{X^a-1}{a}, \text{ if } a\neq 0 \text{ or } Y=\log(X) \text{ if } a=0 \tag{1}$$

is normal. If a = 1, then no transformation is needed; if a = - 1, then an inverse transformation is required; if a = ½, a square root transformation may be appropriate. By convention, a = 0 will refer to a logarithmic data transformation. The usual range for the values of a is between -2 to 2 and the process is by trial and error. The trial and error procedure involved in using the family of Box-Cox transformations makes it unpopular in practice.

Despite the limitation, the Box-Cox transformation remains highly effective in normalizing skewed data, especially when simpler transformations like the log transformation fail [11].

Unlike other methods, such as generalized power transformations (GPA), the Box-Cox method virtually normalizes data by selecting the appropriate power transformation. Osborne [14] demonstrated that the Box-Cox transformation outperforms other techniques in achieving normality. Furthermore, the Box-Cox transformation is more general because the log transformation is a special case of the Box-Cox method. It can handle both positive and negative skewness effectively, making it a powerful tool in data analysis [12, 13].

In addition to normalizing skewed data, the Box-Cox transformation is often used in conjunction with weighted least squares (WLS) techniques to address heteroscedasticity, where the variance of the errors is not constant. Agarwal et al. also developed a weighted least squares model to determine the optimal weighting parameter, further enhancing the Box-Cox transformation's effectiveness in dealing with non-constant variance [15].

A more recent addition to the methodologies for transforming data to normality is the Yeo-Johnson transformation [10]. This method generalizes the Box-Cox approach to handle negative random variables. However, like the Box-Cox method, the Yeo-Johnson transformation still faces similar analytic challenges. For instance, when data are obtained from a uniform distribution on [0, 1], the Yeo-Johnson method may fail to find an appropriate transformation to normality, even though established procedures exist for transforming $U(0,1)$random variables to normally distributed ones.

The search for better and more efficient methods for transforming non-normal data continues. This paper proposes a more general approach to data transformation that does not require trial and error and can be easily implemented with today's faster and more efficient computing power. The proposed method is surprisingly simple and is based on the well-known inverse transform theorem in probability.

II. PRELIMINARIES

The uniform distribution on [0, 1] whose density is given by:

$$f(u)=1, \quad 0\le u\le 1 \tag{2}$$

is the basis for generating random numbers from other distribution. Lemma 1 will be used in the proposed procedure later.

Lemma 1 *Let x be a random variable with density f(x) and cumulative distribution function F(x), then it can be easily shown that F(x) is uniformly distributed on [0,1]. Thus, $U=F(x)$ has a U[0, 1] distribution. It follows that $x=F^{-1}(x)$. If we can generate a uniform random number U, then we can always generate a random variable x from a distribution f(x) by simply following this inversion formula.*

Proof: Let x have the cdf $F(x)$. Then, $P(U\le u)=P(F(x)\le u)=P(x\le F^{-1}(u))=u$ which is the cdf of a uniform random variable. It follows that $F(x)$ is uniformly distributed on [0, 1]. Next, if we have two independent uniform random variables U_1 and U_2, then we can

always use the Box-Mueller [1] method to generate two standard normal random variables Z_0 and Z_1. Suppose U_1 and U_2 are independent random variables that are uniformly distributed in the interval (0, 1]. Let

$$Z_0 = R\cos(\theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2)$$

and

$$Z_1 = R\sin(\theta) = \sqrt{-2\ln U_1}\sin(2\pi U_2).$$

Then, Z_0 and Z_1 are independent random variables with a normal distribution of standard deviation of 1. The proof is done by the method of Jacobians and can be seen in standard textbooks in mathematical Statistics (see for example, Hogg and Craig [3]). ⌐

A combination of these two standard results in Mathematical Statistics provides a way of transforming non-normal observations into normally distributed random variables. Roughly, if X is distributed $F(X)$, then we know that $F(X)$ has a $U(0,1)$ distribution. Let Y be independently drawn from $F(.)$ so that $F(Y)$ will also have the uniform distribution on (0,1). Define $Z = g((F(X),F(Y))$ be the Box-Mueller transformation provided above. Our ability to implement this algorithm depends on a large extent on the availability of a closed-form expression for the cumulative distribution function $F(X)$.

$$F(X) = \int_{-\infty}^{\infty} f(x)\,dx \tag{3}$$

Even for well-known probability densities $f(.)$, a closed-form expression (3) may not be easily obtained e.g. normal densities, the family of beta densities and others. In order to circumvent this problem, we assume that we have a sufficiently large number of observations, x_1, x_2, ..., x_n iid $F(.)$, where $F(.)$ is unknown. We estimate $F(x)$ by the empirical distribution function $F_n(x)$ given by:

$$F_n(X) = \frac{1}{n}\sum I(x_i) \tag{4}$$

where $I(x_i)$ is the indicator function. In effect, the empirical distribution function puts a mass of $\frac{1}{n}$ to each of the observations less than or equal to x_i. Each $I(x_i)$ is a Bernoulli random variable with $\rho = F(x)$ so that by the Law of Large Numbers, we know that $F_n(x)$ converges to $F(x)$ in probability. A stronger result was established independently by Glivenko and Cantelli showing that the convergence to $F(x)$ in fact is uniform.

In the theory of probability, the Glivenko-Cantelli theorem, named after Valery Ivanonich Glivenko[4] and Francisco Paolo Cantelli [2], determines the asymptotic behavior of the empirical distribution function as the number of independent and identically distributed observation grows. This uniform convergence of more general empirical measures becomes an important property of the Glivenko-Cantelli classes of functions and sets. The Glivenko-Cantelli classes arise in Vapnik- Chervonenkis theory [9], with

applications to machine learning. Applications can be found in econometrics making use of M-estimators. The Glivenko-Cantelli Theorem states that:

$$Sup|F_n(x)-F(x)| \to 0 \quad as\, n \to \infty \tag{5}$$

Bounds for the approximation have been established in the past, the latest being that of Massart [7]. The more popular bound, however, is the Dvoretzky-Kiefer- Wolfowitz bound. It is an inequality bounds theprobability that the random function $F_n$differs from F by more than a given constant ε > 0 anywhere on the real line. More precisely, there is the one-sided estimate $\Pr\left(\sup_{x\in\mathbb{R}}(F_n(x)-F(x))>\varepsilon\right)\le e^{-2n\varepsilon^2}$ for every $\varepsilon \ge \sqrt{\frac{1}{2n}\ln 2}$, and the two-sided estimate $\Pr\left(\sup_{x\in\mathbb{R}}(F_n(x)-F(x))>\varepsilon\right)\le 2e^{-2n\varepsilon^2}$ for every $\varepsilon>0$. This strengthens the Glivenko-Cantelli theorem by quantifying the rate of convergence as *n* tends to infinity. It also estimates the tail probability of the Kolmogorov-Smirnov statistic. The Dvoretzky-Kiefer-Wolfowitz inequality provides a convenient way for determining the number of observations needed to estimate $F(x)$ to any desired degree of accuracy with probability 1 – α.

$$n \ge \frac{1}{2\varepsilon^2}\log\frac{\alpha}{2} \tag{6}$$

We illustrate the sample size requirements for a 95% confidence with various accuracy levels in Table 1.

Table 1 Sample Size Requirements (N) for Varying Effect Sizes (Epsilon)

epsilon	Alpha	N
0.01	0.05	18444.4
0.02	0.05	4611.1
0.03	0.05	2049.4
0.04	0.05	1152.8
0.05	0.05	737.8
0.06	0.05	512.3
0.07	0.05	376.4
0.08	0.05	288.2
0.09	0.05	227.7
0.10	0.05	184.4

III. PROPOSED PROCEDURE

We formalize the proposed procedure in this section. Let $x_1, x_2, ..., x_n$ be i.i.d. $F(x)$. We assume that $F(x)$ is absolutely continuous with respect to a Lebesgue measure viz. the density function $f(x)$ exists but is not known. Let $F_n(x)$ be the empirical distribution function of the random sample and whose properties are well-established [8]. Assume *n* is large enough so that the maximum difference between parent distribution and the empirical distribution

function is small, say ε. Such a sample size can be obtained from the Dvoretzky-Kiefer-Wolfowitz inequality. Let:

$$g(u_1,u_2)=\sqrt{-2\ln U_1}\cos(2\pi U_2) \text{ and } h(u_1,u_2)=\sqrt{-2\ln U_1}\sin(2\pi U_2) \tag{7}$$

Lemma 2 *Let $U_1 = F_n(x_1)$ and $U_2 = F_n(x_2)$. We claim that g(.) and h(.) are approximately independent standard normal random variables.*

Proof: It suffices to prove that $g(.)$ is a standard normal random variable. If $U_1 = F(x_1)$ and $U_2 = F(x_2)$, then by the previous result of Box-Mueller, the result follows. We replace F by its empirical estimate F_n:

$$U_1=F_n(x_1) \text{ and } U_2=F_n(x_2)$$

We measure the difference between $g(U_1,U_2)$ and $g(U_1,U_2)$.Now,

$$\left|g(U_1,U_2)-g(U_1,U_2)\right|=\left|g(F(X_1),F(X_2))-g(F_n(X_1),F_n(X_2))\right|=$$
$$\left|\sqrt{-2\ln F(X_1)}\cos 2\pi F(X_2)-\sqrt{-2\ln F_n(X_1)}\cos 2\pi F_n(X_2)\right|=$$
$$\left|\sqrt{-2\ln F(X_1)}\cos 2\pi F(X_2)-\sqrt{-2\ln F_n(X_1)}\cos 2\pi F_n(X_2)+\sqrt{-2\ln F(X_1)}\cos 2\pi F(X_2)-\sqrt{-2\ln F_n(X_1)}\cos 2\pi F_n(X_2)\right|\le$$
$$\left|\sqrt{-2\ln F(X_1)}\cos 2\pi F(X_2)-\cos 2\pi F_n(X_2)+\left|\cos 2\pi F_n(X_1)\sqrt{-2\ln F(X_1)}-\sqrt{-2\ln F_n(X_1)}\right|\right|\le$$
$$\left|\sqrt{-2\ln F(X_1)}\right|\left|F(X_2)-F_n(X_2)\right|+\left|\cos 2\pi F_n(X_n)\right|\left|\sqrt{-2\ln F(X_1)}-\sqrt{-2\ln F_n(X_1)}\right|$$
$$\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=1$$

by the Dvoretzky-Kiefer-Wolfowitz inequality or Glivenko-Cantelli theorem. It follows that $g(U_1,U_2)$ is stochastically close to $g(U_1,U_2)$. Since $g(U_1,U_2)$is a standard normal variate by the by the Box-Mueller, it follows that $g(U_1,U_2)$is approximate normal. We now establish the fact that the maximum deviation of the distribution of $g(U_1,U_2)$ from the standard normal distribution $\phi(x)$is bounded by Dvoretzky-Kiefer-Wolfowitz upper bound. Let $F_n\left(g(U_1,U_2)\right)$ be the empirical distribution of $g(U_1,U_2)$. Then:

$$D=\sup\left|F_n\left(g(U_1,U_2)\right)-\phi(g(U_1,U_2))\right|=\sup\left\{\left|F_n\left(g(U_1,U_2)\right)-F(g(U_1,U_2))\right|+\left|F(g(U_1,U_2))-\phi(g(U_1,U_2))\right|\right\}$$
$$\le\sup\left\{\left|F_n\left(g(U_1,U_2)\right)-F(g(U_1,U_2))\right|+\sup\left|F(g(U_1,U_2))-\phi(g(U_1,U_2))\right|\right\}$$
$$\le\varepsilon$$

It follows that

$$P\left(\sup\left|F_n\left(g(U_1,U_2)\right)-\phi(g(U_1,U_2))\right|\ge\varepsilon\right)\ge 2e^{-2n\varepsilon^2}.$$

The larger the sample size n is, the better approximation of the sampling distribution of the statistic $g(.)$ by a standard normal distribution. ⊓

The Yeo-Johnson method transforms an observation X to: $Y = g(x)$, which is normally distributed. In theory, thus, the original observation X can be recovered from $X = g^{-1}(y)$. For the proposal method, the original observation undergoes a series of transformations:

$$Z_1 = B_1\left(F(X), F(Y)\right), \qquad Z_2 = B_2\left(F(X), F(Y)\right) \tag{8}$$

where $F(.)$ is the cdf of X and B_1 is the Box-Mueller transformation involving the cosine function and B_2 is the transformation involving the sine function. Theoretically, we can solve for $F(x)$ and $F(y)$ given Z_1 and Z_2 by simultaneously solving Equation (8). This leads to $U_1 = F(x)$ and $U_2 = F(y)$ for which $x = F^{-1}(U_1)$ and $y = F^{-1}(U_2)$. The problem $x = F^{-1}(U_1)$ and $y = F^{-1}(U_2)$ is that F is unknown and is replaced by $F_n(x)$. Thus, to affect the inversion process, we need an analytic expression for $F_n(x)$ or an estimated analytic expression as $F_n(x) \cong \frac{A}{1+\exp(-Bx)}$, which is the sigmoidal function. We note that this expression is the form needed to implement a logistic regression.

IV. SIMULATION RESULTS AND DISCUSSION

4.1 Simulation Results

We wish to compare the proposed procedure withthe Yeo-Johnson transformation technique using numerical simulations. We simulated 500 observations from the family of beta densities, Gamma densities and Laplace distribution. For each set of observations, we performed both the proposed procedure and the Yeo-Johnson transformation to transform them into normally distributed random numbers. The results of the transformations were compared using the Kolmogorov-Smirnov deviance statistics. The following distributions were used as base distributions for generating the random observations:

Beta : B(1,2), B(1,3), B(2,1), B(2,2), B(2,3), B(3,1), B(3,2), B(3,3)
Gamma :(1,2), G(1,3), G(2,1), G(2,2), G(2,3), G(3,1), G(3,2), G(3,3)
Laplace : L(1,2), L(1,3), L(2,1), L(2,2), L(2,3), L(3,1), L(3,2), L(3,3).

In order to implement the proposed procedure, we followed the Algorithm 4.1.

Algorithm 1

1. Input random data.
2. Arrange random data from smallest to highest.
3. Assign a weight of $\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \ldots, \frac{n-1}{n}, 1$ to the smallest, second lowest, and third lowest up to the highest data respectively.
4. Put the appropriate weights to the original set of unsorted data.
5. Apply Box-Mueller transformation to the weights in step 4.
6. Test the transformed data for normality by the Kolmogorov-Smirnov statistic.

Whenever feasible, we apply the Yeo-Johnson transformation in step 4 to the original data set and compute the Kolmogorov-Smirnov statistic for the transformed data by this method. Table 2 shows the summary of the simulation results.

Table 2 Comparison of the Yeo-Johnson Algorithm and the Proposed Algorithm

	Proposed Algorithm		Yeo-Johnson Algorithm	
Distribution	P-value	Kolmogorov-Smirnov	P-value	Kolmogorov-Smirnov
Beta(1,3)	>0.15	0.021	>0.15	No Transformation
Beta(1,2)	>0.15	0.027	>0.15	No Transformation
Beta(2,3)	>0.15	0.016	>0.15	No Transformation
Beta(2,2)	>0.15	0.013	>0.15	No Transformation
Beta(2,1)	>0.15	0.030	>0.15	No Transformation
Beta(3,1)	>0.15	0.020	>0.15	No Transformation
Beta(3,2)	>0.15	0.017	>0.15	0.019
Beta(3,3)	>0.15	0.026	>0.15	0.024
Gamma(1,3)	>0.15	0.021	>0.15	0.017
Gamma(1,2)	>0.15	0.028	>0.15	0.016
Gamma(2,3)	>0.15	0.031	>0.15	No Transformation
Gamma(2,2)	>0.15	0.021	>0.15	No Transformation
Gamma(2,1)	>0.15	0.022	>0.15	No Transformation
Gamma(3,1)	>0.15	0.020	>0.15	No Transformation
Gamma(3,2)	>0.15	0.024	>0.15	No Transformation
Gamma(3,3)	>0.15	0.028	>0.15	No Transformation
Laplace(1,3)	>0.15	0.029	>0.15	0.020
Laplace(1,2)	>0.15	0.032	>0.15	0.029
Laplace(2,3)	>0.15	0.026	>0.15	0.025
Laplace(2,2)	>0.15	0.032	>0.15	0.031
Laplace(2,1)	>0.15	0.026	>0.15	0.025
Laplace(3,1)	>0.15	0.019	>0.15	0.015
Laplace(3,2)	>0.15	0.023	>0.15	0.031
Laplace(3,3)	>0.15	0.029	>0.15	0.033

The null hypothesis that the distribution of the transformed data is normal is accepted in all cases for the proposed method. The same observation holds true for the Yeo-Johnson algorithm whenever a transformation is available. Whenever a Yeo-Johnson transformation is available, the computed Kolmogorov statistic or maximum deviation statistic tended to be lower for it than the proposed method. However, the differences observed for the Kolmogorov statistical distances between the proposed method and Yeo-Johnson method are very small indeed showing that the two methods provide equally reliable results.

The main advantage of the proposed method over the Yeo-Johnson algorithm (and Box-Cox Method) is that transformations to normally be always possible for the proposed method while the same way not be available for the Yeo-Johnson algorithm. The disadvantage, however, is the fact that the proposed method requires large number of observations ($n > 100$) for it to work efficiently. The proposed method can easily be coded and incorporated in statistical packages.

4.2 Implications of the Method

The algorithm for generalized conversion to a normal distribution has significant applications across various scientific fields where independent and identically distributed (i.i.d.) random variables are considered. In healthcare and biostatistics, transforming skewed biological measurements, such as biomarkers or survival times, ensures accurate modeling, enabling robust statistical inference and better predictive modeling in clinical trials and epidemiological studies [20, 21]. In economics and finance, where data often exhibit heavy tails and skewness, normalization reduces bias in portfolio optimization, risk modeling, and econometric forecasting [22, 23]. Similarly, environmental science benefits from the transformation of skewed data like pollutant levels or meteorological observations, facilitating effective modeling, trend analysis, and regulatory compliance assessments [17]. In engineering and manufacturing, the method enhances quality control by normalizing defect rates and measurement deviations, improving process control and reliability assessment [24, 25]. Furthermore, social sciences, with their reliance on surveys and psychometric data, leverage this algorithm to improve the accuracy of regression models and structural equation modeling [26, 27].

4.3 Limitations of the Algorithm in Non-IID Cases

When independent and identically distributed (i.i.d.) random variables are involved, the generalized conversion to a normal distribution procedure performs admirably; however, when these conditions are not satisfied, it is less useful. A significant disadvantage of data dependency is that many real-world datasets exhibit correlations rather than independence. For example, in geostatistics or environmental modeling [17], time-series data often exhibit autocorrelation [16], where current values depend on previous observations, and spatial data may exhibit spatial dependence, which defies the independence assumption. Additionally, the algorithm's assumption that all data points have homogeneous distributions might not hold true when data comes from mixtures of distributions, such as combining datasets from different populations or experiments [18], or when distributional shifts occur over time or among subgroups [19].

V. CONCLUSION

The generalized algorithm for converting data to a normal distribution represents a transformative step in addressing the challenges posed by non-normality in statistical analysis. Its capacity to transform diverse datasets into a standard normal distribution, grounded in solid theoretical principles and validated through empirical testing, makes it a robust alternative to established methods such as the Box-Cox and Yeo-Johnson transformations. This versatility

allows for its application across a wide range of fields, including healthcare, finance, environmental science, and engineering, where reliable data normalization is critical.

Despite the algorithm's global applicability and consistent results, its limits in non-i.i.d. and dependence on high sample numbers are its drawbacks. Datasets indicate areas that want improvement. For further developments, the novel combination of the empirical distribution function and the Box-Mueller transformation offers a starting point. In order to enhance data preprocessing and feature engineering for skewed datasets, this study could be extended to high-dimensional domains like machine learning or genomics and image processing, where relationships between variables are common. Potential uses in dynamic systems, like engineering predictive maintenance, adaptive environmental monitoring, and real-time financial modeling, could also be looked into.

It is recommended that future academics expand on this work by changing the algorithm to handle non-i.i.d. data settings, improving its scalability for datasets with high dimensions, or incorporating it into statistical software tools that are automated. Researchers can further unleash the algorithm's potential by expanding and improving this study, opening the door for more creative and extensive applications in both established and developing scientific fields.

Acknowledgements

The authors are grateful to the reviewers for carefully reading the manuscript and for offering substantial comments and suggestions which enabled them to improve the presentation. This work was funded by the authors.

REFERENSI

[1] Box, G. E. P.; Muller, Mervin E. (1958). "A Note on the Generation of Random Normal Deviates". *The Annals of Mathematical Statistics*. 29 (2): 610–611. doi:10.1214/aoms/1177706645. JSTOR 2237361.

[2] Cantelli, F. P. (1933). Sulla determinazione empirica della legge di probabilita. *Giorn.lst. Ital.Attuari* 4, 221-424

[3] Craig, W and Hogg, *An Introduction to Mathematical Statistics*, (Willey and Sons, New York, 2000)

[4] Glivenko, V. (1933). Sulla determinazione empirica della legge di probabilita. *Giorn.lst. Ital.Attuari* 4, 92-99

[5] Huber, P. (1981) *Robust Statistics*. Wiley, New York.

[6] Johnson, R and Wichern, *Applied Multivariate Statistical Analysis* (Willey and Sons, New York, 2000)

[7] Massart, P.(1990) "The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality." *Ann. Probab*. 18 (3) 1269 - 1283. https://doi.org/10.1214/aop/1176990746

[8] Shorack, G. R., Wellner, J. A. (1986) *Empirical Processes with Applications to Statistics*, Wiley.ISBN 0-471-86725-X.

[9] Vapnik, V.N. and Chervonenkis, A. Ya (1971). On uniform convergence of the frequencies of events to their probabilities. *Theor. Prob. Appl*. 16, 264-280

[10] Yeo, In-Kwon and Johnson, Richard (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, 87, 954-959.

[11] Pan, P., Li, R., & Zhang, Y. (2023). Predicting punching shear in RC interior flat slabs with steel and FRP reinforcements using Box-Cox and Yeo-Johnson transformations. *Case Studies in Construction Materials*, 19, e02409. https://doi.org/10.1016/j.cscm.2023.e02409

[12] Thomas K, Peter JV, Christina J, et al. Cost-utility in medical intensive care patients. Rationalizing ongoing care and timing of discharge from intensive care. *Ann Am Thorac Soc.* 2015 Jul;12(7):1058–1065.

[13] Peter JV, Thomas K, Jeyaseelan L, et al. Cost OF intensive care IN India. *Int J Technol Assess Health Care.* 2016 Jan;32(4):241–245.

[14] Osborne J. *Improving your data transformations: applying the Box-Cox transformation. Practical Assessment. Res Eval* [Internet]. 2019 Nov 23;15(1). Available from https://scholarworks.umass.edu/pare/vol15/iss1/12.

[15] Agarwal GG, Pant R. Regression model with power transformation weighting: application to peak expiratory flow rate. *J Reliabil Stat Stud*. 2009 Aug 10:52–59.

[16] Dickey, D. A., & Chatfield, C. (1991). The Analysis of Time Series: An Introduction. *Technometrics*, *33*(3), 363. https://doi.org/10.2307/1268794

[17] Cressie, N. (1992). STATISTICS FOR SPATIAL DATA. *Terra Nova*, *4*(5), 613–617. https://doi.org/10.1111/j.1365-3121.1992.tb00605.x

[18] McLachlan, G., & Peel, D. (2000). *Finite mixture models*. http://ci.nii.ac.jp/ncid/BA49548052

[19] Sugiyama, M., & Kawanabe, M. (2012). *Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation*. http://ci.nii.ac.jp/ncid/BB08941717

[20] Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied Survival Analysis: Regression modeling of Time-to-Event data*. http://cds.cern.ch/record/1555255

[21] Walter, S. D., & Altman, D. G. (1992). Practical Statistics for Medical Research. *Biometrics*, *48*(2), 656. https://doi.org/10.2307/2532320

[22] Campbell, J. Y., Lo, A. W., & MacKinlay, A. (2012). *The econometrics of financial markets*. https://doi.org/10.2307/j.ctt7skm5

[23] Tsay, R. S. (2006). Analysis of Financial Time Series. *Technometrics*, *48*(2), 316. https://doi.org/10.1198/tech.2006.s405

[24] Sauter, R. M., & Montgomery, D. C. (1992). Introduction to statistical quality control. *Technometrics*, *34*(2), 232. https://doi.org/10.2307/1269251

[25] Menten, T., & Phadke, M. (1991). Quality Engineering Using Robust Design. *Technometrics*, *33*(2), 236. https://doi.org/10.2307/1269049

[26] Kline, R. B. (1998). *Principles and practice of structural equation modeling*. http://cds.cern.ch/record/2123621

[27] Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics*. https://www.goodreads.com/work/editions/1559750-using-multivariate-statistics