ISSN 1533-9211

DOI: 10.5281/zenodo.8285749

ON SELECTING POWER TRANSFORMATION PARAMETERS WITH THE PRESENCE OF AN OUTLIER

AZAD ADIL SHAREEF *

Department of Statistics, College of Administration and Economics, University of Duhok, Kurdistan Region, Iraq. Corresponding Author Email: *azada@uod.ac

HAITHEM TAHA MOHAMMED ALI

Department of Economic Sciences, University of Zakho, Department of Economics, Nawroz University, Kurdistan Region, Iraq. Email: haithem.taha@uoz.edu.krd

Abstract

This study assesses the new approach of the Box-Cox Transformation to estimate power parameters using five criteria: the traditional Maximum Likelihood Estimation; coefficient of determination; p-value of Shapiro-Wilk test statistics for the residual's normality of the estimated linear regression of the transformed response vector; p-value related to residual's normality; and the Mean Square Errors of the estimated nonlinear regression of the original response vector. The efficiency of these criteria is studied to determine the optimal transformation parameter in the presence of an outlier within a response variable in simple linear regression. The computational algorithm has been developed and applied to medical data. The authors concluded that it is difficult to obtain a feasible solution for all criteria from which an optimal power parameter can be selected. Therefore, the researcher's experience can be considered a decisive factor in choosing according to the priorities of the comparison between the criteria.

Keywords: Simple linear Regression, Box-Cox Transformation, Maximum Likelihood Estimation, and Outliers

1. INTRODUCTION

Power Transformation (PT) models can offer solutions to problems resulting from a lack of conditions for statistical modeling. In data analysis, transformation is the replacement of a variable by a function of that variable; for example, replacing a variable x by the square root of x or the logarithm of x. In other words, transformation is a shifting process that changes the shape of distribution of the relationship. Regarding the transformation of the response variable in non-linear relationships, PT is a tool to obtain a linear model for the transformed data as a first step. Subsequently, through the back transformation of the PT model, we can re-represent the original data with an efficient nonlinear model estimator. The current study used different methods and criteria to estimate the power parameter of Box-Cox transformation (BCT) in simple linear regression (SLR) model with the presence of an outlier value in response variable dataset. According to Hadi and Chatterjee, ordinary residuals are not appropriate for diagnostic purposes; a transformed version of these is preferable [1].

The Box-Cox transformation (BCT) approach focuses on satisfying the modelling conditions in the Multiple Linear Regression model by using parametric PT [2]. Poirier used the maximum likelihood method to estimate the power parameter of the transformed response model in some life models when the residual is distributed according to the truncated normal distribution [3]. Abd-Rahman and Gerig used the maximum likelihood method to estimate the parameter of PT

The Seybold REPORT
ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

in the general linear model when the linear model had a variance that was proportional to the response by the amount of the PT parameter [4]. Cook and Weisberg proposed their method in 1994. They addressed the problem of selecting a transformation $\varphi(y)$ of a univariate response variable Y so that the regression function $E(\varphi(Y)|X = x)$ was linear in the predictor vector $\boldsymbol{x}$, which aims to find a linear and monotonic transformation of the response variable using the BCT model [5]. Osborne indicated that a method of exploring the outliers with a significantly small proportion of outliers affecting even simple analysis is by summarizing the various potential causes of extreme scores within a data set [6]. Vélez, Correa, and Marmolejo-Ramos proposed a new methodology for estimating BCT parameters, as well as an alternative method for determining plausible values for it. The former is accomplished by first defining a grid of values for and then, running a normality test on the transformed data. The optimal value of λ, is one with the highest p-value from the normality test. After plotting the p-values against the values of λ on the grid, the set of plausible values is determined using the inverse probability method [7]. Pek and Wong applied their research on data transformation to infer with linear regression to achieve the assumption of normally distributed errors in the population [8]. Atkinson and Corbellini introduced a method of BCT family for non-negative response linear models and presented a long and interesting history in both theory and practice that was linked to generalized linear models and log transformed data in terms of the transformation of both sides model. [9]. In 2021, Atkinson, Riani, and Corbellini, studied the BCT of non-negative response in linear regression models. The extensions include both positive or negative of the model transformation and Yeo-Johnson transformation for observations that can be positive or negative [10]. The risk of the methods proposed by Riani, Atkinson, and Corbellini comes from using robust analytics for transforming data and introducing an automatic procedure for transforming the response in regression models to approximate normality. BCT is discussed here, as well as its generalization to the extended Yeo-Johnson transformation, which allows for both positive and negative responses [11].

The aim of the present study is to use PT in the SLR model when the data set contains an outlier. Therefore, it is concerned with using several criteria to select the optimal power parameter and compare its efficiency. The rest of the article is recognized as follows: Section two includes some theoretical aspects on PT and outliers. Section three includes the computational algorithm of the use of BCT model in SLR with the five criteria. Sections four and five include the application and conclusions, respectively.

2. MATHEMATICAL APPROACH OF PARAMETRIC PT AND OUTLIERS

The basic assumption in BCT methodology is that the transformed data is distributed according to the normal distribution. Consequently, the original data will have an unknown probability density function (PDF) and its parameter space will include the scale and location parameters of the normal distribution as well as the transformation parameter. For the positive variable Y, the BCT model is defined as follows [2]:

$$\varphi(y) = \begin{cases} y^{\lambda} - 1/\lambda & if\ \lambda \neq 0 \\ log(y) & if\ \lambda = 0 \end{cases} \tag{1}$$

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

When $\lambda = 1$, the variable is analyzed in its original scale; when $\lambda = 0$ corresponds to the logarithmic transformation. The back transformation of BCT is defined as the following nonlinear model:

$$Y = \begin{cases} (\lambda(\varphi(y)) + 1)^{1/\lambda} & if\ \lambda \neq 0 \\ exp(\varphi(y)) & if\ \lambda = 0 \end{cases} \quad (2)$$

If we suppose that $\varphi(\mathrm{y}) \sim N(\mu, \sigma^2)$, then the PDF of the original random variable Y is of the form:

$$f_Y(y;\ \lambda, \mu, \sigma^2) = f_{\varphi(y)}(\varphi(y);\ \lambda, \mu, \sigma^2)|J| \quad (3)$$

where J is the Jacobian factor. Eq. 3 can be written in the following form:

$$f_Y(y;\ \lambda, \mu, \sigma^2) = \begin{cases} (2\pi\sigma^2)^{-1/2}\ exp\{-(\varphi(y) - \mu)^2/2\sigma^2\}y^{\lambda-1} & if\ \lambda \neq 0 \\ (2\pi\sigma^2)^{-1/2}\ exp\{-(\varphi(y) - \mu)^2/2\sigma^2\}y^{-1} & if\ \lambda = 0 \end{cases} \quad (4)$$

Now, for the SLR model:

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad (5)$$

where Y is the response variable, X is the explanatory variable, β_0 and β_1 represent the parameters to be estimated and the error term ε. For the nonlinear data, the linear regression analysis according to Eq. 5 will be inappropriate for obtaining an efficient model. Therefore, the data can be transformed using PT to estimate a linear model for the transformed data and benefit from the power parameter to estimate a non-linear regression model for the original data.

If we suppose that $\varphi(y) = \beta_0 + \beta_1 X + \varepsilon$ represents the linear model of the transformed data so that $\varphi(y) \sim N(\mu_x, \sigma_\varepsilon^2)$, where, $\mu_x = \beta_0 + \beta_1 X$, and $\mathrm{var}(\varphi(y)) = \sigma_\varepsilon^2$, the Likelihood to estimate the power parameter is as follows [2]:

$$\mathrm{L}(y; \lambda, \beta_0, \beta_1, \sigma_\varepsilon^2) = \prod_{i=1}^{N} f_Y(\varphi(y);\ \lambda, \beta_0 + \beta_1 X, \sigma_\varepsilon^2)|d\varphi(y)/dy|$$

$$= (2\pi\sigma_\varepsilon^2)^{-\mathrm{N}/2}\ exp\{-\textstyle\sum_{i=1}^{N}(\varphi(y_i) - (\beta_0 + \beta_1 X_i))^2 / 2\sigma_\varepsilon^2\} \textstyle\sum_{i=1}^{N}|d\varphi(y_i)/dy_i| \quad (6)$$

Then, the log likelihood is in the following form:

$$\log \mathrm{L}(y; \lambda, \beta_0, \beta_1, \sigma_\varepsilon^2) = -\mathrm{N}/2 \log 2\pi - \mathrm{N}/2 \log \sigma_\varepsilon^2 - \sum \varepsilon^2/2\sigma_\varepsilon^2 + \sum_{i=1}^{N} \log\ |d\varphi(y_i)/dy_i| \quad (7)$$

where:

$$\sigma_\varepsilon^2 = \sum {\varepsilon^2}_i/N \quad (8)$$

Thus, the optimal λ is the value that maximize the variable parameter of equation (7), such that

$$\log L(y; \lambda, \beta_0, \beta_1, \sigma_\varepsilon^2) = A - \mathrm{N}/2 \log \sigma_\varepsilon^2 + \textstyle\sum_{i=1}^{N} \log\ |d\varphi(y_i)/dy_i|, \quad (9)$$

where A is the constant term of equation (9) and equal to $-\mathrm{N}/2\log 2\pi - \mathrm{N}/2$.

In regression analysis, an outlier is an observation for which the residual is large in magnitude

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

compared with other observations in the data set. Hinkley indicated that PT in the presence of outliers is a risky business [12]. With the medical data set, the effect of outliers on the least squares capabilities explains how a single observation may alternate the direction of least square [13]. This illustrates that the presence of an outliers of the regression for these observations would have large scalar residual compared with the other observations.

Qin and Yang proposed a new outlier detection method based on the wavelet transform and local outlier factor algorithm; the experimental results on transformed data from several power transformers showed that the algorithm can detect outliers that exceed the threshold value [14]. In addition, Raymaekers noted that MLE is highly sensitive in estimating parameters when the outliers are presented [15].

3. COMPUTATIONAL ALGORITHM

In this study, the authors used different methods and criteria to estimate the parameter. In addition to the traditional MLE method [15], other test criteria are the Coefficient of Determination (COD) and p-value of Shapiro-Wilk (SW) test statistics of the residual's normality for the following estimated linear regression of the transformed response vector.

$$\widehat{\varphi}(\mathbf{y}) = \widehat{\beta}_0 + \widehat{\beta}_1 x \qquad (10)$$

From the estimated nonlinear regression of the original response resulting from the inverse of BCT according to Eq 2, the tested criteria are the Mean Square Errors (MSE) and p-value of SW test statistics of the residual's normality.

The steps of the computational algorithm are as follows:

Step 1: Suppose that $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda} = \{-\mathbf{2}, -\mathbf{1.9}, \dots, \mathbf{1.9}, \mathbf{2}\}$. This range can be enlarged to $\boldsymbol{\Lambda} = \{-\mathbf{5}, -\mathbf{4.9}, \dots, \mathbf{2.9}, \mathbf{3}\}$ when we do not get a maximum value for MLE (Eq. 7) and COD of the SLR of the transformed data; in other words, when the curves of these indicators are not convex.

Step 2: Estimate SLR model of the transformed response, $\widehat{\boldsymbol{\varphi}}(\mathbf{y}) = \widehat{\boldsymbol{\beta}}_0 + \widehat{\boldsymbol{\beta}}_1 \boldsymbol{x}$ and estimate using the first criteria; COD.

Step 3: Test the normality of the random residual and model for transformed data then calculate the value of p-value.

Step 4: Estimate using the MLE according to Eq. 7.

Step 5: Estimate the nonlinear regression model of the original response Y by the back transformation of BCT according to Eq. 2.

$$\widehat{y} = \begin{cases} \left(\lambda(\widehat{\beta}_0 + \widehat{\beta}_1 x) + 1\right)^{1/\lambda} & \text{if } \lambda \neq 0 \\ \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x) & \text{if } \lambda = 0 \end{cases} \qquad (11)$$

Step 6: Estimate Mean Square Error for estimated nonlinear regression model according to Eq. 11,

The Seybold REPORT
ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

$$MSE(\hat{y}/\lambda) = \sum (y_i - \hat{y}_i)^2 / N - 2 \qquad (12)$$

Step 7: Test the normality of random residual of the estimated nonlinear regression model (Eq. 11) using the SW test.

Step 8: Repeat steps 1 to 7 for all $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$.

4. APPLICATION

BCT was applied to medical dataset and the R software was used to analyze the data. The medical dataset was collected randomly for male and female at the Azadi heart center at the Duhok hospital in the Duhok Governorate in Kurdistan Region of Iraq and contained 30 observations that included a dependent variable Y representing the Systolic Blood Pressure (SBP) and an independent variable (X) as Age (Table 1).

Table 1: Medical dataset of 30 observations between Age (X) and the SBP (Y)

X	39	47	45	47	65	46	67	42	67	56
Y	144	220	138	145	162	142	170	124	158	154
X	64	56	59	34	42	48	45	17	20	19
Y	162	150	140	110	128	130	135	114	116	124
X	36	50	39	21	44	53	63	29	25	69
Y	136	142	120	120	160	158	144	130	125	175

In Figure 1, it can be seen that the second observation of the response variable of SBP is an outlier. Various criteria, such as the Studentized, Mahalanobis Distance, Box Plot, Cooks Distance, and Welsch Distance are used to identify outliers [1] (see figure 1).

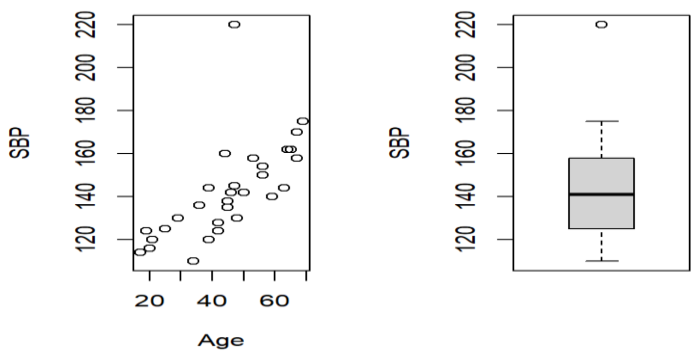


Figure 1: Scatter plot and box-plot of data set

In Table 2, to note the estimation of the power parameter according to the five criteria by applying the proposed algorithm, we can choose that for the power parameter in the range$(\mathbf{-5, -0.1})$, the slope for all intervals is 0.00. Therefore, the estimated model for y according to Eq. 10 will be a straight line parallel to the x-axis. In addition, the logarithm transformation also shows that the slope is 0.00, estimated. However, when the power

ISSN 1533-9211

DOI: 10.5281/zenodo.8285749

parameter was $\boldsymbol{\lambda} \, \boldsymbol{\epsilon} \, (\mathbf{0.1}, \mathbf{0.9})$, we can see that the slope is between (0.00 and 0.06), which is approximately 0, although the p-value of the SW test statistics for the residual's normality of the estimated linear regression for the transformed response vector is 0.00; this means that the residual is not normally distributed. For the values of power parameter in the range$(\mathbf{1}, \mathbf{3})$, it can be shown that the slope appeared between (0.97 and 2E+04), although the p-value of SW test statistics for the error's normality of both estimated linear regression and non-regression are approximately 0.00. Regarding the COD criteria, it is known that the curve $(\boldsymbol{\lambda}, \mathbf{R^2})$ becomes convex. And we restored to expanding the range from (-2,2) to (-5,3) until obtaining the highest value for COD, when the values of the power parameter are in the range $(\mathbf{-5}, \mathbf{-2.5})$, the COD is increased from 0.60 to 0.61 but the slope is 0.00 (Figure 2.b). Based on the MLE method, we obtained the convex curve (Figure 2.a) of function by using Eq. 7; the optimum value of this curve corresponds to the maximum value of the power parameter. Regarding the MSE criterion, it shows that the concave curve (Figure 2.c) and optimum value of this curve corresponding to the minimum value of the power parameter is 1.

Table 2: Results of applying the computational algorithm with the presence of the outlier

$\boldsymbol{\lambda}$	$\mathbf{MLE}(\hat{y}\vert\boldsymbol{\lambda})$	SLR of Transformed Response $\hat{\varphi}(y)\vert X$				Nonlinear Regression $\hat{y}\vert X$	
		$\mathbf{R^2}$	P value ($\boldsymbol{\varepsilon}$)	Intercept	Slope	P value ($\boldsymbol{\varepsilon}$)	$\mathbf{MSE}(\hat{y}\vert\boldsymbol{\lambda})$
(-5.0, -4.4)	(17.5, 19.1)	(0.60, 0.61)	(0.07, 0.12)	(0.20, 0.23)	0.00	(0.10, 0.12)	(337, 351)
(-4.3, -2.5)	(19.3, 21.4)	0.61	(0.09. 0.18)	(0.23, 0.40)	0.00	(0.12, 0.82)	(315, 336)
(-2.4, -1.7)	(21.0, 21.4)	(0.59, 0.60)	(0.01, 0.07)	(0.44, 0.59)	0.00	(0.89, 0. 97)	(314, 314)
(-1.6, -0.1)	(17.4, 21.0)	(0.51, 0.58)	(0.00, 0.01)	(0.63, 3.71)	0.00	(0.38, 0.97)	(301, 309)
Ln Y	17.1	0.51	0.00	4.64	0.00	0.30	301
(0.1, 0.9)	(15.0, 17.1)	(0.44, 0.50)	0.00	(5.9, 68.52)	(0.00, 0.06)	(0.02, 0.24)	(300, 301)
1.0	14.8	0.43	0.00	97.72	0.97	0.01	300
(1.1, 3.0)	(1.1, 14.2)	(0.26, 0.42)	0.00	(140.26, 8E+05)	(1.6, 2E+04)	(0.00, 0.01)	(300, 308)

Table 3: Results according to the optimal λ

Criteria	Optimal λ	Intercept	Slope	Estimated Nonlinear *Model* $\hat{y}\vert X$
Max $(R^2(\varphi(y)\vert X)$	$(-4.3, -2.5)$	(0.23, 0.40)	0.00	$\hat{y} =$ constant
Max (P value $(\varepsilon\vert\hat{y})$	-1.6	0.63	0.00	$\hat{y} =$ constant
Min $(MSE(\hat{y}\vert\lambda))$	1.0	97.72	0.97	$\hat{y} = 98.72 + 0.97\,x$
Max (P value $(\varepsilon\vert\hat{\varphi}(y))$	-3.6	0.278	0.00	$\hat{y} =$ constant
Max (MLE)	-2.4	0.417	0.00	$\hat{y} =$ constant

Table 3 shows that the optimal λ according to the five criterions with the presence of the outlier.

ISSN 1533-9211

DOI: 10.5281/zenodo.8285749

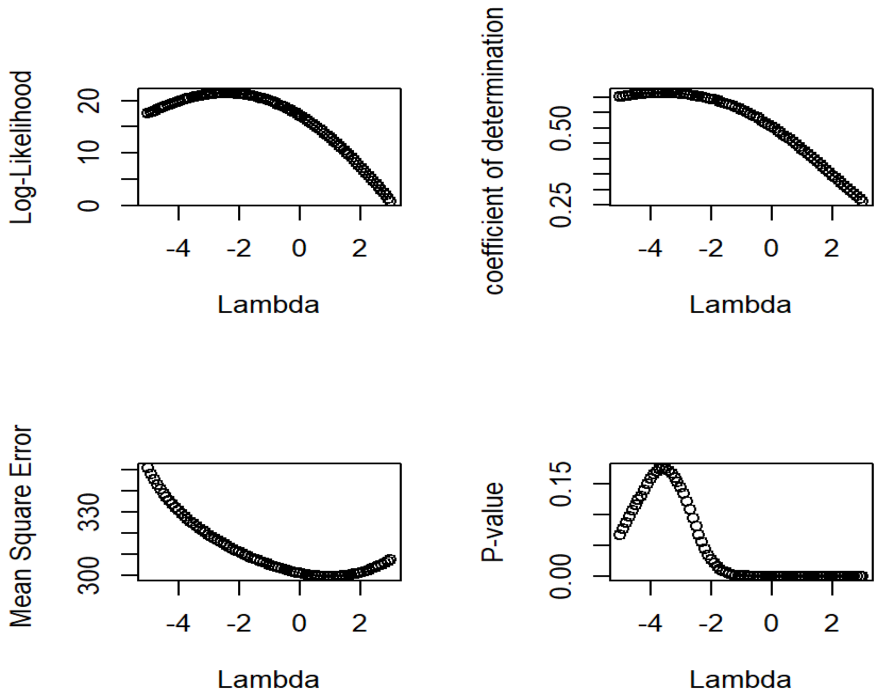


Figure 2: Plots; (a) log-likelihood, (b) COD, (c) MSE plot of the estimated SLR of transformed response variable Z according to BCT, and (d) p-value of SLR of Transformed Response

Table 4 shows the estimations of the optimal power parameter according to the five criteria after deleting the outlier value which is necessary to delete it. The COD for the dataset after removing the outlier is 0.71 when the power parameter$\lambda = 1$, which increases if we compare it with the original dataset and p-value of the SW test statistics for the residual's normality. Both estimated linear regression and non-linear regression are 0.37, indicating that the residual is normally distributed. Conversely, the transformation $Ln(y)$ shows that the p-value of SW test statistics for the residual's normality for both estimated linear regression and non-linear regression are increased by 0.43 and 0.47, respectively. That means that the random residuals are closer to the normal distribution. Another criterion is improved when the power parameter becomes $Ln(y)$, which is MSE. A feasible solution is for the optimal value of the power parameter to be deduced in the *Ln(y).* If L_{max} is the MLE value of PDF of the original random variable Y after deleting the outlier according to Eq. 7, representing the basis for estimating the optimum power parameter, then the optimal value is *Ln(y),* which is shown in Table 5. Thus, we can conclude that the convex curve of the MLE function, p-value of the SW test statistics for the residual's normality by using Eq. 7, and maximum value of this curve correspond to the optimal value of the power parameter (Figure 3). Finally, we found that selecting the optimal value for the PT was possible based on three criteria: the p-value of SW test statistics for the residual's normality of the estimated linear regression of the transformed response vector, MSE

ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

of the estimated nonlinear regression of the original response vector resulting from the inverse of BCT, and MLE. Thus, we obtained the following nonlinear model as suitable for the data:

$$\hat{y} = \exp\{4.62 + 0.01x\} \qquad (13)$$

Resulting from the following equation

$$ln(\hat{y}) = 4.62 + 0.01x \qquad (14)$$

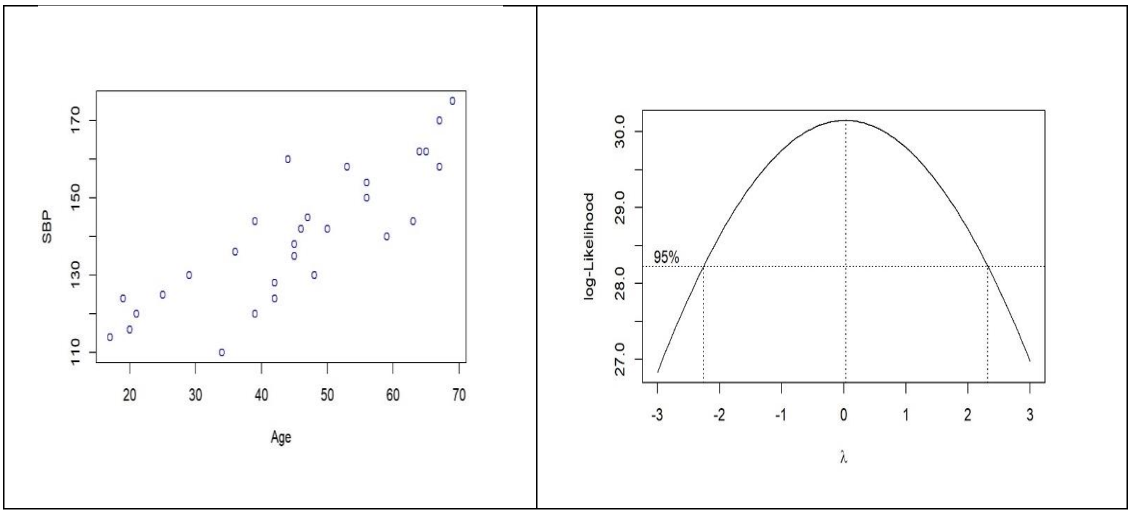
See Figure 4 the plot of observed and fitted value when λ = ln(y).

Table 4: Results of applying the computational algorithm after deleting the outlier

λ	**MLE($\hat{y}$|λ)**	**SLR of Transformed Response $\hat{\varphi}(y)\|X$**				**Nonlinear Regression $\hat{y}\|X$**	
		R^2	**P value (ε)**	**Intercept**	**Slope**	**P value (ε)**	**MSE($\hat{y}$|λ)**
(-3.0, -2.1)	(26.8,28.4)	(0.68, 0.70)	(0.12, 0.28)	(0.33, 0.47)	0.00	(0.51, 0.58)	(83.5,82.9)
(-2.0, -1.3)	(28.6, 29.5)	(0.70, 0.71)	(0.30, 0.39)	(0.50, 0.76)	0.00	(0.59, 0.61)	(82.8, 83.2)
(-1.2, -0.1)	(29.6, 30.1)	(0.71, 0.72)	(0.40, 0.43)	(0.83, 3.70)	0.00	(0.47, 0.62)	(83.2, 84.9)
Ln y	30.2	0.72	0.43	4.62	0.01	0.46	85.1
(0.1, 0.9)	(29.8, 30.1)	(0.71, 0.72)	(0.39, 0.42)	(5.9, 67.5)	(0.01, 0.57)	(0.39, 0.46)	(85.4, 87.8)
1.0	29.79	0.71	0.37	96.07	0.94	0.37	88.1
(1.1,2.0)	(28.7, 29.7)	(0.70, 0.71)	(0.31, 0.37)	(137.1, 3901)	(1.5, 133.7)	(0.32, 0.36)	(88.5, 93.1)
(2.1, 3)	(26.9, 28.6)	(0.68, 0.70)	(0.20, 0.30)	(5641, 96093)	(219.5, 9028)	(0.32, 0.36)	(93.7, 101.1)

Table 5: Results according to the optimal λ after deleting the outlier

Criteria	**Optimal λ**	**Intercept**	**Slope**	**Estimated Nonlinear Model $\hat{y}\|X$**
Max ($R^2(\varphi(y)\|X)$)	(-1.2, 0.9)	(0.83, 67.5)	(0.00, 0.57)	
Max (P value ($\varepsilon\|\hat{y}$))	Ln (y)	4.62	0.01	$\hat{y} = \exp\{4.62 + 0.01xi\}$
Min ($MSE(\hat{y}\|\lambda)$)	Ln (y)	4.62	0.01	$\hat{y} = \exp\{4.62 + 0.01xi\}$
Max (P value ($\varepsilon\|\hat{\varphi}(y)$))	(-3, 0.1)	(0.33, 4.62)	(0.00, 0.01)	
Max (MLE)	Ln (y)	4.62	0.01	$\hat{y} = \exp\{4.62 + 0.01xi\}$



ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

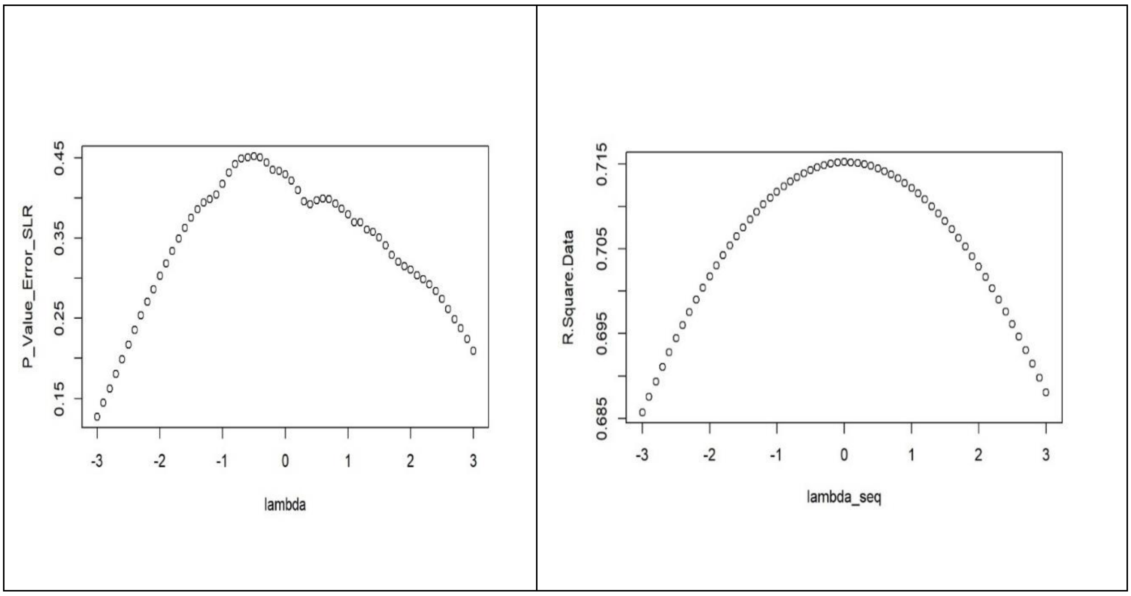


Figure 3: Scatter plot for data set (a) and plots; log-likelihood (b), p-value of SLR of transformed response (c) and COD (d)

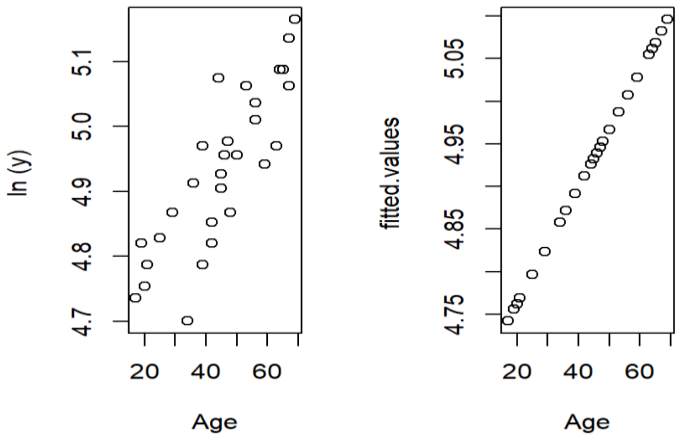


Figure 4: Plot of observed and fitted value

We generated five different criteria for choosing the best value of the transformation parameter by using BCT methods. The authors believe that it is not feasible to acquire an optimal value that meets the conditions of the five criteria: the maximum value of the MLE function, COD, p-value of the SW test of residual vector normality, p-value of SW test for the residual's normality, and MSE of the estimated nonlinear regression model of the original response vector. In several cases, we must evaluate the findings based on the importance and priority of some criteria, as well as what additional criteria might contribute to support the priorities.

ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

5. CONCLUSIONS

This study explored the identification of outliers in a SLR model. Notably, the optimal power parameters for transformation models by using BCT is significantly effective. There are a variety of methods for choosing the best power parameter, which are classified such that the first is the use of well-known estimation methods, such as the MLE method. The second method is to use some efficiency criteria from regression modelling that has been shown in this study, such as the COD and p-value, as decision rules when estimating the power parameter. Furthermore, one can use the p-value of the SW test for the residual's normality and MSE of the estimated nonlinear regression model of the original response vector. We can conclude that there is insufficient evidence to select the best value of the power parameter for the original dataset into five criteria. However, after deletion of the outlier, the result was better than the outlier because the value of MLE, COD, and p-value were increased and the value of MSE decreased, as shown in Tables 4 and 5. We can conclude that a feasible solution of the optimal value of the power parameter can be deduced in the Ln(y) after deleting the outlier; this represents the basis for estimating the optimum power parameter, thus the optimal value is *Ln(y),* which is shown in Table 5. Finally, the authors believe that *Ln(y)* after deleting the outlier of displacement in the original data, which is generated by BCT to choose the optimal power parameter, as an alternative to the parametric method for the hypothesis of normality of transformed response. It is known that it is difficult to find a single solution area for two or more criteria for selecting an optimal value for the PT. However, the multiplicity of criteria provides authors with a wider area for selection and differentiation. This varies according to the type of data and becomes more difficult if there are outlier values within the data.

Data Availability Statement: The medical dataset was collected randomly for male and female at the Azadi heart center at the Duhok hospital in the Duhok Governorate in Kurdistan Region of Iraq and contained 30 observations that included a dependent variable Y representing the Systolic Blood Pressure (SBP) and an independent variable (X) as Age (Table 1).

References

1) A. S. Hadi and S. Chatterjee, *Sensitivity analysis in linear regression*. John Wiley & Sons, 2009.

2) G. E. P. Box and D. R. Cox, "An analysis of transformations," *J. R. Stat. Soc. Ser. B*, vol. 26, no. 2, pp. 211–243, 1964.

3) D. J. Poirier, "The use of the Box-Cox transformation in limited dependent variable models," *J. Am. Stat. Assoc.*, vol. 73, no. 362, pp. 284–287, 1978.

4) M. N. Abd-Rahman and T. M. Gerig, "An estimation in a linear model with response-related error," *Commun. Stat. Comput.*, vol. 14, no. 1, pp. 71–83, 1985.

5) R. D. Cook and S. Weisberg, "Transforming a response variable for linearity," *Biometrika*, vol. 81, no. 4, pp. 731–737, 1994.

6) J. W. Osborne and A. Overbay, "The power of outliers (and why researchers should always check for them)," *Pract. Assessment, Res. Eval.*, vol. 9, no. 1, p. 6, 2004.

7) J. I. Vélez, J. C. Correa, and F. Marmolejo-Ramos, "A new approach to the Box--Cox transformation," *Front. Appl. Math. Stat.*, vol. 1, p. 12, 2015.

ISSN 1533-9211
DOI: 10.5281/zenodo.8285749

8) J. Pek, O. Wong, and A. C. Wong, "Data transformations for inference with linear regression: Clarifications and recommendations," *Pract. assessment, Res. Eval.*, vol. 22, no. 1, p. 9, 2017.

9) A. C. Atkinson, M. Riani, and A. Corbellini, "The analysis of transformations for profit-and-loss data," *J. R. Stat. Soc. Ser. C (Applied Stat.*, vol. 69, no. 2, pp. 251–275, 2020.

10) A. C. Atkinson, M. Riani, and A. Corbellini, *The Box–Cox Transformation: Review and Extensions*, vol. 36, no. 2. 2021, pp. 239–255. doi: 10.1214/20-STS778.

11) M. Riani, A. C. Atkinson, and A. Corbellini, "Automatic robust Box--Cox and extended Yeo--Johnson transformations in regression," *Stat. Methods Appt.*, pp. 1–28, 2022.

12) D. V Hinkley, "On power transformations to symmetry," *Biometrika*, vol. 62, no. 1, pp. 101–111, 1975.

13) J. Neter, M. H. Kutner, C. J. Nachtsheim, W. Wasserman, and others, "Applied linear statistical models," 1996.

14) J. F. Qin, Y. Yang, H. Y. Du, and Z. J. Hong, "Outlier detection for on-line monitoring data of transformer based on wavelet transform and weighted LOF," in *IOP Conference Series: Earth and Environmental Science*, 2019, vol. 354, no. 1, p. 12108.

15) J. Raymaekers and P. J. Rousseeuw, "Transforming variables to central normality," *Mach. Learn.*, pp. 1–23, 2021.