Statistika, Vol. 7 No. 1, 51 – 54
Mei 2007

An Application of Alternative Method of Transformation

Wan Muhamad Amir W Ahmad[A)], Nyi Nyi Naing[B)], Mohd Tengku Ariff Raja Hussein[C)]

[A)]Jabatan Matematik, Fakulti Sains dan Teknologi, Malaysia, Universiti Malaysia Terengganu UMT, 21030 Kuala Terengganu, Terengganu Malaysia,
[B)]Unit Biostatistics and Research Methodology, School of Medical Sciences, Universiti Sains Malaysia, USM, Healthy Campus, 16150 Kubang Kerian, Kelantan, Malaysia,
[C)]Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia, USM, Healthy Campus, 16150 Kubang Kerian, Kelantan, Malaysia,

Abstract

Data transformation plays a major role when behave of the data do not meet the assumption of linear regression and ANOVA. In this paper we will demonstrate an application of the alternative method of transformation, so that the assumptions of ANOVA are met (or violated to a lesser degree). These data were collected from biostatistics experiments. Some of calculation is done by manually and the rest by the MINITAB software. The results will be discussed at the end of the paper.

Keywords: Alternative method of transformation, parameter λ and confidence interval.

1. Introduction to an Alternative Method of Transformation

The original of Box–Cox transformation is define as

$$Y = \rho_\lambda(X) = \begin{cases} \dfrac{X^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln X, & \lambda = 0 \end{cases}$$

In this section we will introduced an alternative method of transformation after make some modification of the original formula. This alternative method is based on Box-Cox method but its involved some of modification on the variable (see Amir et al. 2007, for a detailed discussion).

Power transformation, on the positive line is given by

$$\xi(\lambda, x) = \rho_\lambda(X) = \begin{cases} \dfrac{(X+1)^\lambda - 1}{\lambda} & (x \geq 0, \lambda \neq 0), \\ \log(X+1) & (x \geq 0, \lambda = 0), \end{cases} \tag{1}$$

And for the negative line

$$\xi(\lambda, x) = \rho_\lambda(X) = \begin{cases} -\left\{\dfrac{(-X+1)^{2-\lambda} - 1}{2-\lambda}\right\} & (x < 0, \lambda \neq 2). \\ \log(-X+1) & (x < 0, \lambda = 2), \end{cases} \tag{2}$$

2. Case Study

The data were collected from biostatistics experiments in 1964 by Dr Box and Dr Cox. In this paperwork, we will illustrate the using of an alternative method that gained from some modification of the original Box-Cox formula. We have choose and used the Box-Cox data after considering the results of their study which is very significant in statistics and the used of transformation especially. We will illustrate a case of study using the alternative formula. The data that we use in this case study contains the positive values only. Because of that, we will consider the transformation technique using the formula given in (1) only.

Normal probability plot for the original data have shown that the normality assumption is not fulfilled by the response variables. The structure of the normality plot displays the deflections in point. Thus, there is enough evidence to say that the normality assumption is contravened in this case (see Amir et al. 2006). Table 2.1 shows the dataset of lifetime animal in 3×4 factorial design of experiment with 3 level of poison factor and 4 level of treatment factor.

Table 2.1 Dataset of lifetime animal in 3×4 factorial design of experiment

Poison	Treatment			
	A	B	C	D
I	0.31	0.82	0.43	0.45
	0.45	1.10	0.45	0.71
	0.46	0.88	0.63	0.66
	0.43	0.72	0.76	0.62
II	0.36	0.92	0.44	0.56
	0.29	0.61	0.35	1.02
	0.40	0.49	0.31	0.71
	0.23	1.24	0.40	0.38
III	0.22	0.30	0.23	0.30
	0.21	0.37	0.25	0.36
	0.18	0.38	0.24	0.31
	0.23	0.29	0.22	0.33

Source : Box and Cox 1964

Determination of parameter λ

Using the MINITAB software we can obtain the optimal values of parameter λ. The given value for λ is -0.59. Using the equation given in (1) with the values of $\lambda = -0.59$, we now calculate the transformation values for the dataset of lifetime animal in Table 2.1 and the results is summarized in Table 2.2

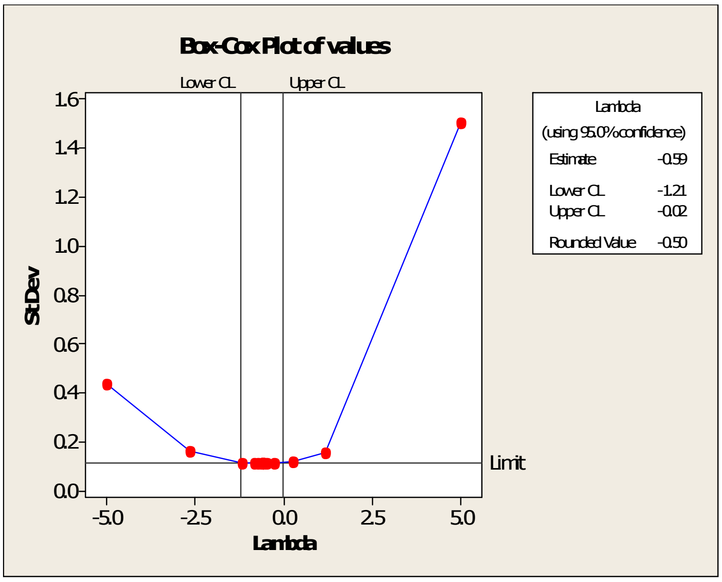

Figure 2.1 Estimation for optimal λ Values

The given control limit in confidence interval in Figure 2.1 indirectly shows the needs of transformation, the poof is given by the 95% confidence interval plot for λ, which is indicate the value without one in the range of the given interval.

The given control limit in confidence interval in Figure 2.1 indirectly shows the needs of transformation, the poof is given by the 95% confidence interval plot for λ, which is indicate the value without one in the range of the given interval.

Table 2.2 Transformed dataset of lifetime animal in 3×4 factorial design of experiment

Poison	Treatment			
	A	B	C	D
I	0.2496	0.5044	0.3225	0.3337
	0.3337	0.6009	0.3337	0.4599
	0.3392	0.5270	0.4245	0.4381
	0.3225	0.4641	0.4807	0.4198
II	0.2812	0.5415	0.3281	0.3911
	0.2364	0.4152	0.2750	0.5755
	0.3052	0.3553	0.2496	0.4599
	0.1949	0.6417	0.3052	0.2933
III	0.1876	0.24307	0.1949	0.2430
	0.1803	0.2873	0.2091	0.2812
	0.1577	0.2933	0.2020	0.2496
	0.1949	0.2364	0.1877	0.2624

Table 2.2 shows the values of transform data by using the alternative method. The formula of alternative method is given in the equation (1) and the selection for the optimal λ is generated by MINITAB software. The optimal value of parameter λ is -0.59.

Diagnostic Checking for Normality Data After using The Alternative Method of Transformation

As usual there a few steps that we need as such are plotting the normal probability plot to the response variable. Figure 2.2 illustrates the normal probability plot of the transform data. From the plot we can see that the assumption of normality is nearly fulfilled by the response variables. Figure 2.3 give the enough evidence to say that the normality assumption is not contravened in this case. To support this interpretation let us look at the Figure 1.2 for detail information.

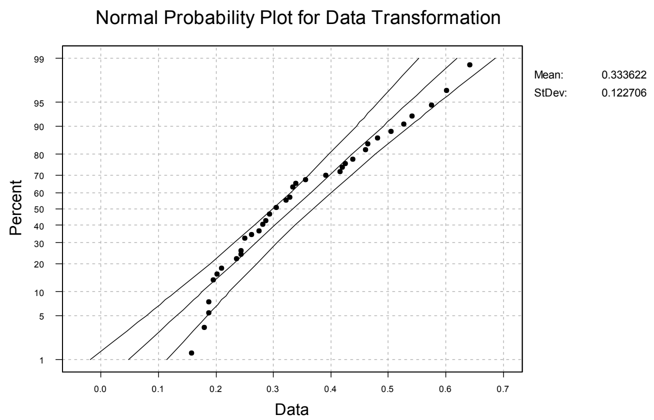

Figure 2.2 Normal probability plot of the transformed data

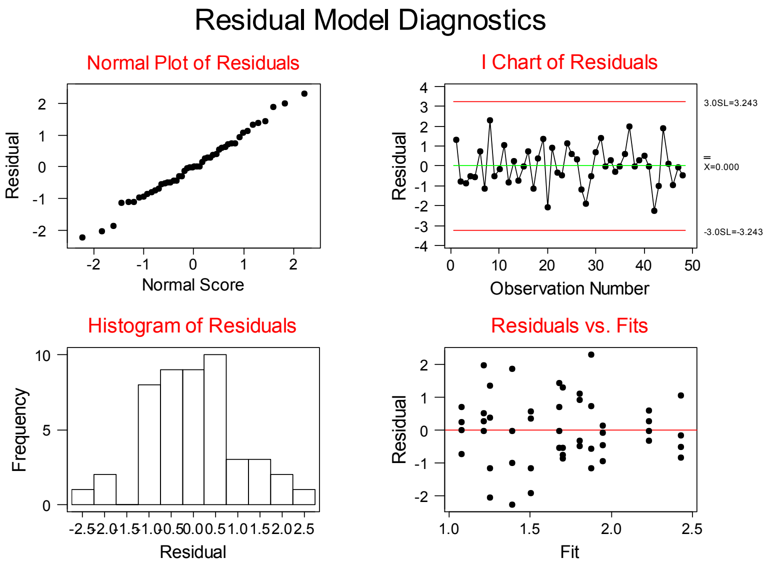

Figure 2.3 Residual Model Diagnostics for Data After Transformation

3. Discussion

The alternative method that we used in this paper is obtains from some modification of the original formula of Box-Cox and provides a simple method to determine the best way to transform the data for reducing heterogeneity of errors. From the case study in section 2, we can see that the alternative method still can provide a good result of transformation. Figure 2.2 shows the normal probability plot and the assumption of normality is nearly fulfilled by the response variables. To support the assumption of normality we again now plot for residual model diagnostics, and all of the results are shown in Figure 2.3. The plot of residual in Figure 2.3 indicates that all the residual are follow the normality assumption. These indicate that our alternative method can achieve the assumption normality as well as Box-Cox formula.

4. Reference

Amir, A., Nyi Nyi Naing, Norhayati Rosli (2006), An Approached of Box-Cox Data Transformation To Biostatistics Experiment, Jurnal STATISTIKA, Jurusan Statistika, Fakultas Matematika & Ilmu Pengetahuan Alam, Universitas Islam Bandung. Volume 6(2),

Andrew, D.F. (1971). A Note on the Selection of Data Transformation. Biometrical 58, 249-254.

Anscombe, F.J. (1961). Examination of Residual. Proc. Fourth Berkeley Symp. On Mathematical Statistics and Probability 1, 1-36. University of California Press, Berkeley.

Atkinson, A.C. (1973). Testing Transformation to Normality. Journal of the Royal Stat. Society Ser. B, 35, 473-479.

Atkinson, A.C. (1986a). Diagnostics Test for Transformations, Technometrics 28, 29-38

Bartlett, M.S. (1947) The Use of Transformation. Biometrical 3 39-52.

Box, G.E.P., and D.R. Cox (1964). An Analysis of Transformation. J.R. Stat Soc. B 26:211-252.

Damon, R. A., Jr., and W. R. Harvey. 1987. Experimental Design, ANOVA and Regression. Harper and Row, New York.

Draper, N. R., and D.R. Cox (1969). On Distribution And Their Transformation to Normality. J.R. Stat Soc. B 31:472-476.

Montgomery, D. C. 1991. Design and Analysis of Experiments (3rd Ed.). John Wiley & Sons, New York.

Peltier, M. R. 1996. Effect of melatonin implantation at the summer solstice and ovarian
status on the annual reproductive rhythm in pony mares. M.S. Thesis. University of Florida, Gainesville.

Sakia R.M. (1992). The Box-Cox Transformation Technique : a review. The Statistician, 41, 169-178.