

An Alternative Numerical Method for Finding the Value of Transformation Parameter in the Box-Cox Transformation

Lakhana Watthanacheewakul

Abstract—To analyze non normal data, the data should be transformed to a normal distribution. The well-known Box-Cox transformation can be used to get the normality. Since the transformation parameter is usually unknown, many statisticians studied the methods of estimation of the transformation parameter, λ , appearing in the Box-Cox transformation. Moreover, we can use the numerical methods for single variable function optimization to find the value of the transformation parameter λ such as bisection method, Newton's method, secant method, and others. The alternative numerical method λ is applied to find the suitable value of λ . It is better than the bisection method in the context of the number of function evaluations and the processing time.

Index Terms— Alternative numerical method, Bisection method, Box-Cox transformation, Transformation parameter

I. INTRODUCTION

In statistical data analysis the basic assumptions are that the model is additive and the errors are randomly, independently, and normally distributed about zero mean and constant variance. If the basic assumptions are not satisfied, then the statistical inference under the normal theory cannot be applied. Tukey [1] suggested that in analyzing data which do not match the assumptions of the conventional method of analysis, we have two alternative ways to go about. We may transform the data to fit the assumptions, or we may develop some new methods of analysis with assumptions fitting the original data. If we can find a satisfactory transformation, it will almost always be easier to use the conventional method of analysis rather than to develop a new one. The well-known Box-Cox transformation is often used to transform non normal data in regression analysis and the analysis of variance to normality with homogeneity of variances. Most statisticians considered the problem of estimation of the transformation parameter, λ , in the Box-Cox transformation [2-5]. Box and Cox [6] proposed the maximum likelihood method to obtain the proper value of λ by plotting the log likelihood against λ for a trial series of values. From this plot the maximizing value $\hat{\lambda}$ may be read off. Moreover, we can

use the numerical methods for single variable function optimization to find the value of transformation parameter λ such as bisection method, Newton's method, secant method, and others [7-9]. In this paper, the alternative numerical method for finding the appropriate value of λ is applied for k exponential populations with one parameter.

II. ESTIMATION OF THE TRANSFORMATION PARAMETER

The Box-Cox transformation for any sets of exponential data to normality

$$Y_{ij} = \begin{cases} \frac{X_{ij}^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln X_{ij}, & \lambda = 0, \end{cases} \quad (1)$$

for $x_{ij} > 0$, where X_{ij} is a random variable in the j th trial from the i th population, Y_{ij} the transformed variable of X_{ij} and λ a transformation parameter. The probability density function of each Y_{ij} is assumed to be normal with mean μ_i and variance σ^2 , i.e.

$$f(y_{ij} | \mu_i, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_{ij} - \mu_i)^2}{2\sigma^2}\right\}, \quad (2)$$

where y_{ij} is the observed value of Y_{ij} .

The likelihood function in terms of μ_i and σ^2 of Y_{ij} , $i = 1, \dots, k; j = 1, \dots, n_i$

$$L(\mu_i, \sigma^2 | y_{ij}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \mu_i)^2\right\}. \quad (3)$$

The maximum likelihood estimators of μ_i and σ^2 are

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij} \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \mu_i)^2, \quad \text{respectively.}$$

The likelihood in relation to the sample observations is given by

$$L(\mu_i, \sigma^2 | x_{ij}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} \left[\frac{x_{ij}^\lambda - 1}{\lambda} - \mu_i\right]^2\right\} J(\lambda; x) \quad (4)$$

$$\text{where} \quad J(\lambda; x) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left| \frac{\partial y_{ij}}{\partial x_{ij}} \right| = \prod_{i=1}^k \prod_{j=1}^{n_i} \left| \frac{\partial}{\partial x_{ij}} \left[\frac{x_{ij}^\lambda - 1}{\lambda} \right] \right| \\ = \prod_{i=1}^k \prod_{j=1}^{n_i} x_{ij}^{\lambda-1}.$$

Manuscript received September 9, 2009.

This work was supported in part by the Faculty of Science, Maejo University, Chiang Mai, Thailand.

Lakhana Watthanacheewakul is with the Faculty of Science, Maejo University, Chiang Mai, Thailand (phone: 66-053-873-540; fax: 66-053-878-225; e-mail: lakhana@mju.ac.th).

Thus, the log likelihood of the observations is

$$\ln L(\mu_i, \sigma^2, \lambda | x_{ij}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} \left[\frac{x_{ij}^\lambda - 1}{\lambda} - \mu_i \right]^2 + (\lambda - 1) \sum_{i=1}^k \sum_{j=1}^{n_i} \ln x_{ij}. \quad (5)$$

The log likelihood function of λ is, except for a constant,

$$\ln L(\lambda | x_{ij}) = -\frac{n}{2} \ln \left(\frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} \left\{ \frac{x_{ij}^\lambda - 1}{\lambda} - \frac{1}{n_i} \sum_{j=1}^{n_i} \left[\frac{x_{ij}^\lambda - 1}{\lambda} \right] \right\}^2 \right) + (\lambda - 1) \sum_{i=1}^k \sum_{j=1}^{n_i} \ln x_{ij}. \quad (6)$$

The partial derivative of $\ln L(\lambda | x_{ij})$ with respect to λ is

$$\frac{d \ln L(\lambda)}{d\lambda} = -n \left[\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^{2\lambda} \ln x_{ij} - \sum_{i=1}^k \frac{1}{n_i} \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \right) \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \ln x_{ij} \right)}{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^{2\lambda} - \sum_{i=1}^k \frac{1}{n_i} \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \right)^2} \right] + \frac{n}{\lambda} + \sum_{i=1}^k \sum_{j=1}^{n_i} \ln x_{ij}. \quad (7)$$

The maximum likelihood estimate of λ is obtained by solving the likelihood equation

$$-n \left[\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^{2\lambda} \ln x_{ij} - \sum_{i=1}^k \frac{1}{n_i} \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \right) \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \ln x_{ij} \right)}{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^{2\lambda} - \sum_{i=1}^k \frac{1}{n_i} \left(\sum_{j=1}^{n_i} x_{ij}^\lambda \right)^2} \right] + \frac{n}{\lambda} + \sum_{i=1}^k \sum_{j=1}^{n_i} \ln x_{ij} = 0. \quad (8)$$

The maximum likelihood estimate of λ is obtained by solving the likelihood equation. This equation is not easy to solve for λ in terms of the observations x_{ij} and, hence, a numerical approach should be considered.

III. AN ALTERNATIVE NUMERICAL METHOD

Since the curvature of the log likelihood function is of a quadratic pattern, it has only one maximum value. If it was ever desired to determine $\hat{\lambda}$ more precisely this could be done by numerically determining the value $\hat{\lambda}$ for which the derivatives with respect to λ are all zero [6]. Numerical methods, such as the bisection method, can be applied, but here an alternative numerical method is applied to find λ . The idea behind this method is that the slope of curvature of the log likelihood function is used to find λ and the transformation parameter is obtained when the slope is nearly zero. The first derivative of the log likelihood function of λ , or the slope given by (7), is denoted by $f'(\lambda)$. In what follows it will always be assumed that $f'(\lambda)$ is continuous on $[\lambda_1, \lambda_2]$. Watthanacheewakul [10] proposed the algorithm for the alternative numerical method to obtain $\hat{\lambda}$ in the alternative transformation is as follows:

Step 1 The Appropriate Interval

Step 1.1 Initialize λ_1, λ_2 , and ε

First of all, the tolerance value for the stopping rule, ε , is set at 0.0001 and λ_1, λ_2 are arbitrary values where $\lambda_1 < \lambda_2$.

Step 1.2 Calculate the value of the first derivatives of the log likelihood of λ_1 and λ_2

The values of the first derivatives of the log likelihood of λ_1 and λ_2 are calculated by replacing λ in $f'(\lambda)$ with λ_1 and λ_2 respectively.

Step 1.3 Compare $|f'(\lambda_1)|$ and $|f'(\lambda_2)|$ with ε

If $|f'(\lambda_1)| < \varepsilon$ then let $\lambda = \lambda_1$. If $|f'(\lambda_1)| \geq \varepsilon$, then consider $|f'(\lambda_2)|$. If $|f'(\lambda_2)| < \varepsilon$, then $\lambda = \lambda_2$. Otherwise, proceed to

Step 1.4.

Step 1.4 Find the appropriate interval $[\lambda_1, \lambda_2]$

If $|f'(\lambda_1)| \geq \varepsilon$ and $f'(\lambda_1) < 0$, let $m = |\lambda_1 - \lambda_2|$, then set $\lambda'_2 = \lambda_1$ and $\lambda'_1 = \lambda'_2 - m/4 = \lambda_1 - m/4$. Set $\lambda_1 = \lambda'_1$. Repeat this process until $f'(\lambda_1) > 0$.

If $|f'(\lambda_2)| \geq \varepsilon$ and $f'(\lambda_2) > 0$, let $m = |\lambda_1 - \lambda_2|$, then set $\lambda'_1 = \lambda_2$ and $\lambda'_2 = \lambda'_1 + m/4 = \lambda_2 + m/4$. Set $\lambda_2 = \lambda'_2$. Repeat this process until $f'(\lambda_2) < 0$.

If $f'(\lambda_1) > 0$ and $f'(\lambda_2) < 0$ then proceed to **Step 2**.

Step 2 The Optimization Value

Step 2.1 Compare $|f'(\lambda_1)|$ and $|f'(\lambda_2)|$

Calculate $m = |\lambda_1 - \lambda_2|$. If $|f'(\lambda_1)| < |f'(\lambda_2)|$, let $\lambda = \lambda_1 + m/4$, otherwise let $\lambda = \lambda_2 - m/4$.

Step 2.2 Check for a stopping rule

If $|f'(\lambda)| < \varepsilon$, stop the iteration. If $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) > 0$, let $\lambda_1 = \lambda$ and go back to **Step 2.1**. If $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) < 0$, let $\lambda_2 = \lambda$ and go back to **Step 2.1**.

This algorithm is applied to find the value of transformation parameter in the Box-Cox transformation.

IV. APPLYING THE ALTERNATIVE NUMERICAL METHOD

The random samples of size 20 were taken from exponential population with scale parameter (β) equals three. The data are as follows 13.51945, 0.071361, 0.347777, 5.577101, 0.605019, 5.295623, 0.487435, 1.374586, 8.99852, 0.91593, 1.540253, 5.55445, 9.484076, 5.394369, 1.776799, 2.992351, 1.597922, 3.602009, 0.179283, 3.424097. They are used for finding the transformation parameter using the alternative numerical method. The procedure is as follows:

Step 1 The Appropriate Interval Round No.1

Step 1.1 Let $\lambda_1 = -1.0$, $\lambda_2 = 0.01$ and $\varepsilon = 0.0001$.

Step 1.2 Calculate $f'(\lambda_1) = 44.442623$

$$\text{and } f'(\lambda_2) = 9.038474.$$

Step 1.3 $|f'(\lambda_1)| \geq \varepsilon$ and $|f'(\lambda_2)| \geq \varepsilon$.

Step 1.4 Since $f'(\lambda_2) > 0$, $m = |\lambda_1 - \lambda_2| = |-1 - 0.01| = 1.01$.

$$\text{Set } \lambda'_1 = 0.01 \text{ and } \lambda'_2 = \lambda'_1 + m/4 = \lambda_2 + m/4$$

$$= 0.01 + 0.2525 = 0.2625. \text{ Set } \lambda_2 = \lambda'_2 \text{ and}$$

$$f'(\lambda_2) = -0.553462.$$

The appropriate interval is [0.01, 0.2625].

Step 2 The optimization value

Round No.1

Step 2.1 Calculate $m = |\lambda_1 - \lambda_2| = |0.01 - 0.2625| = 0.2525$.

$$f'(\lambda_1) = 9.038474 \text{ and } f'(\lambda_2) = -0.553462.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.199375.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = 1.673146$, let

$$\lambda_1 = \lambda = 0.199375.$$

The new interval is [0.199375, 0.262500].

Round No.2

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.199375 - 0.262500| = 0.063125.$$

$$f'(\lambda_1) = 1.673146 \text{ and } f'(\lambda_2) = -0.553462.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.246719.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = -0.008474$, let

$$\lambda_2 = \lambda = 0.246719.$$

The new interval is [0.199375, 0.246719].

Round No.3

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.199375 - 0.246719| = 0.047344.$$

$$f'(\lambda_1) = 1.673146 \text{ and } f'(\lambda_2) = -0.008474.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.234883.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = 0.405402$, let

$$\lambda_1 = \lambda = 0.234883.$$

The new interval is [0.234883, 0.246719].

Round No.4

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.234883 - 0.246719| = 0.011836.$$

$$f'(\lambda_1) = 0.405402 \text{ and } f'(\lambda_2) = -0.008474.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.243760.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = 0.094584$, let

$$\lambda_1 = \lambda = 0.243760.$$

The new interval is [0.243760, 0.246719].

Round No.5

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.243760 - 0.246719| = 0.002959.$$

$$f'(\lambda_1) = 0.094584 \text{ and } f'(\lambda_2) = -0.008474.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.245979$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = 0.017265$, let

$$\lambda_1 = \lambda = 0.245979.$$

The new interval is [0.245979, 0.246719].

Round No.6

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.245979 - 0.246719| = 0.000740.$$

$$f'(\lambda_1) = 0.017265 \text{ and } f'(\lambda_2) = -0.008474.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.246534.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = -0.002040$, let

$$\lambda_2 = \lambda = 0.246534.$$

The new interval is [0.245979, 0.246534].

Round No.7

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.245979 - 0.246534| = 0.000555.$$

$$f'(\lambda_1) = 0.017265 \text{ and } f'(\lambda_2) = -0.002040.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.246395.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = 0.002785$, let

$$\lambda_1 = \lambda = 0.246395.$$

The new interval is [0.246395, 0.246534].

Round No.8

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.246395 - 0.246534| = 0.000139.$$

$$f'(\lambda_1) = 0.002785 \text{ and } f'(\lambda_2) = -0.002040.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.246499.$$

Step 2.2 Since $|f'(\lambda)| \geq \varepsilon$ and $f'(\lambda) = -0.000834$, let

$$\lambda_2 = \lambda = 0.246499.$$

The new interval is [0.246395, 0.246499].

Round No.9

Step 2.1 Calculate

$$m = |\lambda_1 - \lambda_2| = |0.246395 - 0.246499| = 0.000104.$$

$$f'(\lambda_1) = 0.002785 \text{ and } f'(\lambda_2) = -0.000834.$$

$$\text{Since } |f'(\lambda_1)| > |f'(\lambda_2)|, \lambda = \lambda_2 - \frac{m}{4} = 0.246473.$$

Step 2.2 Since $|f'(\lambda)| = 0.000071, |f'(\lambda)| < \varepsilon$. So

$$\lambda = 0.246473.$$

The value of transformation parameter is $\lambda = 0.246473$.

V. THE SIMULATION METHOD

Exponential populations of size $N_i = 5,000$ were generated for the values of parameters β_i , for $i = 1, \dots, k$. From each generated population, 1,000 random samples, each of size n_i , for $i = 1, \dots, k$, were drawn.

The values of parameters are set as follows:

- 1) k = number of the populations = 3
- 2) n_i = sample sizes from the i th exponential population

between 10 and 150, for $i = 1, \dots, k$

3) β_i , the scale parameter of the i th exponential population, is between 1 and 5.

The results are shown in Table I-III. For any k exponential populations, both equal sample sizes and unequal sample sizes, both equal scale parameters and unequal scale parameters, the alternative numerical method for finding the transformation parameter, λ , is better than the bisection method in the context of the number of function evaluations and processing time.

Table I Average of Number of Function Evaluations and Processing Time for the Bisection and Alternative Numerical Methods for k=1

| n | β | Average of Number of Function Evaluations | | Processing Time (Seconds) | |
|-----|---------|---|-------------|---------------------------|-------------|
| | | Bisection | Alternative | Bisection | Alternative |
| 10 | 1 | 16.890 | 12.283 | 5.250 | 4.937 |
| | 3 | 16.873 | 12.255 | 5.250 | 5.062 |
| | 5 | 17.115 | 12.405 | 5.093 | 4.859 |
| 30 | 1 | 19.112 | 13.495 | 6.000 | 5.515 |
| | 3 | 18.913 | 13.515 | 6.235 | 5.672 |
| | 5 | 19.131 | 13.767 | 6.000 | 5.687 |
| 100 | 1 | 21.030 | 15.127 | 9.422 | 8.578 |
| | 3 | 21.326 | 15.378 | 10.141 | 8.766 |
| | 5 | 21.078 | 15.144 | 10.140 | 8.656 |

Table II Average of Number of Function Evaluations and Processing Time for the Bisection and Alternative Numerical Methods for k=2

| n_i | Scale parameter | | Average of Number of Function Evaluations | | Processing Time (Seconds) | |
|------------------------|-----------------|-----------|---|-------------|---------------------------|-------------|
| | β_1 | β_2 | Bisection | Alternative | Bisection | Alternative |
| $n_1 = 10, n_2 = 10$ | 1 | 1 | 18.290 | 13.167 | 5.593 | 5.328 |
| | 3 | 5 | 18.776 | 13.795 | 5.844 | 5.594 |
| $n_1 = 10, n_2 = 15$ | 1 | 1 | 18.685 | 13.463 | 5.921 | 5.485 |
| | 3 | 5 | 18.876 | 13.637 | 5.906 | 5.500 |
| $n_1 = 30, n_2 = 30$ | 1 | 1 | 20.428 | 14.675 | 7.718 | 6.906 |
| | 3 | 5 | 20.420 | 14.627 | 8.063 | 7.141 |
| $n_1 = 30, n_2 = 45$ | 1 | 1 | 20.669 | 14.886 | 8.687 | 7.562 |
| | 3 | 5 | 20.991 | 15.105 | 8.828 | 7.734 |
| $n_1 = 100, n_2 = 100$ | 1 | 1 | 22.111 | 16.061 | 16.297 | 13.422 |
| | 3 | 5 | 22.459 | 16.242 | 17.547 | 13.594 |
| $n_1 = 100, n_2 = 150$ | 1 | 1 | 22.476 | 16.333 | 19.938 | 16.063 |
| | 3 | 5 | 22.593 | 16.425 | 19.765 | 15.969 |

Table III Average of Number of Function Evaluations and Processing Time for the Bisection and Alternative Numerical Methods for k=3

| n_i | Scale parameter | | | Average of Number of Function Evaluations | | Processing Time (Seconds) | |
|-----------------------------------|-----------------|-----------|-----------|---|-------------|---------------------------|-------------|
| | β_1 | β_2 | β_3 | Bisection | Alternative | Bisection | Alternative |
| $n_1 = 10, n_2 = 10, n_3 = 10$ | 1 | 1 | 1 | 19.109 | 13.656 | 6.375 | 5.938 |
| | 1 | 3 | 5 | 19.952 | 14.645 | 6.234 | 5.828 |
| $n_1 = 10, n_2 = 20, n_3 = 30$ | 1 | 1 | 1 | 20.226 | 14.513 | 8.281 | 7.078 |
| | 1 | 3 | 5 | 20.558 | 14.756 | 8.234 | 7.328 |
| $n_1 = 30, n_2 = 30, n_3 = 30$ | 1 | 1 | 1 | 20.978 | 15.249 | 7.515 | 6.360 |
| | 1 | 3 | 5 | 21.501 | 15.414 | 7.593 | 6.407 |
| $n_1 = 30, n_2 = 45, n_3 = 60$ | 1 | 1 | 1 | 21.659 | 15.551 | 10.032 | 8.235 |
| | 1 | 3 | 5 | 22.216 | 15.827 | 10.094 | 8.235 |
| $n_1 = 100, n_2 = 100, n_3 = 100$ | 1 | 1 | 1 | 22.739 | 16.491 | 18.860 | 14.812 |
| | 1 | 3 | 5 | 23.259 | 16.526 | 17.985 | 14.843 |
| $n_1 = 100, n_2 = 150, n_3 = 200$ | 1 | 1 | 1 | 23.519 | 17.095 | 27.750 | 21.484 |
| | 1 | 3 | 5 | 23.826 | 16.767 | 27.829 | 21.015 |

VI. CONCLUSION

For analyzing data which do not match the assumptions of the conventional method of analysis, the data should first be transformed to fit all the assumptions required. The well-known Box-Cox transformation can be used to get the normality. The transformation parameter (λ) is unknown. The alternative numerical method is applied to find the appropriate value of λ . The value of it is obtained when the slope of curvature of the log likelihood function is nearly zero. From simulation results, this method is better than the bisection method in the context of the number of function evaluations and the processing time.

REFERENCES

- [1] W. Tukey, "On the comparative anatomy of transformations," *Annals of Mathematical Statistics*, vol. 23, 1957, pp. 525-540.
- [2] D.F. Andrews, "A Note on the Selection of Data Transformations," *Biometrika*, vol.58, 1971, pp. 249-254.
- [3] R.J. Carroll, "A Robust method for testing transformations to Achieve Approximate Normality," *Journal of the Royal Statistical Society, Ser.B*, vol. 42, 1980, pp. 71-78.
- [4] A.K. Han, "A non-parametric analysis of transformations," *Journal of Econometrics*, vol.35, 1987, pp.191-209.
- [5] D.V. Hinkley, On Power Transformations to Symmetry. *Biometrika*, vol.62(1), 1975, pp. 101-111.
- [6] G. E. P. Box, D. R. Cox, "An analysis of transformations (with discussion)," *Journal of the Royal Statistical Society, Ser.B*. vol. 26, 1964, pp. 211-252
- [7] R. E. Scraton, *Basic Numerical Methods*. London: Edward Arnold, 1984.
- [8] S. Abbasbandy, "Improving Newton-Raphson Method for Nonlinear Equations by Modified Adomian Decomposition Method," *Applied Mathematics and Computation*, vol.145, 2003, pp. 887-893.
- [9] M. Subasi, N. Yildirim, and B. Yildiz, "An Improvement on Fibonacci search method in optimization theory," *Applied Mathematics and Computation*. vol.147, 2004, pp. 893-901.
- [10] L. Wathanacheewakul, *Testing Means of Several Pareto Distributions*. Ph. D. Dissertation, National Institute of Development Administration, 2007, pp. 41-42.