流通科学大学論集—経済・情報・政策編—第 23 巻第 1 号，1-18（2014）

Examination of Transformations to Normality: Part Ⅱ

Takafumi Isogai*

The idea of an $R(x)$ plot in Tarter and Kowalski (1970), which is defined by the ratio of densities with a non-normal distribution and a normal, is developed to examine approximation problems in normalizing transformation theory. A numerical method enables us to detect a functional form of $R(x)$ easily, and to introduce several approximation formulas for $R(x)$. Performance of those approximation formulas is examined by several examples. The accuracy of our approximations is shown to be fairly good.

Keywords: normalizing transformation, $R(x)$ plot, 2nd order differential equation, Runge-Kutta algorithm, distributions of Cauchy, Beta, F, Student's t and sample correlation coefficients

Ⅰ. Introduction

Let Z be a random variable distributed as a normal and let X be a random variable distributed as another distribution. Let us consider the relationship $Z=T(X)$, where T denotes a transformation function to normality. Tarter and Kowalski (1970) proposed an $R(x)$ $(=1/\frac{dz}{dx}=1/\frac{dT}{dx})$ plotting method to detect the functional form of a normalizing transformation T. In this study, $R(x)$ plots and normalizing transformation functions are examined by graphical representations for numerical solutions of appropriate 2nd order differential equations proposed by Kaskey, Kolman, Steinberg and Krishnaiah (1980).

For one-parameter families Efron (1982) considered the existence problem of a transformation T such that $T(x)$ = $\alpha_\theta+\beta_\theta g(x)$, where α_θ and β_θ are functions of the population parameter θ only, and $g(x)$ is independent of the parameter θ. At the same time he gave a diagnostic tool for the existence of such a function $g(x)$. Performance of the Efron's diagnostic method has been examined and comparison with other diagnostic methods has been done in Isogai (2014). The results suggest that in a general case a transformation T satisfying the above Efron's assumption seldom exists.

In the present paper we shall consider how to approximate a normalizing transformation T, when there does not exist a transformation T in the meaning of Efron (1982). There are two approaches to deal with our problem. One is to approximate a transformation function T directly by some heuristic function motivated by $R(x)$ plots and graphical representations of solution curves. The other approach is to

* University of Marketing and Distribution Sciences, Faculty of Commerce, 3-1, Gakuen-nishimachi, Nishi-ku, Kobe 651-2188, JAPAN

（2014 年 3 月 28 日受理）

©2014 UMDS Research Association

approximate $R(x)$ itself.

Thus the organization of the paper is as follows. In Section 2 we shall briefly review the numerical method provided by Kaskey et al. (1980), and the Efron's (1982) diagnostic method. In Section 3 we show heuristic approaches through concrete examples of Cauchy and F distributions. In Section 4 we try to approximate $R(x)$ by some simple functions, and give several examples.

Ⅱ. Review of a differential equation and diagnostic methods

Let $F_\theta(x)$ be a distribution function having the density $f_\theta(x)$ which is positive in some interval [L, U] and zero outside of this interval, and continuously differentiable with respect to x and θ. Suppose that the end points L and U do not depend on θ. Let $\Phi(z)$ be the standard normal distribution function with the density $\phi(z)$. The transformation function $z = T(x)$ is defined by

$$z = T(x) = \Phi^{-1}(F_\theta(x)) \tag{1}$$

or equivalently defined by

$$\int_{-\infty}^{z} \phi(u)\,du = \int_{L}^{x} f_\theta(t)\,dt\,.$$

Differentiating the equation (1) twice with respect to x, we get the following 2nd order differential equation

$$\ddot{z} = z(\dot{z})^2 + \left\{\frac{\partial}{\partial x}\log f_\theta(x)\right\}\dot{z} \tag{2}$$

where we put

$$\begin{cases} \dfrac{dz}{dx} = \dot{z}, \\ \dfrac{d^2 z}{dx^2} = \ddot{z} \end{cases}.$$

The transformation function $z = T(x)$ is given as a solution $z = z(x)$ of the equation (2) with suitable initial conditions. We shall take medians of both distributions $\Phi(z)$ and $F_\theta(x)$ as initial conditions of (2), which are given by

$$\begin{cases} z(x_{0.5}) = 0, \\ \dot{z}(x_{0.5}) = \sqrt{2\pi}\, f_\theta(x_{0.5}) \end{cases} \tag{3}$$

where 100α $(0<\alpha<1)$ percent points z_α and x_α are defined by

$$\Phi(z_\alpha) = \alpha,$$
$$F_\theta(x_\alpha) = \alpha$$

and the solution $z = z(x)$ of (2) satisfies the relationship:

$$z(x_\alpha) = z_\alpha\,.$$

The numerical solution $z = z(x)$ is obtained by the Runge-Kutta algorithm over the interval $[x_{0.001}, x_{0.999}]$.

Efron (1982) considered the existence problem of a transformation function $z = T(x)$ which can be expressed as

$$z = z(x,\ \theta) = \frac{g(x) - \nu_\theta}{\sigma_\theta} \tag{4}$$

with some function $g(x)$ independent of the population parameter θ. Here ν_θ is the median of $g(x)$ and σ_θ the standard deviation of $g(x)$.

The Efron's diagnostic method for the existence of such a function $g(x)$ is to plot

$$\left(z,\ \frac{\partial_\theta z}{\partial_\theta z\left(x_{0.5}\right)}\right)$$

where we put

$$\frac{\partial z}{\partial \theta} = \partial_\theta z .$$

The linearity of this graph means the existence of a function $g(x)$ independent of θ .

If $\partial_\theta z\left(x_{0.5}\right) = 0$, we have to modify the Efron's diagnostic method. The modified Efron's method is to plot

$$\left(z,\ \partial_\theta z - \partial_\theta z\left(x_{0.5}\right)\right) .$$

The linearity of the graph $\left(z,\ \partial_\theta z - \partial_\theta z\left(x_{0.5}\right)\right)$ means the existence of a function $g(x)$ independent of θ .

Our new diagnostic method proposed by Isogai (2014) is to plot

$$\left(z,\ \frac{\partial_\theta \dot{z}}{\dot{z}}\right).$$

A constant tendency of the graph $\left(z,\ \frac{\partial_\theta \dot{z}}{\dot{z}}\right)$ means the existence of a function $g(x)$ independent of θ .

Our new diagnostic method has the most stable performance among three diagnostic methods . These three diagnostic methods are easily extended to a multi-parameter case where the population parameter θ of $F_\theta\left(x\right)$ is a p-dimensional vector $\boldsymbol{\theta}' = \left(\theta_1, \theta_2, \cdots, \theta_p\right)$. For the details, also see Isogai (2014).

III. Heuristic approximations for a normalizing transformation

It happens that the Efron's assumption does not hold, and that $R(x)$ plots and graphical representations of solution curves $z = T(x)$ do not give us much information about their functional forms. In such cases we have to seek a functional form of $z = T(x)$ in a heuristic way. In the following we give two examples of approaches. The first one is a case for the standard Cauchy distribution, where we try to approximate $z =$

$T(x)$ directly. The second one is a case for F distributions, where there is some hypothetical function $g_\theta(x)$ such that $g_\theta(x)$ approximately satisfies $T(x)=(g_\theta(x)-\nu_\theta)/\sigma_\theta$.

1. Trial and error method for the standard Cauchy distribution

Figure 1 shows the solution curve $z = T(x)$ and the $R(x)$ plot for the standard Cauchy distribution, which is also denoted by $t(1)$, Student's t distribution with one degree of freedom. For the density of $t(1)$, see the later equation (28). By the method of trial and error, we have obtained an approximation function $\hat{z}$ for the transformation $z = T(x)$ from the standard Cauchy to a normal distribution as

$$\hat{z}=1.370\ sign(x)\ \log\left\{1+\left(\log\left(1+|x|\right)\right)^{1.220}\right\} \quad \text{for} \quad x_{0.001}\le x\le x_{0.999}\ . \qquad (5)$$

Figure 2 shows the graph of $\hat{z}$ against z and also shows the residual plotting of $\hat{z}$ - z against z. The degree of the difference $\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right|$ is less than 0.0055 for $x_{0.001}\le x\le x_{0.999}$.

Regarding a standard for the degree of our approximation, the precision is expressed as "not so bad", "good" or "excellent", according as the value of $\max_x\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right|$ is equal to 0.025, 0.01 or 0.005.

The above approximation is nearly "excellent".

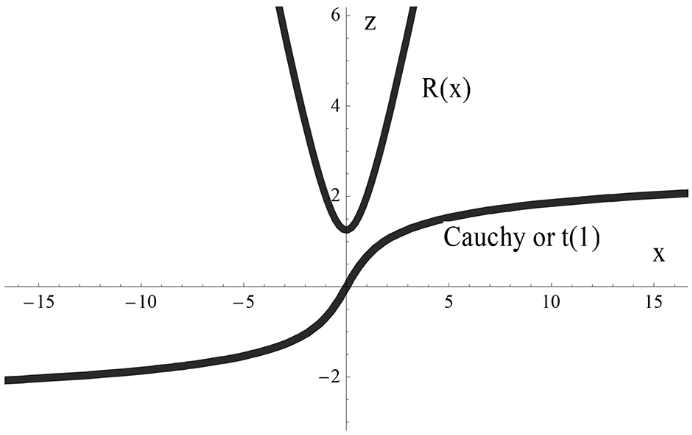

Figure 1. The solution curve $z = T(x)$ of the standard Cauchy distribution.

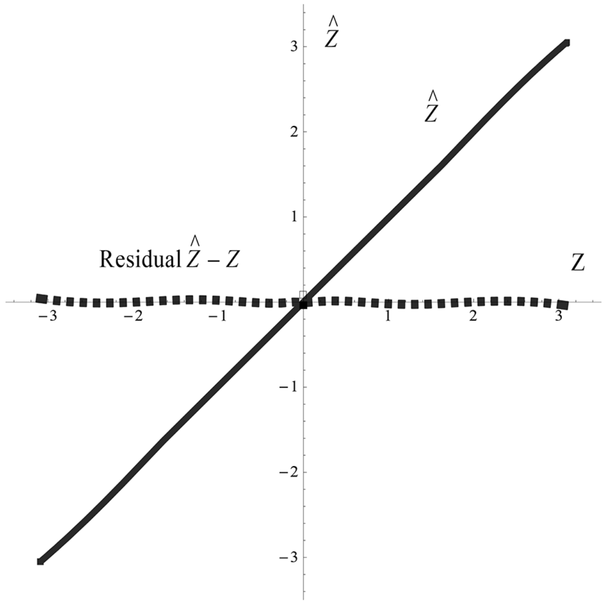

Figure 2. An approximation function $\hat{z}$ (solid line) for the standard Cauchy distribution and the residual $\hat{z}$ - z (dashed line), where $\hat{z} = 1.370\ sign(x)\ \log\left\{1+\left(\log\left(1+|x|\right)\right)^{1.220}\right\}$.

2. Power transformation for F distributions

Suppose that a random variable X is distributed as the central F distribution with (θ_1, θ_2) degrees of freedom (denoted by $\mathrm{F}(\theta_1, \theta_2)$ or $F_{\theta_2}^{\theta_1}$ in the following). The density function of $F_{\theta_2}^{\theta_1}$ is

$$f_F(x) = \frac{(\theta_1/\theta_2)^{\theta_1/2}}{B\left(\frac{1}{2}\theta_1, \frac{1}{2}\theta_2\right)} \frac{x^{(\theta_1/2)-1}}{\left(1+\frac{\theta_1}{\theta_2}x\right)^{(\theta_1+\theta_2)/2}}, \quad x > 0, \tag{6}$$

with $\theta_1 > 0$, $\theta_2 > 0$.

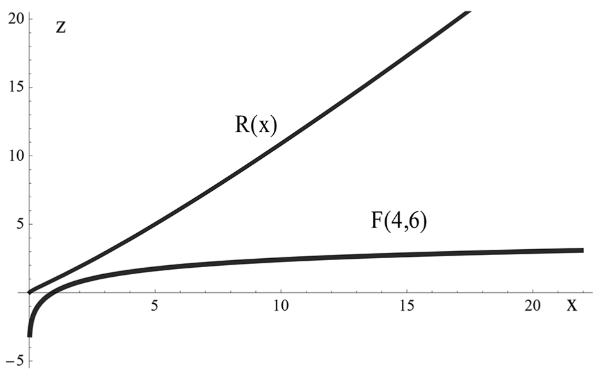

Figure 3. $R(x)$ plot and the solution curve $z = T(x)$ for F(4, 6).

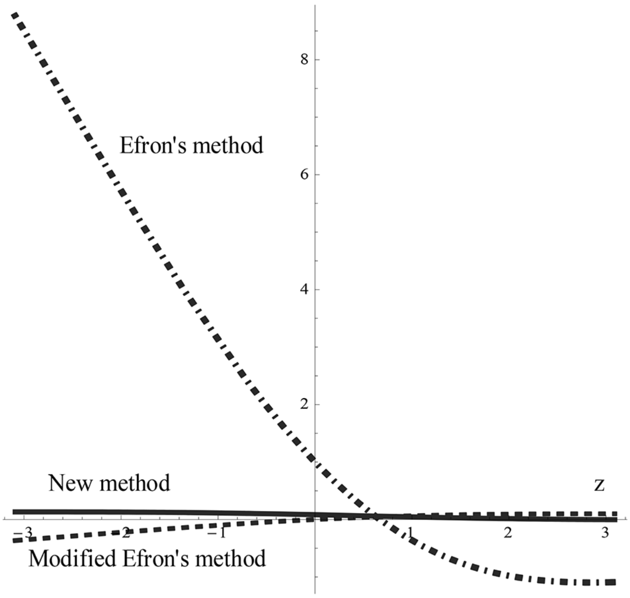

Figure 4. Three diagnostic plots for the upper parameter of F(4,6):
Efron's method (dashed-dotted line), Modified Efron's method (dashed line),
and our new method (solid line).

Figure 3 shows the solution curve $z = T(x)$ and the $R(x)$ plot for F(4,6). Figure 4 also shows three diagnostic plots with respect to the upper parameter θ_1 of F(4,6).

In Figure 4, non-linearity of Efron's and modified Efron's plots and non-constancy of our new method suggest the non-existence of a function $g(x)$ that satisfies $z = T(x) = \left(g(x) - \nu_{\boldsymbol{\theta}}\right)/\sigma_{\boldsymbol{\theta}}$ and is independent of the population parameter $\boldsymbol{\theta} = (\theta_1, \theta_2)$. However, in Figure 3, an almost linear configuration of the $R(x)$ plot and a monotone tendency of the solution curve $z = T(x)$ indicate a possibility that $z = T(x)$ can be approximated by $T(x) = \left(g_{\boldsymbol{\theta}}(x) - \nu_{\boldsymbol{\theta}}\right)/\sigma_{\boldsymbol{\theta}}$ such that $g_{\boldsymbol{\theta}}(x) = x^{h(\boldsymbol{\theta})}$, where $h(\boldsymbol{\theta})$ is a function of the population parameter $\boldsymbol{\theta} = (\theta_1, \theta_2)$.

A functional form of $h(\boldsymbol{\theta})$ is detected by a symmetrization principle (Wilson and Hilferty (1936)). The power parameter $h(\boldsymbol{\theta})$ can be chosen so as to make the third order cumulant of the power transformed F variable $X^{h(\boldsymbol{\theta})}$ equal to zero. We review briefly the results in Isogai (1999) and give several examples.

A simple formula for $h(\boldsymbol{\theta})$ is given by

$$\hat{h}_F = \hat{h}_F(\theta_1, \theta_2) = -\frac{1}{3}\left(\frac{\theta_1 - \theta_2}{\theta_1 + \theta_2}\right) \quad \text{for} \quad \theta_1 > 2\,, \theta_2 > 2\,. \tag{7}$$

Also, a simple formula for the median $\hat{F}_{\theta_2}^{\theta_1}(0.5)$ of $F_{\theta_2}^{\theta_1}$ is given by

$$\hat{F}_{\theta_2}^{\theta_1}(0.5) = \exp\left[-\frac{2}{3}\left(\frac{1}{\theta_1} - \frac{1}{\theta_2}\right)\right] \quad \text{for} \quad \theta_1 > 2\,, \theta_2 > 2\,. \tag{8}$$

Then, a simple formula for a normalizing transformation function $T(x) = \left(g_{\boldsymbol{\theta}}(x) - \nu_{\boldsymbol{\theta}}\right) / \sigma_{\boldsymbol{\theta}}$ is obtained as

$$\hat{z}_F = \hat{T}_F(x) = \frac{x^{\hat{h}_F} - \left\{\hat{F}_{\theta_2}^{\theta_1}(0.5)\right\}^{\hat{h}_F}}{\hat{h}_F\left[2\left(1/\theta_1 + 1/\theta_2\right)\right]^{1/2}} \quad \text{for} \quad \theta_1 > 2\,, \theta_2 > 2\,. \tag{9}$$

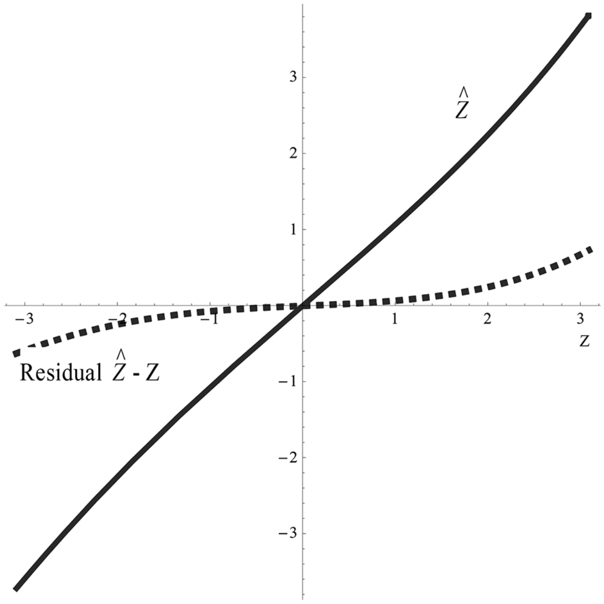

Figure 5. An approximation function $\hat{z}_F$ (solid line) for F(4,6) and the residual $\hat{z}_F - z$ (dashed line).

Figure 5 shows the graph of $\hat{z}_F$ against z for F(4,6) and the residual plotting of $\hat{z}_F$ - z against z. The effect of symmetrization about $\hat{z}_F$ is good. Though the difference $\hat{z}_F$ - z seems large, the degree of the difference $\left|\Phi\left(\hat{z}_F\right) - \Phi\left(z\right)\right|$ is less than 0.0185 for $x_{0.001} \le x \le x_{0.999}$. Our approximation $\hat{z}_F$ is "not so bad". The degree of approximation about $\hat{z}_F$ increases as θ_1 and θ_2 become large (see Isogai (1999)).

As an application of the power transformation for $F_{\theta_2}^{\theta_1}$, we shall consider a family of Beta distributions. A beta random variable is defined by

$$B_{\psi_2}^{\psi_1} = \frac{\psi_1\left(F_{2\psi_2}^{2\psi_1}\right)}{\psi_2 + \psi_1\left(F_{2\psi_2}^{2\psi_1}\right)}, \tag{10}$$

where $F_{2\psi_2}^{2\psi_1}$ denotes a random variable distributed as the central F distribution with $\left(2\psi_1, 2\psi_2\right)$ degrees of freedom. The distribution of $B_{\psi_2}^{\psi_1}$ is called as the standard Beta distribution with parameters ψ_1, ψ_2 and denoted by $\text{Beta}(\psi_1, \psi_2)$. The density function of $\text{Beta}(\psi_1, \psi_2)$ is

$$f_B(x) = \frac{1}{B(\psi_1, \psi_2)} x^{\psi_1 - 1} (1-x)^{\psi_2 - 1}, \quad 0 \le x \le 1, \tag{11}$$

with $\psi_1 > 0$, $\psi_2 > 0$.

Figure 6 shows the solution curve $z = T(x)$ and the $R(x)$ plot for Beta(2,3). Figure 7 also shows three diagnostic plots with respect to the lower parameter ψ_2 of Beta(2,3).

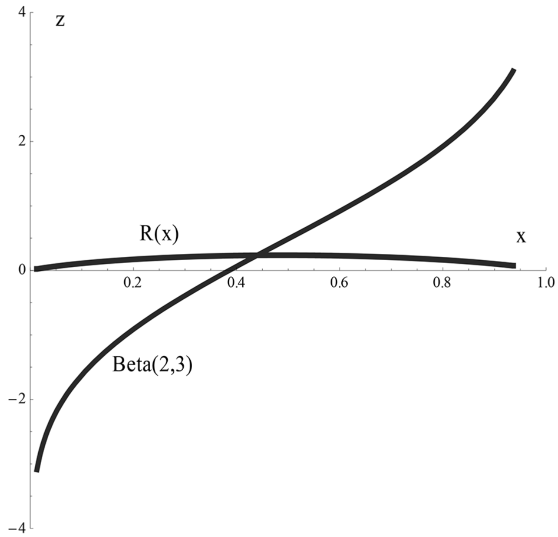

Figure 6. The solution curve $z = T(x)$ and the $R(x)$ plot for Beta(2,3).

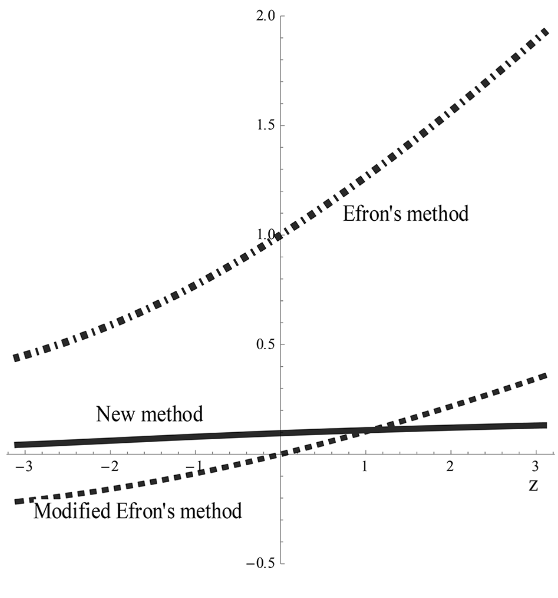

Figure 7. Three diagnostic plots for the lower parameter of Beta(2,3): Efron’s method (dashed-dotted line), Modified Efron’s method (dashed line), and our new method (solid line).

Clearly, in Figure 7, non-linearity of Efron's and modified Efron's plots and non-constancy of our new method suggest the non-existence of a function $g(x)$ that satisfies $z = T(x) = \left(g(x) - \nu_{\psi}\right) / \sigma_{\psi}$ and is independent of the population parameter $\boldsymbol{\psi} = (\psi_1, \psi_2)$. However, using the relationship (10), we can construct simple formulas of symmetrizing transformation for $\mathrm{Beta}(\psi_1, \psi_2)$ through the results for the power transformation for $F_{2\psi_2}^{2\psi_1}$.

A simple formula for the median $\hat{B}_{\psi_2}^{\psi_1}(0.5)$ of $\mathrm{Beta}(\psi_1, \psi_2)$ is given by

$$\hat{B}_{\psi_2}^{\psi_1}(0.5) = \frac{\psi_1\left(\hat{F}_{2\psi_2}^{2\psi_1}(0.5)\right)}{\psi_2 + \psi_1\left(\hat{F}_{2\psi_2}^{2\psi_1}(0.5)\right)} = \frac{\psi_1 \exp\left(-\frac{1}{3\psi_1}\right)}{\psi_1 \exp\left(-\frac{1}{3\psi_1}\right) + \psi_2 \exp\left(-\frac{1}{3\psi_2}\right)}, \quad \psi_1 > 1, \psi_2 > 1. \quad (12)$$

Also, a simple formula of a symmetrizing transformation for $\mathrm{Beta}(\psi_1, \psi_2)$ is given by

$$\hat{z}_B = \hat{T}_B(x) = \frac{\left(\frac{\psi_2 x}{\psi_1(1-x)}\right)^{\hat{h}_B} - \left(F_{2\psi_2}^{2\psi_1}(0.5)\right)^{\hat{h}_B}}{\hat{h}_B\left(1/\psi_1 + 1/\psi_2\right)^{1/2}} \quad (13)$$

where

$$\hat{h}_B = \hat{h}_F(2\psi_1, 2\psi_2) = -\frac{1}{3}\left(\frac{\psi_1 - \psi_2}{\psi_1 + \psi_2}\right). \quad (14)$$

Figure 8 shows the graph of $\hat{z}_B$ against z for Beta(2,3) and the residual plotting of $\hat{z}_B$ - z against z. The effect of symmetrization regarding $\hat{z}_B$ is good. The degree of the difference $\left|\Phi(\hat{z}_B) - \Phi(z)\right|$ is less than 0.0185 for $x_{0.001} \le x \le x_{0.999}$.

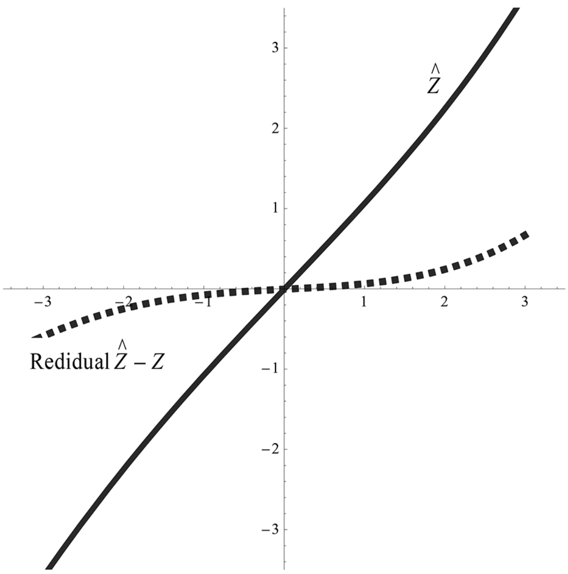

Figure 8. An approximation function $\hat{z}_B$ (solid line) for Beta(2,3) and the residual $\hat{z}_B - z$ (dashed line).

As expected, the result of Figure 8 is the same as that of Figure 5. The power transformation $\hat{z}_B$ for Beta(ψ_1, ψ_2) has the same performance as the power transformation $\hat{z}_F$ for $F_{2\psi_2}^{2\psi_1}$. For other useful applications of the power transformation for $F_{\theta_2}^{\theta_1}$, see Isogai (1999, 2001, 2005).

Ⅳ. Approximation method based on $\dot{z}$

In this section we shall use an $R(x)$ function to examine what type of functions approximate $\dot{z}$. Figures 3 and 6 suggest that there is a possibility of approximating $R(x)$ by the following function

$$\eta(x) = \beta_\theta \left[h(x)\right]^{\gamma(\theta)}, \beta_\theta > 0, h(x) > 0, \tag{15}$$

where β_θ and $\gamma(\theta)$ are functions of θ only, and $h(x)$ is a given function of x only.

Figure 1, including Figure 5 in Isogai (2014), also indicates that the same type of an approximation formula as above is useful, but it needs one more additional term in what follows:

$$\eta(x) = R(x_{0.5}) + \beta_\theta \left[h(x)\right]^{\gamma(\theta)}, \tag{16}$$

where $h(x)\,(\geq 0)$ is symmetric about $x_{0.5}$ and $h(x_{0.5}) = 0$.

As for choices of $h(x)$, we can use $h(x) = x$ in Figure 3, $h(x) = |x|$ in Figure 1, $h(x) = x(1-x)$ in Figure 6 and $h(x) = (1-x)(1+x)$ in Figure 8 of Isogai (2014).

1. Estimation of functions β_θ and $\gamma(\theta)$

First we shall consider approximating $R(x)$ by $\eta(x)$ of (15). Suppose that $R(x) = \eta(x)$. Then we have

$$\frac{1}{\dot{z}} = \beta_\theta \left[h(x)\right]^{\gamma(\theta)}. \tag{17}$$

After taking logarithms of both sides of (17), differentiate it with respect to x. We have

$$\gamma(\theta) = -\frac{\ddot{z}}{\dot{z}}\frac{h(x)}{\dot{h}(x)} = -\left(\frac{\partial}{\partial x}\log f_\theta(x) + z\dot{z}\right)\frac{h(x)}{\dot{h}(x)}, \tag{18}$$

where we have used (2) to derive the last term.

Substitute $x = x_\alpha$ into the last term of (18), where x_α is the 100α percentile of $F_\theta(x)$. Then we obtain an estimator of $\gamma(\theta)$:

$$\hat{\gamma}(\theta) = -\left(\frac{\partial}{\partial x}\log f_\theta(x)\bigg|_{x=x_\alpha} + z(x_\alpha)\dot{z}(x_\alpha)\right)\frac{h(x_\alpha)}{\dot{h}(x_\alpha)}. \tag{19}$$

Using $\hat{\gamma}(\theta)$ of (19), from (17) we also obtain an estimator of β_θ :

$$\hat{\beta}_\theta = \frac{1}{\dot{z}(x_\alpha)\left[h(x_\alpha)\right]^{\hat{\gamma}(\theta)}}. \tag{20}$$

If we use the median $x_{0.5}$ for $x = x_\alpha$ in (19) and (20), we have simpler estimators

$$\hat{\gamma}(\theta) = -\left(\frac{\partial}{\partial x} \log f_\theta(x) \Big|_{x=x_{0.5}} \right) \frac{h(x_{0.5})}{\dot{h}(x_{0.5})}, \tag{21}$$

$$\hat{\beta}_\theta = \frac{\phi(0)}{f_\theta(x_{0.5})\left[h(x_{0.5})\right]^{\hat{\gamma}(\theta)}}. \tag{22}$$

Similarly, we shall consider approximating $R(x)$ by $\eta(x)$ of (16). We have the following estimators of $\gamma(\theta)$ and β_θ :

$$\hat{\gamma}(\theta) = -\frac{\frac{\partial}{\partial x} \log f_\theta(x) \Big|_{x=x_\alpha} + z(x_\alpha)\dot{z}(x_\alpha)}{1 - \dot{z}(x_\alpha) R(x_{0.5})} \left(\frac{h(x_\alpha)}{\dot{h}(x_\alpha)} \right), \tag{23}$$

$$\hat{\beta}_\theta = \frac{R(x_\alpha) - R(x_{0.5})}{\left[h(x_\alpha)\right]^{\hat{\gamma}(\theta)}}, \tag{24}$$

where $\alpha \neq 0.5$. In the following Example 3, we shall set $\alpha = 0.95$ and use $x_{0.95}$ as x_α in (23) and (24).

Using the estimated $\hat{\beta}_\theta$ and $\hat{\gamma}(\theta)$, approximating formulas for the transformation $z = T(x)$ are written as

$$\hat{z} = \hat{T}(x) = \int_{x_{0.5}}^{x} \frac{du}{\hat{\eta}(u)}$$

$$= \frac{1}{\hat{\beta}_\theta} \int_{x_{0.5}}^{x} \left[h(u)\right]^{-\hat{\gamma}(\theta)} du \quad \text{(from (15))} \tag{25}$$

$$= \int_{x_{0.5}}^{x} \frac{du}{R(x_{0.5}) + \hat{\beta}_\theta \left[h(u)\right]^{\hat{\gamma}(\theta)}} \quad \text{(from (16))}. \tag{26}$$

2. Performance of approximation formulas

To check the adequacy of the assumption $R(x) = \eta(x)$ about the model (15), we shall examine the ratio $R(x)/\eta(x)$ through a plotting

$$\left(z, \frac{R(x)}{\hat{\beta}_\theta \left[h(x)\right]^{\hat{\gamma}(\theta)}}\right).$$

A constant tendency around one of this plotting ensures our assumption. We shall denote the function $R(x)/\hat{\eta}(x)$ by $B(x)$ in the following figures.

Example 1. Beta distributions

We shall consider Beta(ψ_1, ψ_2), a family of Beta distributions. For Beta distributions we suppose that $h(x) = x(1-x)$. Figure 9 shows $B(x)$ plots for Beta(2,3), Beta(6,9) and Beta(20,30). The degree of a constant tendency in $B(x)$ plots increases as the parameters ψ_1 and ψ_2 become large.

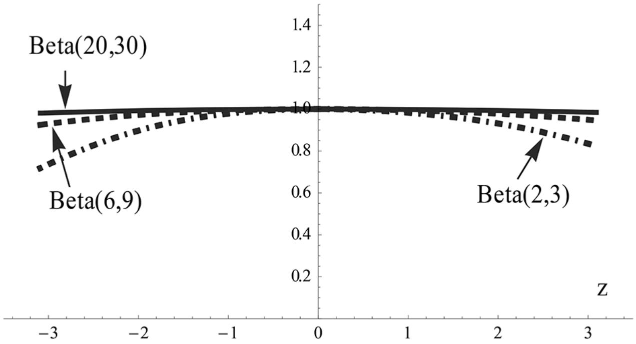

Figure 9. $B(x)\left(=R(x)/\hat{\eta}(x)\right)$ plots for Beta distributions, Beta(2,3) (dashed-dotted line), Beta(6,9) (dashed line), and Beta(20,30) (solid line).

To examine the degree of our approximation, we checked a difference between $\hat{z}$ of (25) and z. Figure 10 shows their residuals $\hat{z}-z$ for Beta(2,3), Beta(6,9) and Beta(20,30). The degree of the difference $\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right|$ over $x_{0.001} \le x \le x_{0.999}$ is as follows. For Beta(2,3), $\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right| \le 0.004$. For Beta(6,9), $\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right| \le 0.0015$. For Beta(20,30), $\left|\Phi\left(\hat{z}\right)-\Phi\left(z\right)\right| \le 0.0004$. Our approximation is "excellent".

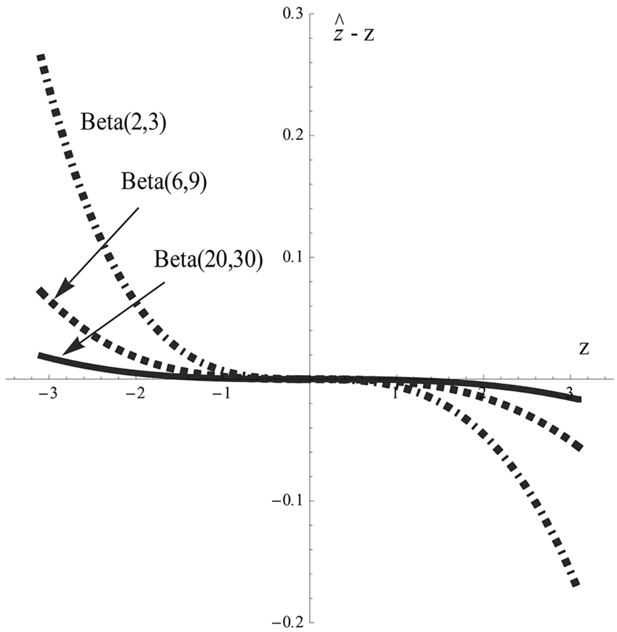

Figure 10. Residual plots of $\hat{z}-z$ for Beta distributions, Beta(2,3) (dashed-dotted line), Beta(6,9) (dashed line), and Beta(20,30) (solid line).

Example 2. Distributions of sample correlation coefficients

For a random sample of size n from a bivariate normal distribution, the density of the sample correlation coefficient $\hat{\rho}$ (= X, say) is given by

$$f_\rho(x) = \frac{(n-2)\left(1-\rho^2\right)^{(n-1)/2}\left(1-x^2\right)^{(n-4)/2}}{\sqrt{2}\,(n-1)\,B\left(\frac{1}{2},\, n-\frac{1}{2}\right)(1-\rho x)^{n-3/2}}\ {}_2F_1\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2};\frac{1+\rho x}{2}\right), \tag{27}$$

where ρ $(-1<\rho<1)$ is the population correlation coefficient, and ${}_2F_1(\cdots)$ is the Gauss hypergeometric function (see Johnson, Kotz and Balakrishnan (1995), p.549).

Figure 11 displays the solution curves $z = T(x)$ and $R(x)$ plots for the distributions of sample correlation coefficients with $\rho = 0.1$ and $\rho = 0.9$ in case of a sample size $n = 5$.

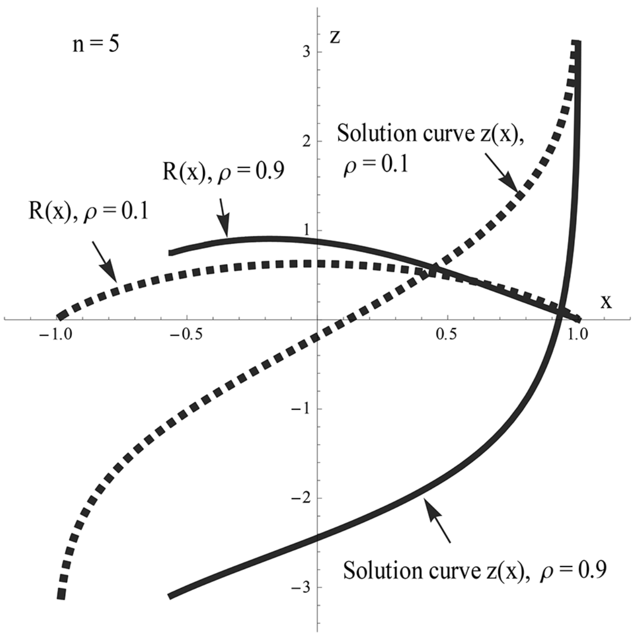

Figure 11. Solution curves $z = T(x)$ and $R(x)$ plots for the distributions of correlation coefficients with $\rho = 0.1$ and $\rho = 0.9$ for $n = 5$.

For distributions of sample correlation coefficients we suppose that $h(x) = (1-x)(1+x)$. Figure 12 shows $B(x)$ plots for the distributions of sample correlation coefficients with $\rho = 0.1$ in cases of sample sizes $n = 5$ and $n = 15$. A constant tendency of a $B(x)$ plot increases as a sample size becomes large.

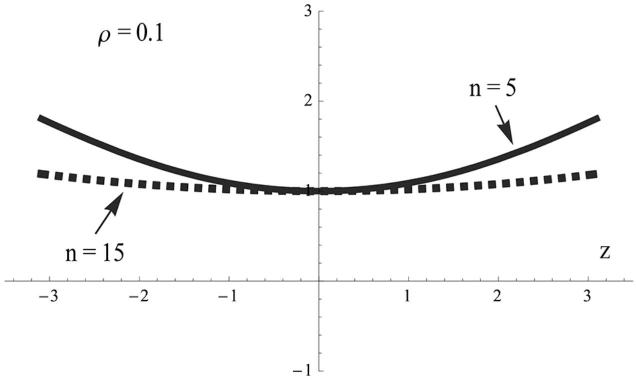

Figure 12. $B(x)\left(=R(x)/\hat{\eta}(x)\right)$ plots for the distributions of sample correlation coefficients with $\rho = 0.1$ in cases of sample sizes n = 5 and n = 15.

Figure 13 displays the difference between $\hat{z}$ of (25) and z for the distributions of sample correlation coefficients with $\rho = 0.1$ in cases of sample sizes n = 5 and n = 15. The accuracy of our approximation increases as a sample size becomes large. The degree of the difference $\left|\Phi(\hat{z})-\Phi(z)\right|$ over $x_{0.001} \le x \le x_{0.999}$ is as follows. $\left|\Phi(\hat{z})-\Phi(z)\right| \le 0.013$ for n = 5, and $\left|\Phi(\hat{z})-\Phi(z)\right| \le 0.003$ for n = 15. Our approximation is fairly "good".

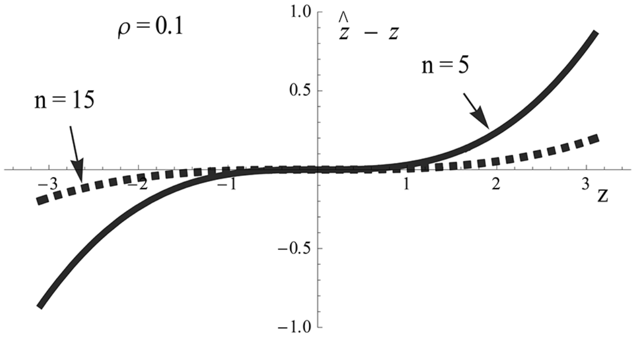

Figure 13. Residual plots of $\hat{z}-z$ for the distributions of sample correlation coefficients with $\rho = 0.1$ for sample sizes n = 5 (solid line) and n = 15 (dashed line).

Similarly, we give $B(x)$ plots (Figure 14) and residual plots (Figure 15) for the distributions of sample correlation coefficients with $\rho = 0.9$ in cases of n = 5 and n = 15.

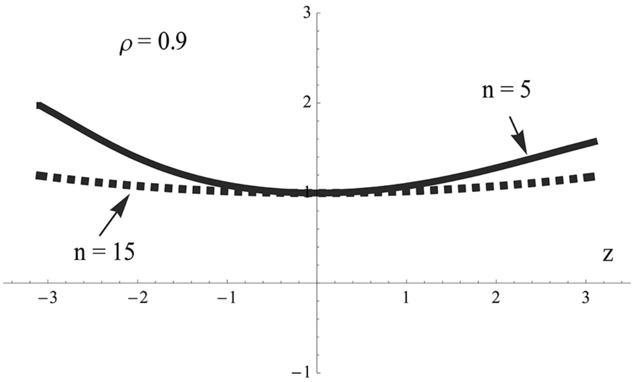

Figure 14. $B(x)\left(=R(x)/\hat{\eta}(x)\right)$ plots for the distributions of sample correlation coefficients with $\rho = 0.9$ in cases of sample sizes n = 5 and n = 15.

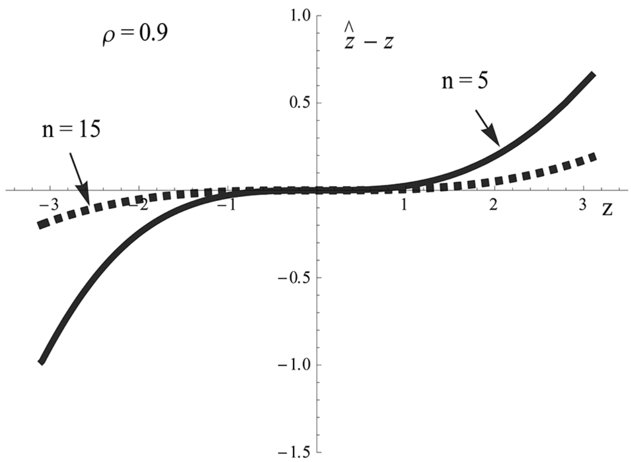

Figure 15. Residual plots of $\hat{z} - z$ for the distributions of sample correlation coefficients with $\rho = 0.9$ for sample sizes n = 5 (solid line) and n = 15 (dashed line).

The degree of the difference $\left|\Phi(\hat{z}) - \Phi(z)\right|$ over $x_{0.001} \le x \le x_{0.999}$ is as follows. $\left|\Phi(\hat{z}) - \Phi(z)\right| \le 0.013$ for n = 5, and $\left|\Phi(\hat{z}) - \Phi(z)\right| \le 0.003$ for n = 15. Our approximation is still fairly “good”.

Example 3. Student’s t distributions

Finally we shall consider a typical symmetric non-normal distribution, that is, Student’s t distribution denoted by $t(\theta)$, where θ (>0) means the degrees of freedom. The density of $t(\theta)$ is given by

$$f_{\theta}(x)=\frac{1}{\sqrt{\theta\pi}}\frac{\Gamma\left(\frac{\theta+1}{2}\right)}{\Gamma\left(\frac{\theta}{2}\right)}\left(1+\frac{x^2}{\theta}\right)^{-\frac{\theta+1}{2}}. \tag{28}$$

Note that $t(1)$ is the standard Cauchy distribution.

For the approximation formula (16) we can examine the ratio $R(x)/\eta(x)$ through a plotting

$$\left(z, \frac{R(x)}{R(x_{0.5})+\hat{\beta}_{\theta}\left[h(x)\right]^{\hat{\gamma}(\theta)}}\right).$$

A constant tendency around one of this plotting ensures our assumption $R(x)=\eta(x)$. We shall denote the function $R(x)/\hat{\eta}(x)$ by $B(x)$ in the following figures. We suppose that $h(x)=|x|$.

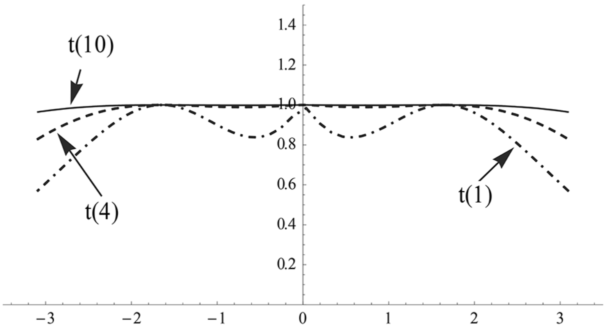

Figure 16. $B(x)\left(=R(x)/\hat{\eta}(x)\right)$ plots for Student's t distributions $t(1)$ (dashed-dotted line), $t(4)$ (dashed line), and $t(10)$ (solid line).

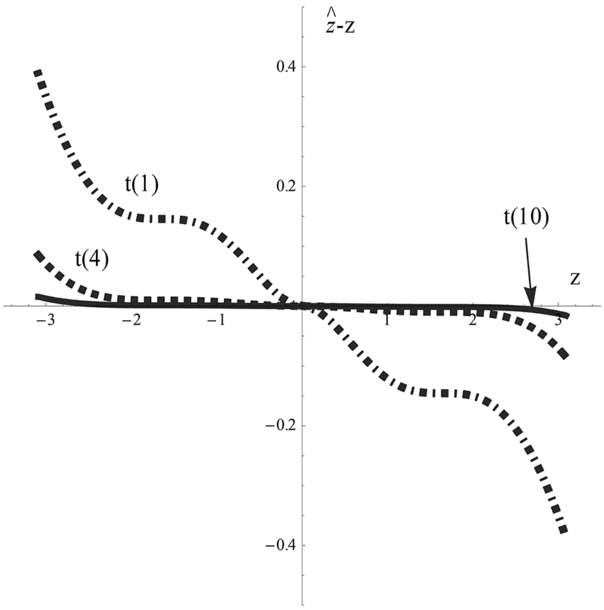

Figure 17. Residual plots of $\hat{z}-z$ for Student's t distributions, $t(1)$ (dashed-dotted line), $t(4)$ (dashed line), and $t(10)$ (solid line).

Figure 16 shows B(x) plots for t(1), t(4) and t(10). A constant tendency of a $B(x)$ plot increases as the parameter θ becomes large. Figure 17 displays the difference between $\hat{z}$ of (26) and z for t(1), t(4) and t(10). The degree of the difference $\left|\Phi(\hat{z})-\Phi(z)\right|$ over $x_{0.001} \le x \le x_{0.999}$ is as follows. For t(1), $\left|\Phi(\hat{z})-\Phi(z)\right| \le 0.035$. For t(4), $\left|\Phi(\hat{z})-\Phi(z)\right| \le 0.002$. For t(10), $\left|\Phi(\hat{z})-\Phi(z)\right| \le 0.0003$. When the degree of freedom θ is more than four, our approximation seems "excellent".

V. Conclusions

In the present paper we have examined the following several approximation methods for a normalizing transformation $z = T(x)$.

(I) A heuristic approach for the standard Cauchy distribution.

(II) A method based on a symmetrization principle, which was applied to F and Beta distributions.

(III) A method based on estimation of a function $R(x)$, which was applied to distributions of Beta, Student's t and sample correlation coefficients.

Our conclusions are as follows.

(1) Methods (I) and (III) have the good accuracy of approximations. As for Method (II), the accuracy of approximation is not so bad.

(2) Method (II) has great potential for applications.

(3) The idea of Method (III) is available for transformation problems from any distribution to another distribution except for normal.

Acknowledgement

The author would like to express his great thanks to Grant-in-Aid for Scientific Research (C) No.24500266 by JSPS.

References

[1] Efron, B. (1982), "*Transformation Theory: How normal is a family of distributions*", Annals of Statistics, Vol.10, pp. 323-339.

[2] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995), "*Continuous Univariate Distributions*", Vol.2, 2nd Edition, Wiley-Interscience.

[3] Isogai, T. (1999), "*Power transformation of the F distribution and a power normal family*", Journal of Applied Statistics, Vol.26, pp.351-367.

[4] Isogai, T. (2001), "*Applications of Power Transformation Formulas of the F distribution and a Power Normal Family*", Journal of the Japanese Society for Quality Control, Vol.31, pp.89-104. (in Japanese)

[5] Isogai, T. (2005), "*Applications of a New Power Normal Family*", Journal of Applied Statistics, Vol.32, No. 4, pp.421-436.

[6] Isogai, T. (2014), "*Examination of Transformations to Normality*", Journal of the University of Marketing and Distribution Sciences −Economics, Informatics & Policy Studies− , Vol.22, pp.87-99.

[7] Kaskay, G., Kolman, B., Krishnaiah, P. R. and Steinberg, L. (1980), "*Transformations to normality*", Handbook of Statistics, Vol.1, North-Holland, pp.321-341.

[8] Tarter, M. E. and Kowalski, C. J. (1972), "*A new test for and class of transformations to normality*", Technometrics, Vol.14, pp.735-744.

[9] Wilson, E. B. and Hilferty, M. M. (1931), "*The distributions of chi-square*", Proceedings of the National Academy of Sciences, Vol.17, pp.684-688.