SOME PROPERTIES OF BOX-COX ESTIMATES OF TRANSFORMING PARAMETERS
IN POWER-NORMAL DISTRIBUTIONS

HIROYUKI UESAKA AND MASASHI GOTO

SHIONOGI KAISEKI CENTER

22-4, 1-CHOME, SUITA, OSAKA 564, JAPAN

SUMMARY

Box-Cox method of parameter estimation in power-normal distribution is investigated. The estimate is shown to be inconsistent except the case of lognormal distribution corresponding to zero transforming parameter in the power-normal distribution. It is also shown that the asymptotic biases of the parameters of the power-normal distribution are determined by true transforming parameter and degree of truncation, and further are derived by those of truncated normal distribution corresponding to unity transforming parameter in the power-normal distribution. Numerical results on biases are illustrated briefly.

Key words: power-normal distribution, Box-Cox transformation,
asymptotic bias

1. Introduction

Many techniques of statistical analysis rest on the normal distribution, above all so do linear models for analysis of variation of observations. In the fields of medicine and biology, however, observations are often found to follow skewed distributions. Two ordinary approaches for analysis of the skewed distributions have been employed. The one is applying transformation to observations for the purpose of transformed observations to have a near-normal distribution. And the other is to adopt statistical models based on the skewed distributions.

In 1964, Box and Cox gave a major step toward a new development of analysis of the skewed distributions. They proposed maximum likelihood estimation of a suitable power transformation to normality.

The basic idea of transformation seems that if the original scale is related to a new scale by a sufficiently smooth one-to-one transformation, we can treat our problems on the new scale and the most convenient scale may be adopted. Such a relationship between the original and transformed scales induces a model on the original scale where we perceive our problem in the physical phenomena, from the model on the transformed scale where we really

analyse our data. Thus the two approaches may be related to two aspects of the same parametric model, one to the manifest model on the original scale and the other to the underlying one on the transformed scale. The power-normal distribution (abb. PND) is obtained by adopting (truncated) normal distribution as the basic distribution for the general model based on Box-Cox transformation family.

Some distributional properties and several applications of the PND have been given by several authors (Goto-Uesaka-Inoue (1979), Uesaka-Goto (1980a, b), (1981) Matsubara-Goto (1980), Goto-Inoue (1980)). The properties of estimates of the PND, however, have not been known satisfactorily. We will discuss asymptotic biases of the estimates obtained by Box-Cox method. We first give a few properties of the PND variate and then, in Section 3, discuss a problem of estimation where we show that the properties of the optimal transformation of any PND except log-normal distribution are closely related to those of a singly truncated normal distribution. In Section 4 we give a few numerical results on the optimal transformation.

2. Definition and some properties of the PND

2.1 Definition

Let us define $X = (0, \infty)$ and $\Lambda = (-\infty, \infty)$. The Box-Cox transformation of $x \in X$ with parameter λ is defined by

$$Z_{\lambda}(x) = \begin{cases} \frac{x^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \log x, & \lambda = 0 \end{cases} \quad (1)$$

The range of $Z_{\lambda}(x)$ is denoted by $\mathcal{J}_{\lambda}$, where

$$\mathcal{J}_{\lambda} = \begin{cases} (-1/\lambda, \infty), & \lambda > 0, \\ (-\infty, \infty), & \lambda = 0, \\ (-\infty, -1/\lambda), & \lambda < 0. \end{cases} \quad (2)$$

For any fixed λ , the transformation $Z_{\lambda}(x)$ maps X onto $\mathcal{J}_{\lambda}$ and is continuously differentiable with respect to x , and for any fixed x , $Z_{\lambda}(x)$ is continuously differentiable with respect to λ . The PND is defined as the distributions defined on X having probability density function (pdf)

$$g(x; \mu, \sigma^2, \lambda) = \begin{cases} \frac{x^{\lambda-1}}{\sqrt{2\pi} \sigma A(\lambda, \mu, \sigma^2)} \exp\left\{-\frac{(Z_{\lambda}(x) - \mu)^2}{2\sigma^2}\right\} & x > 0, \\ 0 & x \leq 0 \end{cases} \quad (3)$$

where

$$A(\lambda, \mu, \sigma^2) = \begin{cases} \Phi\left(\frac{\lambda\mu+1}{\lambda\sigma}\right) & \lambda > 0, \\ 1 & \lambda = 0, \\ \Phi\left(-\frac{\lambda\mu+1}{\lambda\sigma}\right) & \lambda < 0. \end{cases} \quad (4)$$

Let us denote truncated normal distribution defined on the interval (a, b) with location and scale parameter μ, σ by $TN(\mu, \sigma; a, b)$. Then if a random variable x follows $TN(\mu, \sigma; 0, \infty)$, x has a density $g(x; \mu, \sigma, 1)$.

2.2 Some properties of the PND

Several properties of the PND have been given in Uesaka-Goto (1980a, b), (1981) and Goto-Inoue (1980). Here we give a few properties which have not been given there.

Proposition 1. Suppose X has the pdf $g(x; \xi, \theta^2, \lambda)$, $\lambda \neq 0$, and Y has the pdf $TN(K, 1; 0, \infty)$, where

$$K = \frac{\lambda \xi + 1}{|\lambda| \theta}. \quad (5)$$

For any $\rho \neq 0$, let us denote the cumulative distribution function (cdf) of $Z_\rho(Y)$ by $G(z; K, \rho)$. Then the cdf of $Z_\rho(X)$ is

$$G\left(\frac{\lambda}{(|\lambda| \theta)^{\rho/\lambda}} Z_\rho(x) - Z_{-\rho/\lambda}(|\lambda| \theta)\right). \quad (6)$$

(See Appendix 1 for proof of the Proposition.)

Corollary 1. Suppose X and Y have pdf's $g(x; \xi, \theta^2, \lambda)$ and $g(y; -\xi, \theta^2, -\lambda)$, respectively. Then for any $p \neq 0$, X^p and Y^{-p} have the same distribution.

We denote a family of subsets of Λ by

$$I(\lambda) = \begin{cases} (-\lambda, \infty) & \lambda > 0, \\ (-\infty, \infty) & \lambda = 0, \\ (-\infty, -\lambda) & \lambda < 0. \end{cases} \quad (7)$$

Then we have next proposition.

Proposition 2. Let X be a random variable with pdf $g(x; \xi, \theta, \lambda)$ and $\lambda \neq 0$. For any $p \in I(\lambda)$ and for any $k=0, 1, 2, \dots$, we have

$$E(X^p (\log X)^k) = (\sqrt{2}|\lambda|\sigma)^{p/\lambda} \frac{1}{\sqrt{2}} \frac{\varphi(K)}{\Phi(K)} \frac{1}{(2\lambda)^k} \sum_{n=0}^{\infty} \frac{(\sqrt{2}K)^n}{n!} \sum_{j=0}^k C_j [2 \log(\sqrt{2}|\lambda|\sigma)]^{k-j} \Gamma^{(j)} \left(\frac{n+1}{2} + \frac{p}{2\lambda} \right), \quad (8)$$

where $K = (\xi\lambda+1)/|\lambda|\theta$ and $\Gamma^{(j)}(z)$ is the j -th derivative of gamma function $\Gamma(z)$. If $p \notin I(\lambda)$, then we have $E(X^p) = +\infty$.

Proposition 3. When $\lambda=0$, $g(x; \xi, \theta^2, \lambda)$ becomes the pdf of a log-normal distribution and

$$E X^p (\log X)^k = \exp(p\xi + \frac{p^2\theta^2}{2}) \sum_{n=0}^{[k/2]} C_{2n} (\xi + p\theta^2)^{k-2n} (2n-1)!! \theta^{2n}, \quad (9)$$

where $[k/2]$ denote the greatest integer not exceeding $k/2$, and

$$(2n-1)!! = (2n-1)(2n-3) \dots 3 \cdot 1. \quad (10)$$

From Proposition 2, it follows that for any $\lambda > 0$, any moment of positive order exists and the moments standardized by the standard deviation do depend on ξ and θ only through λ and K .

3. Parameter estimation

Box and Cox (1964) proposed to estimate the parameters ξ , θ^2 and λ by maximizing the likelihood function derived on the assumption that the transformed variate $Z_\rho(X)$ has a normal distribution. Since then, the problem of parameter estimation in power transformation has been studied by many authors (Draper-Cox (1969), Andrews (1971), Carroll (1980), Hernandez-Johnson (1980), Bickel-Doksum (1981)). Hernandez and Johnson (1980) discussed the asymptotic behavior of the Box-Cox estimates and the distribution of the transformed variate in four cases of original distributions.

In this paper we investigate the optimal transformation of the power-normal observations. The term "optimal" means that the transformation is determined by an asymptotic value of the transforming parameter obtained by Box-Cox method.

Suppose that X_i , $i=1, \dots, n$ are independently and identically distributed (iid) with pdf $g(x; \xi, \theta^2, \lambda)$, where ξ , θ and λ are unknown parameters. In Box-Cox method, we assume that $Z_\rho(X_i)$, $i=1, \dots, n$ are iid with density $N(\mu, \sigma^2)$. Now we write $X_n = (X_1, \dots, X_n)'$ and denote the log-likelihood of Box-Cox by $L(\mu, \sigma^2, \lambda | X_n)$. Then we have

$$\begin{aligned} L(\mu, \sigma^2, \rho | X_n) &= (\rho-1) \sum_{j=1}^n \log X_i - \frac{1}{2\sigma^2} \sum_{j=1}^n (Z_\rho(X_i) - \mu)^2 \\ &\quad - \frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2. \end{aligned} \quad (11)$$

This is obtained by deleting the term $n \log A(\rho, \xi, \sigma^2)$ from the full likelihood of PND. Taking derivatives with respect to μ , σ^2 and ρ on $L(\mu, \sigma^2, \rho | X_n)$, we have

$$\begin{aligned} \frac{\partial L}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (Z_\rho(X_i) - \mu), \\ \frac{\partial L}{\partial \sigma^2} &= -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^n (Z_\rho(X_i) - \mu)^2, \end{aligned} \quad (12)$$

and

$$\frac{\partial L}{\partial \mu} = \sum_{i=1}^n \log X_i - \frac{1}{\sigma^2} \sum_{j=1}^n (Z_\rho(X_i) - \mu) \frac{\partial Z_\rho(X_i)}{\partial \rho}.$$

Let us write

$$\hat{\mu}_n(\rho) = \frac{1}{n} \sum_{i=1}^n Z_\rho(X_i), \quad (13)$$

and

$$\hat{\sigma}_n^2(\rho) = \frac{1}{n} \sum_{i=1}^n (Z_\rho(X_i) - \hat{\mu}_n(\rho))^2. \quad (14)$$

These are the maximum likelihood (ML) estimates of μ and σ^2 when ρ is known. If ρ is unknown, the ML estimate of ρ is a solution to equation

$$\sum_{i=1}^n \log X_i - \frac{1}{\hat{\sigma}_n^2(\rho)} \sum_{i=1}^n (Z_\rho(X_i) - \hat{\mu}_n(\rho)) \frac{\partial Z_\rho(X_i)}{\partial \rho} = 0, \quad (15)$$

and is denoted by $\hat{\rho}_n$. Corresponding to the estimate $\hat{\rho}_n$, we denote the ML estimates of μ and σ^2 by $\hat{\mu}_n$ and $\hat{\sigma}_n^2$, respectively.

Then we have

$$\hat{\mu}_n = \hat{\mu}_n(\hat{\rho}_n) \text{ and } \hat{\sigma}_n^2 = \hat{\sigma}_n^2(\hat{\rho}_n). \quad (16)$$

Now let us evaluate the asymptotic value of $\hat{\mu}_n$, $\hat{\sigma}_n^2$ and $\hat{\rho}_n$ when n tends to infinity. If $\rho \in I(\lambda/2)$, the second order moment of $Z_\rho(X)$ exists, so that $\hat{\mu}_n(\rho)$ and $\hat{\sigma}_n^2(\rho)$ converge in probability to

$$\mu(\rho) = \text{Eg } Z_\rho(X) \text{ and } \sigma^2(\rho) = \text{Eg}(Z_\rho(X) - \mu(\rho))^2, \quad (17)$$

respectively.

Let us write

$$\begin{aligned}
 L_n(\rho | X_n) &= \max_{\mu, \sigma} \frac{1}{n} \ln(\mu, \sigma^2, \rho | X_n) \\
 &= (\rho - 1) \frac{1}{n} \sum_{i=1}^n \log x_i - \frac{1}{2} \log \hat{\sigma}_n^2(\rho) \\
 &\quad - \frac{1}{2} - \frac{1}{2} \log 2\pi
 \end{aligned} \tag{18}$$

and

$$L(\rho; \xi, \theta^2, \lambda) = (\rho - 1) E_g \log x - \frac{1}{2} \log \sigma^2(\rho) - \frac{1}{2} (1 + \log 2\pi), \tag{19}$$

where $E_g(\cdot)$ denote the expectation under g . If $\rho \in I(\lambda/2)$, $L_n(\rho | X_n)$ converges in probability to $L(\rho; \xi, \theta^2, \lambda)$.

When $\lambda=0$, we have

$$L(\rho; \xi, \theta^2, \lambda) = \begin{cases} -\frac{\theta^2 \rho}{2} - \frac{1}{2} \log \left(\frac{e^{\theta^2 \rho^2} - 1}{\rho^2} \right) - \frac{1}{2} (1 + \log 2\pi) & \text{for } \rho \neq 0, \\ -\frac{1}{2} \log \theta^2 - \frac{1}{2} (1 + \log 2\pi) & \text{for } \rho = 0. \end{cases} \tag{20}$$

For any λ , that $L(\rho; \xi, \theta^2, \lambda)$ is bounded above is shown in the same manner as the proof of boundedness of Kullback-Leibler information number (Kullback (1959, p.15)). If $L(\rho; \xi, \theta^2, \lambda)$ has the maximum

value in a finite range of ρ , the solution to the equation $\hat{\rho}_n$ converges in probability to ρ^* , the value of ρ which maximize $L(\rho; \xi, \sigma^2, \lambda)$. Then $\hat{\mu}_n(\hat{\rho}_n)$ and $\hat{\sigma}_n^2(\hat{\rho}_n)$ converge in probability to $\mu(\rho^*)$ and $\sigma^2(\rho^*)$, respectively. We write $\mu^* = \mu(\rho^*)$ and $\sigma^{*2} = \sigma^2(\rho^*)$, these are the asymptotic value of the estimated location and scale of the PND. When $\lambda=0$, it follows, from (20) that $\rho^* = \lambda, \mu^* = \xi$ and $\sigma^* = \theta$. For any $\lambda \neq 0$, $\rho^* \neq \lambda, \mu^* \neq \xi$ and $\sigma^* \neq \theta$. As it seems difficult to evaluate the biases of ρ^*, μ^* and σ^{*2} , we will numerically investigate the biases in Section 4.

Here we give important results on the optimal transformation.

Proposition 4. The function of $\rho; L(\rho; \xi, \theta, \lambda)$ depends only on λ and $K = (\xi\lambda + 1)/|\lambda|\theta$.

This is shown by expanding $L(\rho; \xi, \theta, \lambda)$ as

$$\begin{aligned} L(\rho; \xi, \theta, \lambda) = & \frac{\rho}{2\lambda} \psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma^{(1)}\left(\frac{n+1}{2}\right) + \frac{1}{2} \log \rho^2 \\ & - \frac{1}{2} \log \left[\psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma\left(\frac{n+1}{2} + \frac{\rho}{\lambda}\right) \right. \\ & \left. - \left\{ \psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma\left(\frac{n+1}{2} + \frac{\rho}{2\lambda}\right) \right\}^2 \right] \\ & - \text{Eg}(\log X) + \frac{1}{2} \log 2, \end{aligned} \quad (21)$$

where

$$\psi(K) = \frac{1}{\sqrt{2}} \frac{\varphi(K)}{\Phi(K)} \text{ and } \varphi(K) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{K^2}{2}\right).$$

From this proposition we can observe that for any given λ , the bias of ρ^* depends only on K , in other words degree of truncation $1-\Phi(K)$. Thus the ρ^* associated with the pdf $g(x; \xi, \theta^2, \lambda)$ may be written as $\rho_K^*(\lambda)$. Then we have next proposition.

Proposition 5. For any $\lambda \neq 0$, $\rho_K^*(\lambda) = \lambda \rho_K^*(1)$. (See Appendix 2 for proof of the Proposition.)

If we simply write $\rho_K^* = \rho_K^*(1)$, the optimal transforming parameter of $TN(K, 1; 0, \infty)$, $\rho_K^*(\lambda) = \lambda \rho_K^*$.

This proposition states that the optimal transformation of any PND except for log-normal distribution is determined by the one for the normal distribution singly truncated from below at zero. Combining Propositions 1 and 5, we have next proposition.

Proposition 6. Suppose random variables X and Y have pdf's $g(x; \xi, \theta^2, \lambda)$ and $TN(K, 1; 0, \infty)$, respectively, and $K = (\lambda\xi + 1)/|\lambda|\theta$. Further we denote the optimal transforming parameter of X and Y by ρ^* and ρ_K^* , respectively. If $U = Z_{\rho^*}(X)$ has cdf $F(u)$ and $V = Z_{\rho_K^*}(Y)$ has cdf $G(v)$, then we have

$$1) \rho^* = \lambda \rho_K^*,$$

$$2) F(u) = G\left(\frac{\lambda}{(|\lambda|\theta)\rho_K^*} u - Z_{-\rho_K^*}(|\lambda|\theta)\right).$$

Thus for any $\lambda \neq 0$, all properties of optimally transformed variables of the power-normal variate are determined by the optimal transformation of $TN(K, 1; 0, \infty)$.

3.3 Illustration of biases in Box-Cox estimation

Here we give a typical example which shows the bias of optimal transformation. Let Y be a standard normal deviate. Then $X=Y^2$ has the chi-square distribution with unit degree of freedom, and $Z_{0.5}(x)=2(\sqrt{x}-1)$ has $TN(-2, 4, -2, \infty)$, whose degree of truncation is 0.5. Thus X has the pdf $g(x; -2, 4, 0.5)$. The optimal transforming parameter is approximately 0.2, which is greatly different from the true value 0.5. We have

$$\mu(0.2) = 5(\sqrt{2} \Gamma(0.7)/\sqrt{\pi} - 1) = -0.386$$

and

$$\sigma^2(0.2) = 5^2 \cdot 2^{2/5} (\Gamma(0.9)/\sqrt{\pi} - \Gamma(0.7)^2/\sqrt{\pi}) = 2.097.$$

The skewness γ_1^* and kurtosis γ_2^* are $\gamma_1^* = -0.134$ and $\gamma_2^* = 2.053$. Thus for a sufficiently large sample, Box-Cox method give $g(x; -0.386, 2.097, 0.2)$ as the fitted PND. Figure 1 shows the two pdf's, $g(x; -0.386, 2.097, 0.2)$ and $g(x; -2, 4, 0.5)$, that is, the pdf of the chi-square with unit degree of freedom.

4. Numerical evaluation of biases of the optimal transformation

As we saw in the preceding section, the optimal transformation of PND and its properties are determined by those of the corresponding truncated normal distribution $TN(K, 1; 0, \infty)$. We therefore investigate ρ_K^* and moments of transformed variate of $TN(K, 1; 0, \infty)$'s. Table 1 gives the transforming parameter, location, scale, skewness, kurtosis of the optimal transformation,

$$K^* = (\rho_K^* \mu^{*+1}) / \rho_K^* \sigma^* \text{ and } A^* = \Phi(K^*),$$

for a number of cases of K 's between 0 and 5.0.

The results shows that as the degree of truncation gets large the transforming parameter tends to zero. The skewness is negative but small and less than 0.12 in absolute values. The kurtosis are between 2.5 and 3.0, thus the distributions are somewhat platykurtic. Further degrees of truncation of the estimated PND are less than 1.5%.

The explanation of these facts may be that Box-Cox method give the distribution nearest to a full normal distribution by the criterion of Kullback-Leibler information number, and consequently the degree of truncation is reduced as far as possible to attain symmetry of the distribution.

In order to check the degree of approximation of the estimated PND to the true truncated normal distribution, we give pdf's and

cdf's of the two distributions for $K=2.5, 2, 1$ and 0 in the Figure 2.

Difference between the two distributions is clear when $K \leq 2$, and remarkable at the range of x where the values of cdf are over 0.8 .

5. Concluding remarks

In the application of Box-Cox transformation in data analyses, we intend not only analysis of transformed variables but also describing variation of observations by fitting distributions to data over whole range of the original scale. The likelihood treated by Box-Cox (Box and Cox (1964), equation (5)) is, however, incomplete and does not define any proper distribution of x . By completing the likelihood, we get a distribution with pdf on $\mathcal{X}$ defined by the equation (3) which is referred to as PND. In the present paper, we obtained some results on the properties of Box-Cox transformation of PND and its Box-Cox estimation and the biases of the estimated parameters by Box-Cox method is numerically investigated. These results are not enough for PND to be established as a useful distribution for data analysis and further examinations should be required.

Figure 1 Pdf's and cdf's of chi-square distribution with unit degree of freedom and fitted PND with $\mu^* = -0.386$, $\sigma^{*2} = 2.097$ and $\rho^* = 0.2$

$G(X|1.999, 0.882, 0.877) ; TN(2.5, 1)$

$G(X|0.819, 0.821, 0.777) ; TN(2, 1)$

$K=2.5$

$K=2.0$

$G(X|0.159, 0.750, 0.581) ; TN(1, 1)$

$G(X|-0.990, 0.799, 0.417) ; TN(0, 1)$

$K=1.0$

$K=0.0$

Figure 2 Truncated normal distributions and the fitted PND

Table 1

Asymptotic biases of optimal transformations of PND
with $\lambda=1$, $\theta=1$ and $\xi=k-1$.

K	ρ_K^*	ξ	μ^*	σ^{*2}	γ_1^*	γ_2^*	K^*	A^*
5.00	1.000	4.000	4.000	1.000	0.000	3.000	5.000	1.000
4.50	0.999	3.500	3.497	0.998	-0.000	2.999	4.501	1.000
4.00	0.996	3.000	2.989	0.994	-0.001	2.994	4.005	1.000
3.50	0.982	2.500	2.465	0.977	-0.007	2.978	3.524	1.000
3.00	0.946	2.000	1.927	0.939	-0.022	2.935	3.079	0.999
2.50	0.877	1.500	1.399	0.882	-0.049	2.856	2.705	0.997
2.00	0.777	1.000	0.919	0.821	-0.082	2.746	2.436	0.993
1.50	0.661	0.500	0.506	0.774	-0.113	2.635	2.294	0.989
1.00	0.561	0.0	0.159	0.750	-0.115	2.554	2.242	0.988
0.50	0.478	-0.500	-0.137	0.739	-0.107	2.513	2.275	0.989
0.30	0.451	-0.700	-0.243	0.736	-0.103	2.515	2.301	0.989
0.25	0.445	-0.750	-0.268	0.736	-0.102	2.508	2.307	0.989
0.10	0.427	-0.900	-0.342	0.734	-0.100	2.509	2.333	0.990
0.05	0.422	-0.950	-0.366	0.734	-0.097	2.510	2.339	0.990
0.0	0.417	-1.000	-0.390	0.733	-0.096	2.510	2.346	0.991

References

- Andrews, D.F. (1971). A note on the selection of data transformations. *Biometrika*, 58, 249 ~ 254.
- Bickel, P.J. and Doksum, K.A. (1981). An analysis of transformations revisited. *J. Amer. Statist. Assoc.*, 76, 296 ~ 311.
- Box, G.P.E. and Cox, D.R. (1964). An analysis of transformations. *J. Roy. Statist. Soc., B* 26, 211 ~ 252.
- Carroll, R.T. (1980). A robust method for testing transformations to achieve approximate normality. *J. Roy. Statist. Soc., B* 42, 71 ~ 78.
- Draper, N.R. and Cox, D.R. (1969). On distributions and their transformation to normality, *J. Roy. Statist. Soc., B* 31, 472 ~ 476.
- Goto, M., Uesaka, H. and Inoue, T. (1979). Some linear models for power transformed data. Invited presentation at the 10th International Biometric Conference at Brazil August, 6 ~ 10. (Res. Rep. No. 93, Res. Instit. Fund. Infor. Soc., Kyushu University)
- Goto, M. and Inoue, T. (1980). Some properties of power normal distribution. *Bulletin of Biometric Society of Japan*, 1, 28 ~ 54.

- Hernandez, F. and Johnson, R.A. (1980). The large-sample behavior of transformation to normality. J. Amer. Statist. Assoc., 75, 855 ~ 861.
- Matsubara, Y. and Goto, M. (1980). A censored regression analysis of survival time data based on power-normal distribution. Japanese Journal of Applied Statistics, 9, 43 ~ 51.
- Uesaka, H. and Goto, M. (1980a). An analysis of clinical chemistry data based on power-normal distribution. Japanese Journal of Applied Statistics, 9, 23 ~ 33.
- Uesaka, H. and Goto, M. (1980b). Statistical analysis of clinical chemistry data based on power-normal distribution. MEDINFO 80, Proceedings of the Third World Conference on Medical Informatics. eds, Lindberg and Kaihara, September 29 ~ October 4, 1980, Tokyo. 1014 ~ 1018, North-Holland, Amsterdam.
- Uesaka, H. and Goto, M. (1981). Power-normal distribution and its applications. Abstract of International Symposium on Computational Statistics in Fukuoka, October 19, 1981.

Appendix 1 Proof of the proposition 1

As X has pdf $g(x; \xi, \theta^2, \lambda)$, $Z_\lambda(X)$ has pdf $TN(\xi, \theta^2; -1/\lambda, \infty)$ for $X > 0$, and $TN(\xi, \theta^2; -\infty, -1/\lambda)$ for $\lambda < 0$.

Let us define

$$V = X^\lambda / |\lambda| \theta,$$

then V has the pdf $TN(K, 1; 0, \infty)$. For $\rho \neq 0$, $Z_\rho(X)$ is rewritten as

$$\begin{aligned} Z_\rho(X) &= \frac{1}{\lambda} \left\{ \frac{(X^\lambda)^{\rho/\lambda} - 1}{(\rho/\lambda)} \right\} \\ &= \frac{(|\lambda| \theta)^{\rho/\lambda}}{\lambda} \cdot Z_{\rho/\lambda}(V) + \frac{(|\lambda| \theta)^{\rho/\lambda} - 1}{\rho}. \end{aligned}$$

Thus we have

$$\begin{aligned} Z_{\rho/\lambda}(V) &= \frac{\lambda}{(|\lambda| \theta)^{\rho/\lambda}} \cdot Z_\rho(X) + \frac{(|\lambda| \theta)^{-\rho/\lambda} - 1}{\rho/\lambda} \\ &= \frac{\lambda}{(|\lambda| \theta)^{\rho/\lambda}} Z_\rho(X) - Z_{-\rho/\lambda}(|\lambda| \theta). \end{aligned}$$

Q.E.D.

Appendix 2 Proof of the proposition 5.

If we rewrite (21) by denoting $\tau = \rho/\lambda$, we have

$$\begin{aligned} L(\tau\lambda; \xi, \theta, \lambda) &= \frac{\tau}{2} \psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma\left(\frac{n+1}{2}\right) + \log \tau \\ &\quad - \frac{1}{2} \log \left[\psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma\left(\frac{n+1}{2} + \tau\right) \right. \\ &\quad \left. - \left\{ \psi(K) \sum_{n=0}^{\infty} \frac{(\sqrt{2K})^n}{n!} \Gamma\left(\frac{n+1}{2} + \frac{\tau}{2}\right) \right\}^2 \right] \\ &\quad + C_0, \end{aligned}$$

where

$$C_0 = \frac{1}{2} \log 2 + \log \lambda - E g(\log X).$$

This equation is independent of λ and C_0 is free from τ . Let us denote the value of τ which maximizes $L(\tau\lambda; \xi, \theta, \lambda)$ by τ^* . The above equation shows that τ^* is determined by K for any λ . This means that $\tau^* = \rho_K^*$. Thus we have $\rho_K^*(\lambda) = \lambda \tau^* = \lambda \rho_K^*$.

Q.E.D.