流通科学大学論集—経済・情報・政策編—第 26 巻第 2 号，35-59（2018）

Some Empirical Remarks on Using the Box-Cox Power Normal Family — Applications of a New Power Normal Family —

Takafumi Isogai*

We focus on a behavior of the maximum likelihood estimate $\hat{\lambda}$ for a power parameter λ of the Box-Cox power normal family. The m.l.e. $\hat{\lambda}$ is invariant under a common scale transformation of the relevant random variables, but in two-sample as well as multi-sample cases its scale-invariant property is broken under individual scale transformations of each random variable. In the present paper, by using the generalized F distribution and its approximate statistical model, that is, a new power normal family introduced by the author, we exhibit various examples to show unexpected behaviors of $\hat{\lambda}$ and other parameters' estimates under scale transformations.

Keywords: non-normality; bootstrap generalized information criteria; generalized F distribution; Box-Cox power normal family; new power normal family

I. Introduction

In practical fields, such as quality control and social sciences, we often come across non-normal data comprised of positive numbers, of which distributions are unimodal and positively or negatively skewed. A well-known statistical model to deal with such non-normal data is the Box-Cox power normal family (shortly, BC family) due to its easy handling (Box and Cox [1]). Recently, a new power normal family (shortly, NP family) has been introduced through the generalized F distribution, and is expected to have the same performance as BC family (Isogai [4]).

In regression analyses with respect to both families, our case studies (Isogai [4] and [8]) indicate that both families have the same performance in the meaning of goodness of fit, but there is a striking difference in estimation of power parameters, that is, estimates of power parameters have opposite signs. Furthermore, absolute values of estimates of power parameters with BC family are small. This fact suggests that an appropriate transformation is Logarithmic transformation under BC family. A similar situation often happens in a multi-sample case under BC family. Therefore, in the present paper, by using statistical properties of NP family, we shall investigate possible situations that cause a fluctuation of signs or small absolute values in estimation of power parameters with respect to BC family.

NP family has been constructed by Isogai [3] from the generalized F distribution (see Prentice [12])

*Faculty of Commerce University of Marketing and Distribution Science, 3-1 Gakuen Nishi-machi, Nishi-ku, Kobe city, Hyogo, 6512188, Japan

（2017 年 8 月 29 日受理）

through a Wilson-Hilferty type power transformation of an F random variable. A random variable X of the generalized F distribution is defined by

$$X = \eta\left(F_{2\phi_2}^{2\phi_1}\right)^{\gamma} \tag{1}$$

for $\eta > 0, \gamma > 0, \phi_1 > 1$ and $\phi_2 > 1$, where $F_{2\phi_2}^{2\phi_1}$ be the random variable of a central F distribution with degrees of freedom ($2\phi_1, 2\phi_2$) . The definition of NP family is that the following random variable

$$\frac{1}{\delta}\left\{\left(\frac{X}{\eta}\right)^{\frac{\delta}{\sigma}} - \exp\left(-\delta^2\right)\right\} \tag{2a}$$

or equivalently

$$\frac{1}{\delta}\left\{\exp\left(\delta\left(\frac{\log X - \mu}{\sigma}\right)\right) - \exp\left(-\delta^2\right)\right\} \tag{2b}$$

is approximately distributed as the standard normal variable for $\eta > 0, \sigma > 0, \mu = \log\eta$ and

$$-\frac{1}{3} < \delta < \frac{1}{3}, \tag{3}$$

where we put

$$\begin{cases} \delta = \delta_1\delta_2, \\ \sigma = \gamma\delta_2, \\ \delta_1 = \dfrac{1}{3}\left(\dfrac{1}{\phi_1} - \dfrac{1}{\phi_2}\right) \Big/ \left(\dfrac{1}{\phi_1} + \dfrac{1}{\phi_2}\right), \\ \delta_2 = \left(\dfrac{1}{\phi_1} + \dfrac{1}{\phi_2}\right)^{1/2}. \end{cases} \tag{4}$$

In the following, when we need not distinguish between our equivalent models (2a) and (2b), we call them simply NP family (2) and denote NP family (2) by NP$\left(\delta, \sigma, \eta\right)$.

As δ tends to zero, $\log X$ is nearly distributed as the normal distribution with mean μ (= $\log\eta$) and variance σ^2 . When δ is near $-1/3$, the distribution of X is approximated by the extreme value distribution of Type 2 (see Johnson, Kotz and Balakrishnan [9]):

$$\Pr\left(X \le x\right) = \exp\left\{-\left(\frac{x}{\eta}\right)^{-\frac{1}{\sigma}}\right\}$$

for $x \ge 0$. When δ is near $1/3$, the distribution of X is approximated by the Weibull distribution, of which distribution function is given by

$$\Pr\left(X \le x\right) = 1 - \exp\left\{-\left(\frac{x}{\eta}\right)^{\frac{1}{\sigma}}\right\}$$

for $x \geq 0$. Under a scale transformation of $X \to \theta X \ (\theta > 0)$, δ and σ of NP family (2) are invariant.

On the other hand, the Box-Cox power normal family (BC family) is defined by normality of the following random variable

$$\frac{1}{\tau}\left\{\frac{X^{\lambda}-1}{\lambda}-\xi\right\} \tag{5}$$

for $\tau > 0$ (for the details, see Box and Cox [1]). The random variable $\left(X^{\lambda}-1\right)/\lambda$ is supposed to be approximately distributed as normal with mean ξ and variance τ^2. We denote BC family (5) by BC(λ,τ,ξ).

BC family is closely related to NP family, because a different expression of parameters in NP family (2a) gives us BC family. A comparison between (2a) and (5) gives us the following parameter transformation from (δ,σ,η) to (λ,τ,ξ):

$$\lambda = \frac{\delta}{\sigma}, \quad \tau = \sigma\eta^{\lambda}, \begin{cases} \xi = \dfrac{\eta^{\lambda}\exp\left(-\delta^2\right)-1}{\lambda} & (\delta \neq 0) \\ \xi = \log\eta & (\delta = 0) \end{cases}. \tag{6}$$

We remark that λ of BC family (5) is invariant under a scale transformation of $X \to \theta X \ (\theta > 0)$.

Furthermore, for a comparative purpose we shall consider an intermediate family between NP and BC families. Replace $\exp\left(-\delta^2\right)$ by one and put $\delta/\sigma = \lambda$ only in NP family (2a), then we can define a modified new power normal family (shortly, MNP family hereafter) by normality of the following random variable

$$\frac{1}{\sigma}\frac{\left\{\left(\dfrac{X}{\eta}\right)^{\lambda}-1\right\}}{\lambda}. \tag{7}$$

The random variable $\left\{(X/\eta)^{\lambda}-1\right\}/\lambda$ is supposed to be approximately distributed as normal with mean zero and variance σ^2. MNP family is similar to BC family (5) in the meaning that λ is the same power parameter. We also remark that λ of MNP family (7) is invariant under a scale transformation of $X \to \theta X \ (\theta > 0)$. MNP family has shown its good performance in practical applications (Isogai [7]). We denote MNP family (7) by MNP(λ,σ,η).

In our simulation study, regarding a true population, we adopt some kind of the generalized F distribution (1), of which parameters are given by parameters $\{\delta,\sigma,\eta\}$ of NP family (2), i.e. NP(δ,σ,η). In Section 2 we define this generalized F distribution and denote it by GF(δ,σ,η).

In Section 3, we consider a one-sample problem based on sampling data coming from GF (δ,σ,η). First we show the equivalence of NP (δ,σ,η) and BC (λ,τ,ξ) in the meaning that their parameter spaces have a one-to-one mapping. This one-to-one mapping suggests the existence of some non-linear relationship among parameters of BC (λ,τ,ξ), which causes great variations of estimates of parameters τ and ξ under a scale transformation of $X \to \theta X$ $(\theta > 0)$. In the last of Section 3, we exhibit simulation results for statistical modeling, in which comparison with BC, NP and MNP models is made by bootstrap generalized information criteria under a scale transformation of *X*. The degree of prediction error with BC family varies greatly according to choices of a scale transformation of *X*, but there is no fluctuation in the degree of prediction error with respect to NP and MNP families.

In Section 4 we focus on a two-sample problem based on sampling data coming from two populations GF (δ,σ,η_1) and GF (δ,σ,η_2). In case of $\eta_1 \neq \eta_2$ we show that signs in estimates of the power parameter λ of BC family are apt to change, and at the same time a tendency of Logarithmic transformation appears in BC family.

In Section 5, we examine a multi-sample case, especially, multifactor designs.

Finally, in Section 6, we shall consider a more general regression structure than multifactor designs. In the general regression situation, a scale-invariant property of estimates of δ and λ in NP and MNP families does not hold any longer under individual scale transformations of each random variable. However, using the fact that effects of scale transformations of random variables can be attained by tuning design matrices in regression structures, we can examine the maximum likelihood estimate of the power parameter λ of BC family as well as estimates of power parameters δ and λ in NP and MNP families. Changes of a design matrix cause large fluctuations of an estimate of λ in BC family, but estimates of power parameters δ and λ in NP and MNP families are hardly affected and remain stable.

II. Choice of a True Population for a Simulation Study

We should choose freely one of NP, BC and MNP families as a true population, but we cannot use them directly in our simulation, because these families are incomplete distributions. On the other hand, we know that NP family (2) is an approximate distribution to the generalized F distribution (1), but BC and MNP families do not have such a property. Therefore, we first specify parameters of NP family (2), and then simulate random numbers from the corresponding generalized F distribution (1).

NP family (2) has three parameters δ, σ and η, and the generalized F distribution (1) has four parameters ϕ_1, ϕ_2, γ and η. As η is common, we have only to decide how to specify ϕ_1, ϕ_2 and γ for given δ and σ.

Before specification of parameters ϕ_1, ϕ_2 and γ, we recall some restrictions concerning δ_2 and σ:

(i) $0 < \delta_2 < 1 \quad \left(\delta_2 = \left(\dfrac{1}{\phi_1} + \dfrac{1}{\phi_2} \right)^{1/2} \right)$,

(ii) $0 < \sigma < 1$.

These restrictions were introduced to evaluate approximate formulas about the expectation E[X] and variance V(X) of NP family (2) (see Isogai [5]). Especially, the condition (ii) $0 < \sigma < 1$ ensures unimodality of the generalized F distribution (1) (also see Isogai [4]).

Furthermore, from the definitions of δ_1 and δ_2 in (4), we have relationships between ϕ_i (i =1,2) and $\delta \left(= \delta_1 \delta_2\right)$:

$$\begin{cases} \phi_1^{-1} = \left(1/2\right)\delta_2^2(1+3\delta_1) = \left(1/2\right)\delta_2(\delta_2+3\delta) \\ \phi_2^{-1} = \left(1/2\right)\delta_2^2(1-3\delta_1) = \left(1/2\right)\delta_2(\delta_2-3\delta) \end{cases}.$$

1. Random number generation from the generalized F distribution

For given δ ($-1/3 < \delta < 1/3$) and σ ($0 < \sigma < 1$), we specify parameters ϕ_1, ϕ_2 and γ as follows:

Step one. Choose δ_2 so as to satisfy $3\left|\delta\right| < \delta_2 < 1$. For example, we set

$$\delta_2 = 3\left|\delta\right| + \frac{1-3\left|\delta\right|}{100},$$

where the number 100 is selected appropriately to make δ_2 closer to $3\left|\delta\right|$. This setting is motivated to make ϕ_1 and ϕ_2 as large as possible because NP family (2) is a good approximation to the generalized F distribution (1) with large ϕ_1 and ϕ_2 (Isogai [3]).

Step two. Put $\gamma = \dfrac{\sigma}{\delta_2}$ and $\delta_1 = \dfrac{\delta}{\delta_2}$.

Step three. Put

$$\phi_1 = \left[\frac{\delta_2^2\left(1+3\delta_1\right)}{2} \right]^{-1} \text{ and } \phi_2 = \left[\frac{\delta_2^2\left(1-3\delta_1\right)}{2} \right]^{-1}.$$

Step four. For any given η (> 0), generate a random number

$$X = \eta \left(F_{2\phi_2}^{2\phi_1} \right)^{\gamma}.$$

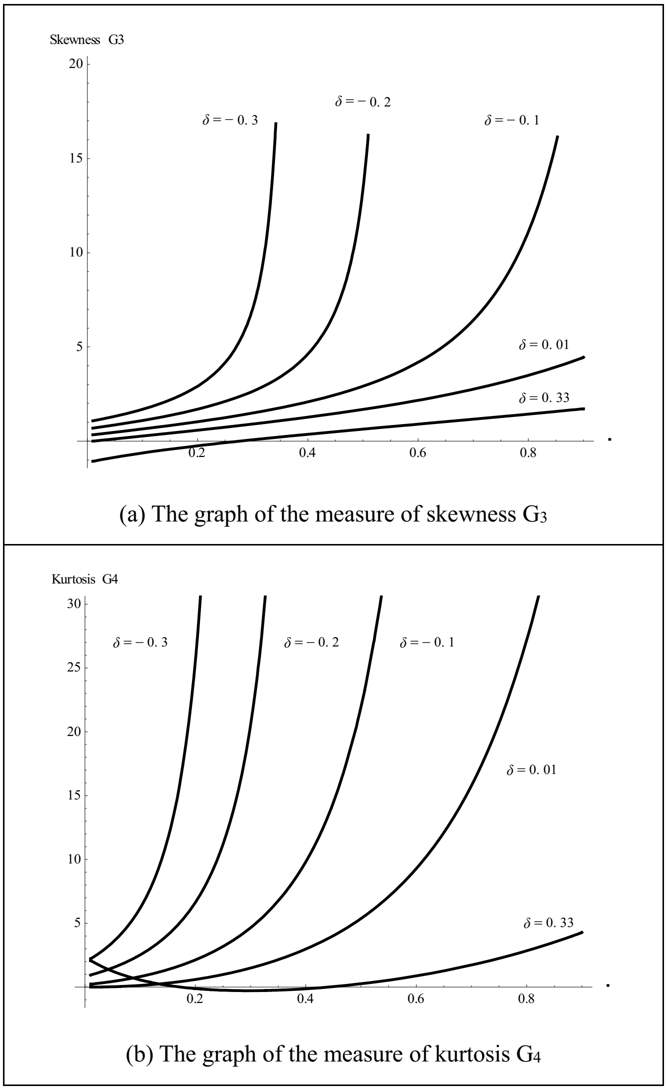
We denote this generalized F distribution corresponding to given δ and σ by GF$\left(\delta, \sigma, \eta\right)$.

2. To use the (n+1)-quantiles of the generalized F distribution $\mathrm{GF}(\delta,\sigma,\eta)$ as a sample

Here we explain our sampling scheme in the following simulation study. Instead of random sampling we use the (n+1)-quantiles as a sample of size n. The (n+1)-quantiles of size n, i.e. $\{x_{(1)},x_{(2)},\cdots,x_{(n)}\}$ from $\mathrm{GF}(\delta,\sigma,\eta)$ with any η are defined by

$$x_{(i)}=\eta\left\{F^{-1}\left(\frac{i}{n+1}\right)\right\}^{\gamma}\ (i=1,\cdots,n), \tag{8}$$

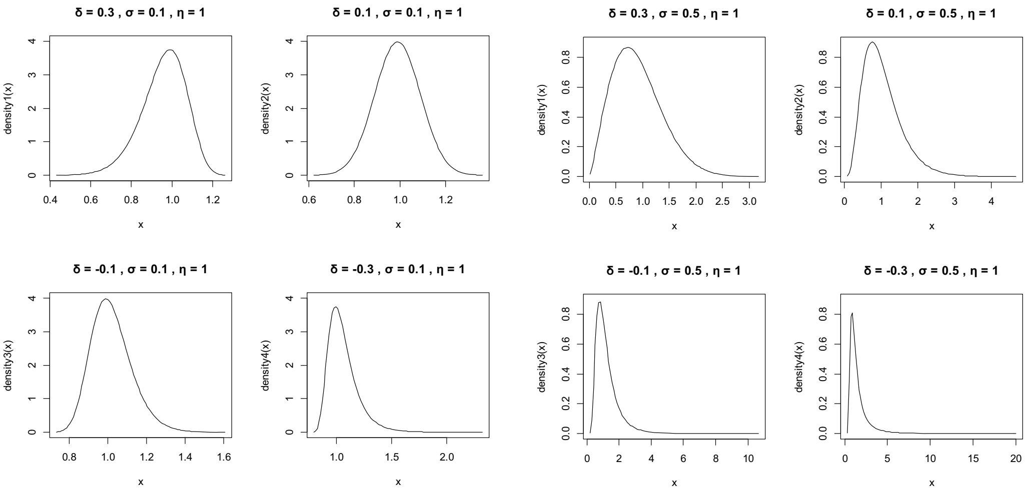
where $F^{-1}(\cdot)$ is the inverse of the cumulative distribution function $F(x)$ of $F_{2\phi_2}^{2\phi_1}$.



(a) The graph of the measure of skewness G_3

(b) The graph of the measure of kurtosis G_4

Figure 1. Graphs of measures G_3 and G_4.



Case (1): $\sigma = 0.1, \eta = 1$ Case (2): $\sigma = 0.5, \eta = 1$

Figure 2. Two cases of density functions of the generalized F distribution GF (δ, σ, η) .

3. Shapes of the generalized F distribution GF(δ, σ, η)

We give some concrete images about the generalized F distribution GF (δ, σ, η) corresponding to given δ ($-1/3 < \delta < 1/3$) and σ ($0 < \sigma < 1$).

Figure 1 shows graphs of the measures of skewness G3 and kurtosis G4 with respect to GF (δ, σ, η) for some δ and σ ($0 < \sigma < 1$), where G3 and G4 are defined by

$$G_3 = \frac{E\left[\left(X - E[X]\right)^3\right]}{\left\{E\left[\left(X - E[X]\right)^2\right]\right\}^{3/2}} \quad \text{and} \quad G_4 = \frac{E\left[\left(X - E[X]\right)^4\right]}{\left\{E\left[\left(X - E[X]\right)^2\right]\right\}^{2}} - 3 .$$

From Figure 1 we know that values of G3 and G4 are positive in the most part of the parameter domain δ ($-1/3 < \delta < 1/3$) and σ ($0 < \sigma < 1$). The degree of non-normality is low when σ is small, and non-normality seems more serious as σ becomes larger.

In Figure 2 we give two cases of density functions of GF (δ, σ, η) corresponding to $\sigma = 0.1$ and $\sigma = 0.5$. From Case (2) in Figure 2 we know that when δ approaches to – 1/3 in case of large σ , the right-hand tails of densities spread out more and more. These tendencies seem to reflect well large positive G3 and G4.

III. One-Sample Problem

1. Moments of BC and MNP families under the generalized F distribution $\mathrm{GF}(\delta,\sigma,\eta)$

Let X be a random variable of the generalized F distribution $\mathrm{GF}(\delta,\sigma,\eta)$. We give the first three approximate cumulants κ_1, κ_2 and κ_3 of X^λ :

$$(1)\ \kappa_1 = \mathrm{E}_{\mathrm{GF}}\left[X^\lambda\right] = \eta^\lambda \exp\left(-\frac{3\delta\sigma}{2}\lambda + \frac{\sigma^2}{2}\lambda^2 + \sigma\lambda O\left(\delta_2{}^2\right)\right)$$

$$= \eta^\lambda \left(1 - \frac{3\delta\sigma}{2}\lambda + \frac{\sigma^2}{2}\lambda^2 + \sigma\lambda O\left(\delta_2{}^2\right)\right), \tag{9}$$

$$(2)\ \kappa_2 = \mathrm{V}_{\mathrm{GF}}\left(X^\lambda\right) = \eta^{2\lambda} \exp\left(-3\sigma\delta\lambda + \sigma^2\lambda^2 + \sigma\lambda O\left(\delta_2{}^2\right)\right)\left(\exp\left\{\sigma^2\lambda^2 + \sigma^2\lambda^2 O\left(\delta_2{}^2\right)\right\} - 1\right)$$

$$= \eta^{2\lambda}\left(\sigma^2\lambda^2 + \sigma^2\lambda^2 O\left(\delta_2{}^2\right)\right), \tag{10}$$

$$\text{Especially,}\quad V_{\mathrm{GF}}\left(X^\lambda\right) \approx \lambda^2\sigma^2\left(E_{\mathrm{GF}}\left[X^\lambda\right]\right)^2, \tag{11}$$

$$(3)\ \kappa_3 = \mathrm{E}_{\mathrm{GF}}\left[\left(X^\lambda - \mathrm{E}_{\mathrm{G}}\left[X^\lambda\right]\right)^3\right] = \eta^{3\lambda}\left\{3\lambda^3\sigma^3(\lambda\sigma - \delta) + \lambda^3\sigma^3 O\left(\delta_2{}^2\right)\right\}, \tag{12}$$

where a symbol $O\left(\delta_2{}^2\right)$ stands for so-called "Landau notation" or "Big-O notation".

To prove the above results, we have only to notice an expression $X^\lambda = \eta^\lambda\left(F_{2\phi_1}^{2\phi_2}\right)^{\lambda\gamma}$ in the appendix of Isogai [5]. Using these cumulants κ_1, κ_2 and κ_3 of X^λ, we can obtain the first three cumulants $\kappa_1{}^{\mathrm{BC}}$, $\kappa_2{}^{\mathrm{BC}}$ and $\kappa_3{}^{\mathrm{BC}}$ of $\left(X^\lambda - 1\right)/\lambda$ for BC family as follows:

$$\kappa_1{}^{\mathrm{BC}} = \mathrm{E}_{\mathrm{GF}}\left[\frac{X^\lambda - 1}{\lambda}\right] = \frac{\eta^\lambda - 1}{\lambda} + \eta^\lambda\left\{-\frac{3\delta\sigma}{2} + \frac{\sigma^2}{2}\lambda + \sigma O\left(\delta_2{}^2\right)\right\}, \tag{13}$$

$$\kappa_2{}^{\mathrm{BC}} = \mathrm{V}_{\mathrm{GF}}\left(\frac{X^\lambda - 1}{\lambda}\right) = \eta^{2\lambda}\sigma^2\left\{1 + O\left(\delta_2{}^2\right)\right\}, \tag{14}$$

$$\kappa_3{}^{\mathrm{BC}} = \mathrm{E}_{\mathrm{GF}}\left[\left(\frac{X^\lambda - 1}{\lambda} - \mathrm{E}_{\mathrm{GF}}\left[\frac{X^\lambda - 1}{\lambda}\right]\right)^3\right] = \eta^{3\lambda}\sigma^3\left\{3(\lambda\sigma - \delta) + O\left(\delta_2{}^2\right)\right\}. \tag{15}$$

Similarly, the first three cumulants κ_1^{MNP} , κ_2^{MNP} and κ_3^{MNP} of $\left((X/\eta)^{\lambda}-1\right)/\lambda$ for MNP family are as follows:

$$\kappa_1^{\mathrm{MNP}}=\mathrm{E}_{\mathrm{GF}}\left[\frac{(X/\eta)^{\lambda}-1}{\lambda}\right]=-\frac{3\delta\sigma}{2}+\frac{\sigma^2}{2}\lambda+\sigma O\left(\delta_2^{\ 2}\right), \tag{16}$$

$$\kappa_2^{\mathrm{MNP}}=\mathrm{V}_{\mathrm{GF}}\left(\frac{(X/\eta)^{\lambda}-1}{\lambda}\right)=\sigma^2\left\{1+O\left(\delta_2^{\ 2}\right)\right\}, \tag{17}$$

$$\kappa_3^{\mathrm{MNP}}=\mathrm{E}_{\mathrm{GF}}\left[\left(\frac{(X/\eta)^{\lambda}-1}{\lambda}-\mathrm{E}_{\mathrm{GF}}\left[\frac{(X/\eta)^{\lambda}-1}{\lambda}\right]\right)^3\right]=\sigma^3\left\{3(\lambda\sigma-\delta)+O\left(\delta_2^{\ 2}\right)\right\}. \tag{18}$$

When $\lambda=\delta/\sigma$, BC and MNP families attain normalizing transformation in the meaning that the standardized third order cumulants $\kappa_3^{\mathrm{BC}}/\left(\kappa_2^{\mathrm{BC}}\right)^{3/2}$ and $\kappa_3^{\mathrm{MNP}}/\left(\kappa_2^{\mathrm{MNP}}\right)^{3/2}$ are nearly equal to zero. In case of $\lambda=0$, Logarithmic transformation appears, and $\kappa_3^{\mathrm{BC}}/\left(\kappa_2^{\mathrm{BC}}\right)^{3/2}$ and $\kappa_3^{\mathrm{MNP}}/\left(\kappa_2^{\mathrm{MNP}}\right)^{3/2}$ are nearly equal to -3δ. This fact means that Logarithmic transformation never attains normalizing transformation unless $\delta=0$. The formula κ_2^{MNP} tells us that it is natural to consider σ as a dispersion parameter in our setting.

2. One-to-one mapping between parameter spaces of BC and NP families

For a parameter space of NP (δ,σ,η), we put

$$\Pi_{\mathrm{NP}}=\left\{(\delta,\sigma,\eta)\middle|-\infty<\delta<\infty,\sigma>0,\eta>0\right\}. \tag{19}$$

To construct a parameter space of BC (λ,τ,ξ), from the last equation of (6) we need a restriction

$$1+\lambda\xi>0, \tag{20}$$

and so we put

$$\Pi_{\mathrm{BC}}=\left\{(\lambda,\tau,\xi)\middle|-\infty<\lambda<\infty,\tau>0,-\infty<\xi<\infty,1+\lambda\xi>0\right\}. \tag{21}$$

For given any point (λ,τ,ξ) in Π_{BC}, the corresponding point (δ,σ,η) in Π_{NP} is given by the following equations

$$\sigma\exp\left(\lambda^2\sigma^2\right)=\frac{\tau}{1+\lambda\xi},\ \delta=\lambda\sigma,\ \begin{cases}\eta=\exp\left(\lambda\sigma^2\right)(1+\lambda\xi)^{1/\lambda} & (\lambda\neq 0)\\ \eta=\exp\xi & (\lambda=0)\end{cases}. \tag{22}$$

A function $\sigma \exp\left(\lambda^2\sigma^2\right)$ in the first equation of (22) is a strongly increasing function of σ, and so σ is determined uniquely. Parameter spaces Π_{NP} and Π_{BC} have a one to one mapping (6) and its inverse mapping (22). Thus, BC (λ,τ,ξ) with Π_{BC} and NP (δ,σ,η) with Π_{NP} are equivalent models with different parameter representations.

The parameter space Π_{BC} has some singular property. Parameterization (6) indicates the existence of some functional relationship among ξ, λ and τ such that

$$\sigma \exp\left(\delta^2\right) = \frac{\tau}{1+\lambda\xi}. \tag{23}$$

For fixed δ and σ, under a scale transformation of $X \to \theta X$ $(\theta > 0)$, λ is supposed to be invariant, and $\sigma \exp\left(\delta^2\right)$ in the left hand side of (23) is constant. Thus, ξ and τ are functionally related to each other. This fact is reflected in parameter estimation of ξ, λ and τ ,that is, estimates of ξ and τ are highly correlated.

3. Effects of fluctuations of η in a one-sample problem

For a given sample of size n from GF (δ,σ,η), we shall examine maximum likelihood estimation of parameters in BC (λ,τ,ξ), NP (δ,σ,η) and MNP (λ,σ,η) families. Let us denote a product of n densities with respect to GF (δ,σ,η) by $\mathrm{GF}(\delta,\sigma,\eta)^n$. Similarly, $\mathrm{BC}(\lambda,\tau,\xi)^n$, $\mathrm{NP}(\delta,\sigma,\eta)^n$ and $\mathrm{MNP}(\lambda,\sigma,\eta)^n$ express products of n BC, n NP and n MNP densities respectively. Also, let us denote maximum likelihood estimates of δ, σ and $\mu(=\log(\eta))$ of $\mathrm{NP}(\delta,\sigma,\eta)^n$ by $\hat{\delta}$, $\hat{\sigma}_{\mathrm{NP}}$ and $\hat{\mu}_{\mathrm{NP}}\left(=\log\left(\hat{\eta}_{\mathrm{NP}}\right)\right)$. Similarly, let us denote maximum likelihood estimates of λ in $\mathrm{BC}(\lambda,\tau,\xi)^n$ and $\mathrm{MNP}(\lambda,\sigma,\eta)^n$ by $\hat{\lambda}_{\mathrm{BC}}$ and $\hat{\lambda}_{\mathrm{MNP}}$ respectively. The other estimates are denoted by $\hat{\tau}$, $\hat{\xi}$, $\hat{\sigma}_{\mathrm{MNP}}$ and $\hat{\mu}_{\mathrm{MNP}}\left(=\log(\hat{\eta}_{\mathrm{MNP}})\right)$.

Here we give a remark concerning the maximum likelihood estimation of $\{\lambda,\tau,\xi\}$ in $\mathrm{BC}(\lambda,\tau,\xi)^n$. For a given random sample $\{X_1=x_1, X_2=x_2,\cdots,X_n=x_n\}$ from GF (δ,σ,η), we use the standardized power transform

$$x_i^{(\lambda)} = \begin{cases} \dfrac{x_i^{\lambda}-1}{\lambda\dot{x}^{\lambda-1}}, \text{for } \lambda \neq 0, \\ \dot{x}\log x_i, \text{for } \lambda = 0, \end{cases}$$

where

$$\dot{x} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

is the geometric mean of the observations $x_1, x_2, \cdots, x_n$. In the following sections, to estimate parameters in BC families we always use the above power transform. For the details of the power transform, see Box and Cox [1], and Draper and Smith [2].

Now we give an example. Using the formula (8) we first draw the (n+1)-quantiles of size n=30 from GF (δ,σ,η) with $\delta=-0.3$, $\sigma=0.1$ and $\eta=1$. Then, we move η in the form of $\eta=\exp(Index)$ $(-5 \le Index \le 5)$, and examine behaviors of estimates $\hat{\lambda}_{\mathrm{BC}}$, $\hat{\xi}$ and $\hat{\tau}$ of $\mathrm{BC}(\lambda,\tau,\xi)^n$ as well as a ratio $\hat{\tau}/\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}\right)$. The result is given in Table 1.

Clearly $\hat{\xi}$ and $\hat{\tau}$ change together extremely according to varying η. On the other hand, the ratio $\hat{\tau}/\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}\right)$ seems to be constant as expected. Estimates of power parameters $\hat{\lambda}_{\mathrm{BC}}$, $\hat{\delta}$ and $\hat{\lambda}_{\mathrm{MNP}}$ are invariant under a scale transformation of *X*. Values of $\hat{\lambda}_{\mathrm{BC}}$ and $\hat{\lambda}_{\mathrm{MNP}}$ are very close. Also, estimates of dispersion parameters $\hat{\sigma}_{\mathrm{NP}}$ and $\hat{\sigma}_{\mathrm{MNP}}$ are invariant and close. Values of $\hat{\mu}_{\mathrm{NP}}$ and $\hat{\mu}_{\mathrm{MNP}}$ move in a similar way. There is no difference in minimum values of negative log-likelihood functions with respect to BC, NP and MNP families.

Table 1. Example on behaviors of power parameter estimates and the other parameter estimates in BC, NP and MNP families for the one-sample case: GF $(\delta,\sigma,\eta)^n$ with $n=30$, $\delta=-0.3, \sigma=0.1$, $\eta=\exp(Index)$ ($Index=-5,-4,\cdots,4,5$) and Ratio $=\hat{\tau}/\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}\right)$.

Index	$\hat{\lambda}_{\mathrm{BC}}$	$\hat{\tau}$	$\hat{\xi}$	Ratio	$\hat{\delta}$	$\hat{\sigma}_{\mathrm{NP}}$	$\hat{\mu}_{\mathrm{NP}}$	$\hat{\lambda}_{\mathrm{MNP}}$	$\hat{\sigma}_{\mathrm{MNP}}$	$\hat{\mu}_{\mathrm{MNP}}$	MinBC	MinNP	MinMNP
-5	-2.402	1.512E+04	-6.300E+04	0.0999	-0.227	0.0949	-4.987	-2.402	0.0999	-4.966	-174.181	-174.181	-174.181
-4	-2.402	1.369E+03	-5.705E+03	0.0999	-0.227	0.0949	-3.987	-2.402	0.0999	-3.966	-144.181	-144.181	-144.181
-3	-2.402	1.240E+02	-5.162E+02	0.0999	-0.227	0.0949	-2.987	-2.402	0.0999	-2.966	-114.181	-114.181	-114.181
-2	-2.402	1.122E+01	-4.636E+01	0.0999	-0.227	0.0949	-1.987	-2.402	0.0999	-1.966	-84.181	-84.181	-84.181
-1	-2.402	1.016E+00	-3.819E+00	0.0999	-0.227	0.0949	-0.987	-2.402	0.0999	-0.966	-54.181	-54.181	-54.181
0	-2.402	9.204E-02	3.279E-02	0.0999	-0.227	0.0949	0.013	-2.402	0.0999	0.034	-24.181	-24.181	-24.181
1	-2.402	8.334E-03	3.816E-01	0.0999	-0.227	0.0949	1.013	-2.402	0.0999	1.034	5.819	5.819	5.819
2	-2.402	7.547E-04	4.132E-01	0.0999	-0.227	0.0949	2.013	-2.402	0.0999	2.034	35.819	35.819	35.819
3	-2.402	6.834E-05	4.161E-01	0.0999	-0.227	0.0949	3.013	-2.402	0.0999	3.034	65.819	65.819	65.819
4	-2.402	6.188E-06	4.163E-01	0.0999	-0.227	0.0949	4.013	-2.402	0.0999	4.034	95.819	95.819	95.819
5	-2.402	5.603E-07	4.163E-01	0.0999	-0.227	0.0949	5.013	-2.402	0.0999	5.034	125.819	125.819	125.819

(Note: MinBC, MinNP and MinMNP mean the minimum values of negative log-likelihood functions with respect to BC, NP and MNP families respectively.)

4. Model selection using bootstrap generalized information criteria

Let us denote one of densities with respect to BC, NP and MNP families by $g(x|\boldsymbol{\theta})$, where a parameter vector $\boldsymbol{\theta}$ means one of corresponding parameter vectors (λ,τ,ξ), (δ,σ,η) and (λ,σ,η). For a given random sample $\mathbf{x}^{(n)}=\{x_1,x_2,\cdots,x_n\}$, put

$$L\left(\mathbf{x}^{(n)}\middle|\boldsymbol{\theta}\right)=\frac{1}{n}\sum_{\alpha=1}^{n}\log g\left(x_\alpha\middle|\boldsymbol{\theta}\right). \tag{24}$$

Maximization of $L\left(\mathbf{x}^{(n)}\middle|\boldsymbol{\theta}\right)$ with respect to $\boldsymbol{\theta}$ gives us the maximum likelihood estimate $\hat{\boldsymbol{\theta}}\left(\mathbf{x}^{(n)}\right)$.

In statistical modeling, when we encounter a problem of model selection, we want to choose a model having a small prediction error. Our prediction error can be evaluated by the following risk

$$E_{\mathbf{x}^{(n)}}E_Z\left[-\log g\left(Z\middle|\hat{\boldsymbol{\theta}}\left(\mathbf{x}^{(n)}\right)\right)\right], \tag{25}$$

where random variables Z and $\mathbf{x}^{(n)}$ are independent, and expectations with respect to Z and $\mathbf{x}^{(n)}$ are defined by their common unobservable true population. The above risk is so-called "information criterion" based on "Kullback-Leibler information" (see Kullback [11]). A well-known estimate of the risk is Akaike's Information Criterion, i.e. AIC, which is given by

$$\mathrm{AIC}=-L\left(\mathbf{x}^{(n)}\middle|\hat{\boldsymbol{\theta}}\left(\mathbf{x}^{(n)}\right)\right)+\widehat{\mathrm{Bias}}, \tag{26}$$

where

$$\widehat{\mathrm{Bias}}=\frac{1}{n}\dim\left(\boldsymbol{\theta}\right). \tag{27}$$

The derivation of $\widehat{\mathrm{Bias}}$ in AIC depends heavily on regular properties of the density $g(x|\boldsymbol{\theta})$. On the other hand, our density $g(x|\boldsymbol{\theta})$ is an approximate and incomplete one, and so we cannot estimate a bias term analytically. However, using bootstrap methods (see Efron and Tibshirani [3]), we can indirectly estimate the bias term. Thus, due to improper properties of the density $g(x|\boldsymbol{\theta})$, our risk is considered as "generalized information criterion". For the details of "information criteria" and related topics, see Konishi and Kitagawa [9].

Let us denote B bootstrap samples generated from the original data set $\mathbf{x}^{(n)}=\{x_1,x_2,\cdots,x_n\}$ by $\mathbf{y}^{(n)}(b)=\{y_1(b),y_2(b),\cdots,y_n(b)\}$ $(b=1,2,\cdots,B)$, where $y_\alpha(b)$ $(\alpha=1,\cdots,n)$ are obtained by randomly sampling n times, with replacement, from $\mathbf{x}^{(n)}=\{x_1,x_2,\cdots,x_n\}$. For the b-th bootstrap sample $\mathbf{y}^{(n)}(b)=\{y_1(b),y_2(b),\cdots,y_n(b)\}$, maximization of $L\left(\mathbf{y}^{(n)}(b)\middle|\boldsymbol{\theta}\right)$ with respect to $\boldsymbol{\theta}$ gives us the maximum likelihood estimate $\hat{\boldsymbol{\theta}}\left(\mathbf{y}^{(n)}(b)\right)$. Then, an estimate of "generalized information criterion", i.e. GIC is given by

$$\mathrm{GIC}=-L\left(\mathbf{x}^{(n)}\middle|\hat{\boldsymbol{\theta}}\left(\mathbf{x}^{(n)}\right)\right)+\widehat{\mathrm{Bias}}, \tag{28}$$

where a bootstrap bias estimate is constructed by

$$\widehat{\mathrm{Bias}}=\frac{1}{B}\sum_{b=1}^{B}\left\{L\left(\mathbf{y}^{(n)}(b)\middle|\hat{\boldsymbol{\theta}}\left(\mathbf{y}^{(n)}(b)\right)\right)-L\left(\mathbf{x}^{(n)}\middle|\hat{\boldsymbol{\theta}}\left(\mathbf{y}^{(n)}(b)\right)\right)\right\}. \tag{29}$$

Table 2. Example on behaviors of bootstrap generalized information criteria $n\,\mathrm{GIC}_{\mathrm{BC}}$, $n\,\mathrm{GIC}_{\mathrm{NP}}$ and $n\,\mathrm{GIC}_{\mathrm{MNP}}$ with respect to BC, NP and MNP families for the one-sample case: GF $(\delta,\sigma,\eta)^n$ with $n=30$, $\delta=-0.3,\sigma=0.1$, $\eta=10^{Index}$ ($-5\le Index\le 5$). $n\,\mathrm{GIC}_{(*)}=\mathrm{Min}(*)+n\,\widehat{\mathrm{Bias}}_{(*)}$ $(*=\mathrm{BC},\mathrm{NP},\mathrm{MNP})$ and calculation of $n\,\widehat{\mathrm{Bias}}_{(*)}$ is based on 10000(=B) bootstrap samples.

Index	MinBC	MinNP	MinMBC	$n\,\widehat{\mathrm{Bias}}_{\mathrm{BC}}$	$n\,\widehat{\mathrm{Bias}}_{\mathrm{NP}}$	$n\,\widehat{\mathrm{Bias}}_{\mathrm{MNP}}$	$n\,\mathrm{GIC}_{\mathrm{BC}}$	$n\,\mathrm{GIC}_{\mathrm{NP}}$	$n\,\mathrm{GIC}_{\mathrm{MNP}}$
-5	-369.569	-369.569	-369.569	1.62E+08	2.743	2.776	1.62E+08	-366.826	-366.793
-4	-300.491	-300.491	-300.491	4.58E+06	2.541	2.570	4.58E+06	-297.950	-297.921
-3.7	-279.768	-279.768	-279.768	2.18E+04	2.568	2.598	2.16E+04	-277.199	-277.170
-3.5	-265.952	-265.952	-265.952	2.630	2.622	2.650	-263.322	-263.330	-263.302
-3.2	-245.229	-245.229	-245.229	2.650	2.637	2.665	-242.579	-242.592	-242.564
-3.1	-238.321	-238.321	-238.321	2.702	2.695	2.724	-235.620	-235.626	-235.597
-3	-231.413	-231.413	-231.413	2.689	2.683	2.713	-228.725	-228.730	-228.700
-2	-162.336	-162.336	-162.336	2.660	2.656	2.683	-159.676	-159.680	-159.652
-1	-93.258	-93.258	-93.258	2.566	2.558	2.583	-90.692	-90.700	-90.675
0	-24.181	-24.181	-24.181	2.570	2.568	2.592	-21.611	-21.613	-21.588
1	44.897	44.897	44.897	2.686	2.684	2.713	47.583	47.581	47.610
2	113.974	113.974	113.974	2.554	2.550	2.577	116.528	116.524	116.552
3	183.052	183.052	183.052	2.892	2.610	2.636	185.944	185.662	185.688
3.1	189.960	189.960	189.960	5.532	2.549	2.574	195.491	192.508	192.534
3.2	196.867	196.867	196.867	61.618	2.610	2.640	258.486	199.478	199.507
3.5	217.591	217.591	217.591	3.90E+04	2.640	2.667	3.92E+04	220.231	220.257
3.7	231.406	231.406	231.406	5.61E+06	2.639	2.671	5.61E+06	234.045	234.077
4	252.129	252.129	252.129	5.44E+09	2.651	2.682	5.44E+09	254.781	254.812
5	321.207	321.207	321.207	4.83E+19	2.668	2.698	4.83E+19	323.875	323.905

(Note: MinBC, MinNP and MinMNP mean the minimum values of negative log-likelihood functions with respect to BC, NP and MNP families respectively.)

Now we check GIC with respect to BC, NP and MNP families. We first draw the (n+1)-quantiles of size n=30 from $\mathrm{GF}(\delta,\sigma,\eta)$ with $\delta=-0.3$, $\sigma=0.1$ and $\eta=1$, which is the same sample in the preceding subsection. Then, we move η as $\eta=10^{Index}$ $(-5\le Index\le 5)$, and, by using 10000(=B) bootstrap samples generated from the original sample quantiles of size n=30, examine behaviors of $n\,\mathrm{GIC}_{\mathrm{BC}}$, $n\,\mathrm{GIC}_{\mathrm{NP}}$ and $n\,\mathrm{GIC}_{\mathrm{MNP}}$, where $n\,\mathrm{GIC}$ means GIC times n. The results are given in Table 2.

Clearly $n\,\mathrm{GIC}_{\mathrm{BC}}$ becomes larger than $n\,\mathrm{GIC}_{\mathrm{NP}}$ and $n\,\mathrm{GIC}_{\mathrm{MNP}}$, as the absolute value of *Index* deviates from 3. Especially $n\,\widehat{\mathrm{Bias}}_{\mathrm{BC}}$ becomes considerably larger than $n\,\widehat{\mathrm{Bias}}_{\mathrm{NP}}$ and $n\,\widehat{\mathrm{Bias}}_{\mathrm{MNP}}$, as η becomes fairly large or fairly small. On the other hand, $n\,\widehat{\mathrm{Bias}}_{\mathrm{NP}}$ and $n\,\widehat{\mathrm{Bias}}_{\mathrm{MNP}}$ are small and stable in all cases of η. The degree of prediction error of BC family is larger than that of NP and MNP families.

IV. Two-Sample Problem

We consider two populations $\mathrm{GF}(\delta,\sigma,\eta_1)$ and $\mathrm{GF}(\delta,\sigma,\eta_2)$, where δ and σ are common but $\eta's$ are different, and distinguish their random variables X_1 and X_2 by subscripts 1 and 2. For given two random samples of sample sizes n_1 and n_2 from $\mathrm{GF}(\delta,\sigma,\eta_1)$ and $\mathrm{GF}(\delta,\sigma,\eta_2)$, we denote their joint densities by $\mathrm{GF}(\delta,\sigma,\eta_1)^{n_1}$ and $\mathrm{GF}(\delta,\sigma,\eta_2)^{n_2}$ respectively. We shall examine maximum likelihood estimation of parameters in BC families $\prod_{i=1}^{2}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, NP families $\prod_{i=1}^{2}\mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$ and MNP families $\prod_{i=1}^{2}\mathrm{MNP}(\delta,\sigma,\eta_i)^{n_i}$.

In applications of BC families $\prod_{i=1}^{2}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, we assume equality of variances with respect to transformed variables X_1^{λ} and X_2^{λ}. Their nominal means and variances with respect to X_1^{λ} and X_2^{λ} are given by

$$E_{\mathrm{BC}}\left[X_i^{\lambda}\right]=1+\lambda\xi_i\ (i=1,2), V_{\mathrm{BC}}\left(X_i^{\lambda}\right)=\lambda^2\tau^2\ (i=1,2). \tag{30}$$

1. Inevitability of Logarithmic transformation

We examine a case where variances of X_1^{λ} and X_2^{λ} are equal to each other. Put

$$V_{\mathrm{GF}}\left(\frac{X_1^{\lambda}-1}{\lambda}\right)=V_{\mathrm{GF}}\left(\frac{X_2^{\lambda}-1}{\lambda}\right). \tag{31}$$

Then, from the formula (14), the above equation leads to

$$\lambda\left(\log\eta_1 - \log\eta_2\right) = 0 \tag{32}$$

Unless $\eta_1 = \eta_2$, λ is always equal to zero. A situation $\eta_1 \neq \eta_2$ is apt to suggest that Logarithmic transformation is appropriate for BC families $\prod_{i=1}^{2} \mathrm{BC}\left(\lambda,\tau,\xi_i\right)^{n_i}$.

2. Some relationship among parameters of BC families

For joint densities $\mathrm{GF}\left(\delta,\sigma,\eta_1\right)^{n_1}$ and $\mathrm{GF}\left(\delta,\sigma,\eta_2\right)^{n_2}$, the total of variances with respect to X_1^{λ} and X_2^{λ} is

$$TV_{\mathrm{GF}} = n_1 V_{\mathrm{GF}}\left(X_1^{\lambda}\right) + n_2 V_{\mathrm{GF}}\left(X_2^{\lambda}\right) = n_1\lambda^2\sigma^2\left(E_{\mathrm{GF}}\left[X_1^{\lambda}\right]\right)^2 + n_2\lambda^2\sigma^2\left(E_{\mathrm{GF}}\left[X_2^{\lambda}\right]\right)^2, \tag{33}$$

where we have used the formula (11). In applying BC families $\prod_{i=1}^{2} \mathrm{BC}\left(\lambda,\tau,\xi_i\right)^{n_i}$, the total of variances for BC families $\prod_{i=1}^{2} \mathrm{BC}\left(\lambda,\tau,\xi_i\right)^{n_i}$ is

$$TV_{\mathrm{BC}} = n_1 V_{\mathrm{BC}}\left(X_1^{\lambda}\right) + n_2 V_{\mathrm{BC}}\left(X_2^{\lambda}\right) = \left(n_1 + n_2\right)\lambda^2\tau^2 . \tag{34}$$

Here, we suppose that

$$TV_{\mathrm{GF}} = TV_{\mathrm{BC}}, \quad E_{\mathrm{GF}}\left[X_i^{\lambda}\right] = E_{\mathrm{BC}}\left[X_i^{\lambda}\right] \quad \left(i = 1,2\right). \tag{35}$$

Then, we have the following relationship

$$\sigma = \frac{\tau}{\sqrt{\dfrac{n_1}{n_1+n_2}\left(1+\lambda\xi_1\right)^2 + \dfrac{n_2}{n_1+n_2}\left(1+\lambda\xi_2\right)^2}} . \tag{36}$$

Parameters $\{\lambda,\tau,\xi_1,\xi_2\}$ of $\mathrm{BC}\left(\lambda,\tau,\xi_1\right)$ and $\mathrm{BC}\left(\lambda,\tau,\xi_2\right)$ are functionally related to each other. There is a possibility of high correlations among estimates of parameters $\{\lambda,\tau,\xi_1,\xi_2\}$. The relationship (36) seems to be an extension of (23).

3. Effects of fluctuations of η_1 and η_2 in two samples from GF $\left(\delta,\sigma,\eta_1\right)$ and GF $\left(\delta,\sigma,\eta_2\right)$

Now we give a simple example. For the case $\delta = -0.3$, $\sigma = 0.1$ and $\eta = 1$ we first generate two sets of (n+1)-quantiles of sizes $n_1 = 30$ and $n_2 = 20$ from $\mathrm{GF}\left(\delta,\sigma,\eta\right)$. Then, we construct two samples from GF $\left(\delta,\sigma,\eta_1\right)$ and GF $\left(\delta,\sigma,\eta_2\right)$ with fixed η_1 and varying η_2 . Here we put $\eta_1 = \exp(-3)$ and

$\eta_2 = \eta_1 \exp(Index)$, where *Index* takes integer values from -5 to 5. For given two sets of (n_i +1)-quantiles $(i = 1,2)$ from GF (δ, σ, η_1) and GF (δ, σ, η_2) with $\eta_1 (= \exp(\mu_1))$ and η_2 $(= \exp(\mu_2))$, we estimate $\{\lambda, \tau, \xi_1, \xi_2\}$ of $\prod_{i=1}^{2} \mathrm{BC}(\lambda, \tau, \xi_i)^{n_i}$ and examine behaviors of not only estimates $\hat{\lambda}_{\mathrm{BC}}$, $\hat{\xi}_i$ (i = 1,2) and $\hat{\tau}$ but also a ratio $\hat{\tau} / \sqrt{\frac{n_1}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_1\right)^2 + \frac{n_2}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_2\right)^2}$. The result is given in Table 3.

Clearly, in a situation $\eta_1 \neq \eta_2$ (or $\mu_1 \neq \mu_2$), $\hat{\lambda}_{\mathrm{BC}}$ moves around zero, which indicates Logarithmic transformation, and has positive and negative signs. Also, $\hat{\xi}_1$, $\hat{\xi}_2$ and $\hat{\tau}$ change together solidly according to varying η_2 . On the other hand, the case $\eta_1 = \eta_2$ (or $\mu_1 = \mu_2$) suggests a true value of $\hat{\lambda}_{\mathrm{BC}}$, which is far from Logarithmic transformation. Absolute values of $\hat{\xi}_1$, $\hat{\xi}_2$ and $\hat{\tau}$ become very large. The ratio $\hat{\tau} / \sqrt{\frac{n_1}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_1\right)^2 + \frac{n_2}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_2\right)^2}$ seems to be almost constant as expected. Behaviors of $\hat{\lambda}_{\mathrm{BC}}$ show that $\hat{\lambda}_{\mathrm{BC}}$ is not invariant under scale transformations of $X_1 \to \theta_1 X_1$ and $X_2 \to \theta_2 X_2$ $(\theta_1 \neq \theta_2, \theta_1 > 0, \theta_2 > 0)$.

Table 3. Example on behaviors of power parameter estimates and the other parameter estimates in BC, NP and MNP families for the two-sample case.

μ_1	μ_2	$\hat{\lambda}_{\mathrm{BC}}$	$\hat{\tau}$	$\hat{\xi}_1$	$\hat{\xi}_2$	Ratio	$\hat{\delta}$	$\hat{\sigma}_{\mathrm{NP}}$	$\hat{\mu}_1^{\mathrm{NP}}$	$\hat{\mu}_2^{\mathrm{NP}}$	$\hat{\lambda}_{\mathrm{MNP}}$	$\hat{\sigma}_{\mathrm{MNP}}$	$\hat{\mu}_1^{\mathrm{MNP}}$	$\hat{\mu}_2^{\mathrm{MNP}}$	MinBC	MinNP	MinMNP
-3	-8	-0.0113	0.1112	-3.003	-8.323	0.1050	-0.2225	0.0938	-2.987	-7.986	-2.396	0.0985	-2.966	-7.965	-289.35	-291.07	-291.07
-3	-7	-0.0149	0.1125	-3.019	-7.327	0.1050	-0.2225	0.0938	-2.987	-6.986	-2.392	0.0986	-2.966	-6.965	-269.35	-271.07	-271.07
-3	-6	-0.0216	0.1149	-3.049	-6.353	0.1050	-0.2225	0.0938	-2.987	-5.986	-2.394	0.0985	-2.966	-5.965	-249.36	-251.07	-251.07
-3	-5	-0.0372	0.1208	-3.121	-5.440	0.1049	-0.2225	0.0938	-2.987	-4.986	-2.392	0.0985	-2.966	-4.965	-229.38	-231.07	-231.07
-3	-4	-0.1023	0.1479	-3.448	-4.874	0.1047	-0.2225	0.0938	-2.987	-3.986	-2.391	0.0986	-2.966	-3.965	-209.44	-211.07	-211.07
-3	-3	-2.3906	118.2664	-501.551	-501.651	0.0985	-0.2225	0.0938	-2.987	-2.986	-2.393	0.0985	-2.966	-2.965	-191.07	-191.07	-191.07
-3	-2	-0.0190	0.1104	-3.037	-1.991	0.1052	-0.2225	0.0938	-2.987	-1.986	-2.394	0.0985	-2.966	-1.965	-169.31	-171.07	-171.07
-3	-1	0.0064	0.1038	-2.925	-0.951	0.1052	-0.2225	0.0938	-2.987	-0.986	-2.393	0.0985	-2.966	-0.965	-149.31	-151.07	-151.07
-3	0	0.0078	0.1038	-2.919	0.046	0.1052	-0.2243	0.0937	-2.987	0.014	-2.394	0.0985	-2.966	0.035	-129.32	-131.07	-131.07
-3	1	0.0072	0.1042	-2.922	1.050	0.1052	-0.2225	0.0938	-2.987	1.014	-2.393	0.0985	-2.966	1.035	-109.32	-111.07	-111.07
-3	2	0.0063	0.1045	-2.925	2.059	0.1052	-0.2225	0.0938	-2.987	2.014	-2.392	0.0985	-2.966	2.035	-89.32	-91.07	-91.07

(Note: The true densities are $\prod_{i=1}^{2} \mathrm{GF}(\delta, \sigma, \eta_i)^{n_i}$ with $n_1 = 30, n_2 = 20, \delta = -0.3, \sigma = 0.1,\ \eta_1 = \exp(-3)$, $\eta_2 = \eta_1 \exp(Index)$ $(Index = -5, -4, \cdots, 5)$. Ratio $= \hat{\tau} / \sqrt{\frac{n_1}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_1\right)^2 + \frac{n_2}{n_1+n_2}\left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_2\right)^2}$,

$\mu_1 = \log\eta_1$ and $\mu_2 = \log\eta_2$. MinBC, MinNP and MinMNP mean the minimum values of negative log-likelihood functions with respect to BC, NP and MNP families respectively.)

Regarding estimates for $\{\delta,\sigma,\mu_1,\mu_2\}$ of NP families $\prod_{i=1}^{2}\mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$ and estimates for $\{\lambda,\sigma,\mu_1,\mu_2\}$ of MNP families $\prod_{i=1}^{2}\mathrm{MNP}(\delta,\sigma,\eta_i)^{n_i}$, values of $\hat{\delta}$, $\hat{\lambda}_{\mathrm{MNP}}$, $\hat{\sigma}_{\mathrm{NP}}$ and $\hat{\sigma}_{\mathrm{MNP}}$ seem to be invariant under scale transformations of $X_1 \to \theta_1 X_1$ and $X_2 \to \theta_2 X_2$ $(\theta_1 > 0, \theta_2 > 0)$.

Also, $\hat{\mu}_1^{\mathrm{NP}}$ and $\hat{\mu}_1^{\mathrm{MNP}}$ as well as $\hat{\mu}_2^{\mathrm{NP}}$ and $\hat{\mu}_2^{\mathrm{MNP}}$ are close, and their values seem near to true values. There is hardly any difference in minimum values of negative log-likelihood functions with respect to BC, NP and MNP families.

V. Multi-Sample Problem

Our results in the preceding section are easily extended to a multi-sample case. Now suppose that we have M random samples of their sample sizes n_i drawn from $\mathrm{GF}(\delta,\sigma,\eta_i)$ (i = 1,$\cdots$, M), and denote their joint density by $\prod_{i=1}^{M}\mathrm{GF}(\delta,\sigma,\eta_i)^{n_i}$. We shall examine maximum likelihood estimation of parameters in BC families $\prod_{i=1}^{M}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, NP families $\prod_{i=1}^{M}\mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$ and MNP families $\prod_{i=1}^{M}\mathrm{MNP}(\delta,\sigma,\eta_i)^{n_i}$.

Regarding parameters in BC families $\prod_{i=1}^{M}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, a similar discussion about the formula (36) leads to the following relationship

$$\sigma = \frac{\tau}{\sqrt{\frac{1}{n_1+n_2+\cdots+n_M}\sum_{i=1}^{M} n_i\left(1+\lambda\xi_i\right)^2}} . \tag{37}$$

Parameters $\{\lambda,\tau,\xi_1,\xi_2,\cdots,\xi_M\}$ of $\prod_{i=1}^{M}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$ are functionally related to each other.

Similarly to two-sample cases in the preceding section, an invariant property of the maximum likelihood estimate $\hat{\lambda}_{\mathrm{BC}}$ may be expected to be broken under scale transformations of $X_i \to \theta_i X_i$ $(\theta_i > 0)$ $(i = 1,\cdots,M)$. Furthermore, the absolute value of $\hat{\lambda}_{\mathrm{BC}}$ may become small as discrepancies in $\eta_i's$ of $\prod_{i=1}^{M}\mathrm{GF}(\delta,\sigma,\eta_i)^{n_i}$ become large.

1. Four-sample case

To check our conjecture on the behavior of $\hat{\lambda}_{\mathrm{BC}}$, we give an example of a four-sample case. For the case $\delta = -0.3$, $\sigma = 0.1$, we generate (n+1)-quantiles of size $n = 8$ from each of $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i = 1,2,3,4)$, where we suppose that

$$\eta_1 = \exp(\beta_0),\ \eta_2 = \exp(\beta_0 + \beta_1),$$
$$\eta_3 = \exp(\beta_0 + \beta_2),\ \eta_4 = \exp(\beta_0 + \beta_1 + \beta_2),$$
$$\beta_0 = 5,\ \beta_1 = 0.1 \times Index,\ \beta_2 = 0.2 \times Index,$$

and *Index* takes integer values from 0 to 10. Note that the above setting is equivalent to a simple version of Example A3 [Two-factor design] in the Appendix, where we have only to put $a = 2$ and $b = 2$ in Case (ii).

As *Index* becomes large, discrepancies in $\eta_i's$ become large. We also put parameters ξ_i $(i = 1,2,3,4)$ of $\mathrm{BC}(\lambda,\tau,\xi_i)$ $(i = 1,2,3,4)$ as

$$\xi_1 = \psi_0,\ \xi_2 = \psi_0 + \psi_1,\ \xi_3 = \psi_0 + \psi_2,\ \xi_4 = \psi_0 + \psi_1 + \psi_2.$$

According to each *Index*, using four samples from $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i = 1,2,3,4)$ mentioned above, we estimate $\{\lambda,\tau,\psi_0,\psi_1,\psi_2\}$ of $\prod_{i=1}^{4}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$ and examine behaviors of not only estimates $\hat{\lambda}_{\mathrm{BC}}$, $\hat{\psi}_i$ (i =0,1,2) and $\hat{\tau}$ but also a ratio $\hat{\tau} / \left\{\sum_{i=1}^{4} n_i \left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_i\right)^2 / (n_1+n_2+n_3+n_4)\right\}^{1/2}$, where $n_1 = n_2 = n_3 = n_4 = 8$, and we put $\hat{\xi}_1 = \hat{\psi}_0, \hat{\xi}_2 = \hat{\psi}_0 + \hat{\psi}_1, \hat{\xi}_3 = \hat{\psi}_0 + \hat{\psi}_2, \hat{\xi}_4 = \hat{\psi}_0 + \hat{\psi}_1 + \hat{\psi}_2$. At the same time, we estimate $\{\delta,\sigma,\beta_0,\beta_1,\beta_2\}$ of $\prod_{i=1}^{4}\mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$ and $\{\lambda,\sigma,\beta_0,\beta_1,\beta_2\}$ of $\prod_{i=1}^{4}\mathrm{MNP}(\lambda,\sigma,\eta_i)^{n_i}$. Results are given in Table 4.

Clearly a change of *Index* causes fluctuations of $\hat{\lambda}_{\mathrm{BC}}$ with respect to $\prod_{i=1}^{4}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, and large values of *Index* indicate Logarithmic transformation. Also, in comparison to $\hat{\beta}_0$, $\hat{\psi}_0$ changes greatly according to *Index*.

On the other hand, the ratio $\hat{\tau} / \left\{\sum_{i=1}^{4} n_i \left(1+\hat{\lambda}_{\mathrm{BC}}\hat{\xi}_i\right)^2 / \left(\sum_{i=1}^{4} n_i\right)\right\}^{1/2}$ seems to be nearly constant as expected. Also, $\hat{\delta}$, $\hat{\lambda}_{\mathrm{MNP}}$, $\hat{\sigma}_{\mathrm{NP}}$ and $\hat{\sigma}_{\mathrm{MNP}}$ are invariant under scale transformations. The values of $\hat{\beta}_i^{\mathrm{NP}}$ and $\hat{\beta}_i^{\mathrm{MNP}}$ $(i = 0,1,2)$ are close, and seem near to true values. There is hardly any difference in minimum values of negative log-likelihood functions with respect to BC, NP and MNP families.

Table 4. Example on behaviors of power parameter estimates and the other parameter estimates in BC, NP and MNP families for the four-sample case.

β_0	β_1	β_2	$\hat{\lambda}_{BC}$	$\hat{\tau}$	$\hat{\psi}_0$	$\hat{\psi}_1$	$\hat{\psi}_2$	Ratio	$\hat{\delta}$	$\hat{\sigma}_{NP}$	$\hat{\beta}_0^{NP}$	$\hat{\beta}_1^{NP}$	$\hat{\beta}_2^{NP}$	$\hat{\lambda}_{MNP}$	$\hat{\sigma}_{MNP}$	$\hat{\beta}_0^{MNP}$	$\hat{\beta}_1^{MNP}$	$\hat{\beta}_2^{MNP}$	MinBC	MinNP	MinMN
5	0	0	-2.22274	1.17E-06	0.450	2.55E-17	2.55E-17	0.0844	-0.1814	0.0817	5.020	2.57E-07	-8E-07	-2.222	0.0844	5.034	-3.5E-05	-2.3E-05	128.25	128.25	128.2
5	0.1	0.2	-0.43976	0.00888	2.026	0.0102	0.0204	0.0869	-0.1814	0.0817	5.020	0.1000	0.2000	-2.222	0.0844	5.034	0.1000	0.2000	133.49	133.05	133.0
5	0.2	0.4	-0.09743	0.05192	3.985	0.1189	0.2377	0.0873	-0.1814	0.0817	5.020	0.2000	0.4000	-2.222	0.0844	5.034	0.2000	0.4000	138.37	137.85	137.8
5	0.3	0.6	-0.03126	0.07363	4.666	0.2527	0.5054	0.0874	-0.1814	0.0817	5.020	0.3000	0.6000	-2.222	0.0844	5.034	0.3000	0.6000	143.19	142.65	142.6
5	0.4	0.8	-0.01249	0.08148	4.888	0.3728	0.7456	0.0874	-0.1814	0.0817	5.020	0.4000	0.8000	-2.224	0.0844	5.034	0.4000	0.8000	147.99	147.45	147.4
5	0.5	1	-0.00581	0.08455	4.970	0.4835	0.9669	0.0874	-0.1814	0.0817	5.020	0.5000	1.0000	-2.222	0.0844	5.034	0.5000	1.0000	152.79	152.25	152.2
5	0.6	1.2	-0.00302	0.08588	5.005	0.5893	1.1787	0.0874	-0.1814	0.0817	5.020	0.6000	1.2000	-2.222	0.0844	5.034	0.6000	1.2000	157.59	157.05	157.0
5	0.7	1.4	-0.00171	0.08653	5.021	0.6927	1.3855	0.0874	-0.1814	0.0817	5.020	0.7000	1.4000	-2.222	0.0844	5.034	0.7000	1.4000	162.39	161.85	161.8
5	0.8	1.6	-0.00104	0.08687	5.030	0.7948	1.5896	0.0874	-0.1814	0.0817	5.020	0.8000	1.6000	-2.222	0.0844	5.034	0.8000	1.6000	167.20	166.65	166.6
5	0.9	1.8	-0.00066	0.08707	5.035	0.8962	1.7924	0.0874	-0.1814	0.0817	5.020	0.9000	1.8000	-2.222	0.0844	5.034	0.9000	1.8000	172.00	171.45	171.4
5	1	2	-0.00044	0.08719	5.037	0.9971	1.9943	0.0874	-0.1814	0.0816	5.020	1.0000	2.0000	-2.222	0.0844	5.034	1.0000	2.0000	176.80	176.25	176.2

(Note: The true densities are $\prod_{i=1}^{4} \mathrm{GF}(\delta,\sigma,\eta_i)^{n_i}$ with $n_1 = n_2 = n_3 = n_4 = 8, \delta = -0.3, \quad \sigma = 0.1,$

$\eta_1 = \exp(\beta_0),\ \eta_2 = \exp(\beta_0 + \beta_1), \eta_3 = \exp(\beta_0 + \beta_2), \eta_4 = \exp(\beta_0 + \beta_1 + \beta_2),\ \beta_0 = 5, \quad \beta_1 = 0.1 \times \mathit{Index},$

$\beta_2 = 0.2 \times \mathit{Index}$ ($\mathit{Index} = 0,1,2,\cdots,10$). As for estimates $\{\hat{\lambda},\hat{\tau},\hat{\psi}_0,\hat{\psi}_1,\hat{\psi}_2\}$ of $\prod_{i=1}^{4} \mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, we put

$\hat{\xi}_1 = \hat{\psi}_0,\ \hat{\xi}_2 = \hat{\psi}_0 + \hat{\psi}_1,\ \hat{\xi}_3 = \hat{\psi}_0 + \hat{\psi}_2,\ \hat{\xi}_4 = \hat{\psi}_0 + \hat{\psi}_1 + \hat{\psi}_2$ and $\mathrm{Ratio} = \hat{\tau} / \left\{ \sum_{i=1}^{4} n_i \left(1 + \hat{\lambda}_{BC} \hat{\xi}_i\right)^2 / \sum_{i=1}^{4} n_i \right\}^{1/2}$.

MinBC, MinNP and MinMNP mean the minimum values of negative log-likelihood functions with respect to BC, NP and MNP families respectively.)

2. Three-factor design: Revisited example

Variations of signs in estimates of power parameters appeared in Example 1 (Isogai [3]), where an estimate $\hat{\lambda}_{BC}$ of λ in BC family (5) had an opposite sign compared to an estimate $\hat{\delta}$ in NP family (2). This data is given by Box and Cox [1] (p.223), and we reanalyze it in our setting.

The data set comes from 3^3 factorial experiment on three factors: length of test specimen (denoted by A), amplitude of loading cycle (denoted by B) and load (denoted by C). Three factors A, B and C have three levels denoted by $\mathrm{A_i}$, $\mathrm{B_j}$ and $\mathrm{C_k}$ $(i, j, k = 1,2,3)$. The response variable Y_{ijk} $(i, j, k = 1,2,3)$ is the number of cycles to failure of worsted yarn, and is supposed to be distributed as $\mathrm{GF}(\delta,\sigma,\eta_{\mathrm{ijk}})$.

By using $\mathrm{NP}(\delta,\sigma,\eta_{ijk})$, $\mathrm{BC}(\lambda,\tau,\xi_{ijk})$ and $\mathrm{MNP}(\lambda,\sigma,\eta_{ijk})$ $(i, j, k = 1,2,3)$, we consider regression analyses on the data. We suppose that $\log \eta_{ijk}$ and ξ_{ijk} have linear structures with main effects (denoted by α_i, β_j and γ_k for i , j, k =1, 2, 3), a constant term (denoted by μ) and no interaction. That is,

$\log\eta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k$ and $\xi_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k$.

To remove any redundancy of parameters, we assume that parameters corresponding to the third levels are equal to zero: $\alpha_3 = 0$, $\beta_3 = 0$, $\gamma_3 = 0$.

To examine the behavior of $\hat{\lambda}_{\mathrm{BC}}$ with respect to $\mathrm{BC}\left(\lambda,\tau,\xi_{ijk}\right)$ $\left(i,j,k = 1,2,3\right)$ under scale transformations of $Y_{ijk} \to Y_{ijk} / \left\{\theta_i \rho_j \zeta_k\right\}$ $\left(\theta_i > 0, \rho_j > 0, \zeta_k > 0\right)$, we consider the following case

$$\theta_i = \begin{cases} 1 & \left(i = 1,2\right) \\ \exp(Index) & \left(i = 3\right) \end{cases}, \quad \rho_j = 1 \ \left(j = 1,2,3\right) \ \text{ and } \ \zeta_k = 1 \ \left(k = 1,2,3\right),$$

where *Index* takes integer values from -10 to 10.

Table 5. Results on regression analyses for the number of cycles to failure data with respect to NP $\left(\delta,\sigma,\eta_{ijk}\right)$, $\mathrm{BC}\left(\lambda,\tau,\xi_{ijk}\right)$ and $\mathrm{MNP}\left(\lambda,\sigma,\eta_{ijk}\right)$ with $\log\eta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k$ and $\xi_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k$ $\left(i,j,k = 1,2,3\right)$ under scale transformations $Y_{ijk} \to Y_{ijk} / \theta_i$, where $\theta_i = 1$ $\left(i = 1,2\right)$ and $\theta_3 = \exp\left(Index\right)$ $\left(Index = -10,-9,\cdots,9,10\right)$. $\mathrm{Ratio} = \hat{\tau} / \left\{\sum_{i,j,k=1}^{3}\left(1 + \hat{\lambda}_{\mathrm{BC}}\hat{\xi}_{ijk}\right)^2 / 27\right\}^{1/2}$ with $\hat{\xi}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k$ $\left(i,j,k = 1,2,3\right)$.

Index	$\hat{\lambda}_{\mathrm{BC}}$	$\hat{\tau}$	Ratio	$\hat{\alpha}_1^{\mathrm{BC}}$	$\hat{\alpha}_2^{\mathrm{BC}}$	$\hat{\alpha}_1^{\mathrm{BC}} - \hat{\alpha}_2^{\mathrm{BC}}$	$\hat{\beta}_1^{\mathrm{BC}}$	$\hat{\beta}_2^{\mathrm{BC}}$	$\hat{\gamma}_1^{\mathrm{BC}}$	$\hat{\gamma}_2^{\mathrm{BC}}$	$\hat{\delta}$	$\hat{\sigma}_{\mathrm{NP}}$	$\hat{\alpha}_1^{\mathrm{NP}}$	$\hat{\alpha}_2^{\mathrm{NP}}$	$\hat{\alpha}_1^{\mathrm{NP}} - \hat{\alpha}_2^{\mathrm{NP}}$	$\hat{\beta}_1^{\mathrm{NP}}$
-10	-0.0207	0.1201	0.1451	-9.2505	-8.4386	-0.8119	1.0317	0.5205	0.6382	0.3672	0.3301	0.1180	-11.8095	-10.9057	-0.9038	1.2598
-9	-0.0227	0.1188	0.1452	-8.3639	-7.5615	-0.8024	1.0198	0.5148	0.6308	0.3631	0.3333	0.1174	-10.8108	-9.9071	-0.9038	1.2598
-8	-0.0251	0.1172	0.1453	-7.4800	-6.6889	-0.7912	1.0058	0.5080	0.6222	0.3582	0.3333	0.1174	-9.8109	-8.9071	-0.9038	1.2598
-7	-0.0280	0.1153	0.1454	-6.5970	-5.8195	-0.7775	0.9886	0.4998	0.6116	0.3523	0.3333	0.1174	-8.8108	-7.9071	-0.9038	1.2597
-6	-0.0317	0.1130	0.1455	-5.7190	-4.9583	-0.7607	0.9679	0.4897	0.5988	0.3451	0.3333	0.1174	-7.8109	-6.9071	-0.9038	1.2598
-5	-0.0365	0.1101	0.1458	-4.8432	-4.1038	-0.7394	0.9414	0.4769	0.5825	0.3360	0.3333	0.1173	-6.8107	-5.9069	-0.9038	1.2597
-4	-0.0429	0.1063	0.1461	-3.9751	-3.2632	-0.7120	0.9074	0.4603	0.5616	0.3243	0.3333	0.1174	-5.8109	-4.9071	-0.9038	1.2597
-3	-0.0518	0.1014	0.1466	-3.1191	-2.4437	-0.6754	0.8624	0.4383	0.5339	0.3088	0.3333	0.1174	-4.8109	-3.9071	-0.9038	1.2597
-2	-0.0647	0.0949	0.1476	-2.2881	-1.6622	-0.6259	0.8020	0.4085	0.4967	0.2880	0.3333	0.1174	-3.8108	-2.9070	-0.9038	1.2598
-1	-0.0834	0.0867	0.1496	-1.5118	-0.9513	-0.5605	0.7238	0.3694	0.4488	0.2614	0.3333	0.1174	-2.8109	-1.9071	-0.9038	1.2598
0	-0.1006	0.0825	0.1547	-0.8808	-0.3744	-0.5063	0.6675	0.3389	0.4149	0.2436	0.3333	0.1174	-1.8108	-0.9070	-0.9038	1.2598
1	0.0072	0.1702	0.1630	-0.6936	0.2643	-0.9579	1.3173	0.6317	0.8193	0.4797	0.3333	0.1174	-0.8108	0.0930	-0.9037	1.2598
2	0.1362	0.3308	0.1510	0.6687	2.7365	-2.0678	2.7320	1.2824	1.6905	0.9643	0.3333	0.1174	0.1891	1.0929	-0.9038	1.2598
3	0.1031	0.2563	0.1460	2.1835	3.8802	-1.6966	2.1854	1.0515	1.3513	0.7688	0.3333	0.1174	1.1891	2.0929	-0.9038	1.2598
4	0.0756	0.2139	0.1447	3.2326	4.6722	-1.4395	1.8392	0.8969	1.1370	0.6476	0.3333	0.1174	2.1891	3.0929	-0.9038	1.2598
5	0.0585	0.1920	0.1443	4.1704	5.4706	-1.3002	1.6558	0.8134	1.0236	0.5837	0.3333	0.1174	3.1891	4.0929	-0.9038	1.2598
6	0.0474	0.1792	0.1442	5.0768	6.2935	-1.2167	1.5472	0.7634	0.9565	0.5459	0.3333	0.1174	4.1891	5.0929	-0.9038	1.2598
7	0.0397	0.1710	0.1441	5.9733	7.1356	-1.1624	1.4769	0.7309	0.9130	0.5215	0.3333	0.1174	5.1891	6.0929	-0.9038	1.2598
8	0.0341	0.1653	0.1441	6.8652	7.9894	-1.1242	1.4278	0.7080	0.8827	0.5044	0.3333	0.1174	6.1891	7.0929	-0.9038	1.2598
9	0.0299	0.1611	0.1441	7.7560	8.8523	-1.0963	1.3920	0.6913	0.8605	0.4920	0.3333	0.1174	7.1891	8.0929	-0.9038	1.2598

Table 5. (Continued.)

Index	$\hat{\beta}_2^{\text{NP}}$	$\hat{\gamma}_1^{\text{NP}}$	$\hat{\gamma}_2^{\text{NP}}$	$\hat{\lambda}_{\text{MNP}}$	$\hat{\sigma}_{\text{MNP}}$	$\hat{\alpha}_1^{\text{MNP}}$	$\hat{\alpha}_2^{\text{MNP}}$	$\hat{\alpha}_1^{\text{MNP}} - \hat{\alpha}_2^{\text{MNP}}$	$\hat{\beta}_1^{\text{MNP}}$	$\hat{\beta}_2^{\text{MNP}}$	$\hat{\gamma}_1^{\text{MNP}}$	$\hat{\gamma}_2^{\text{MNP}}$	MinBC	MinNP	MinMNP
-10	0.7228	0.7051	0.4049	3.6547	0.1269	-11.8279	-10.9257	-0.9022	1.2606	0.7433	0.7025	0.4030	247.54	248.78	248.51
-9	0.7242	0.7048	0.4047	3.6611	0.1268	-10.8280	-9.9258	-0.9022	1.2607	0.7435	0.7025	0.4031	238.55	239.76	239.51
-8	0.7243	0.7047	0.4047	3.6542	0.1269	-9.8277	-8.9254	-0.9023	1.2607	0.7432	0.7023	0.4030	229.57	230.76	230.51
-7	0.7242	0.7048	0.4047	3.6613	0.1268	-8.8280	-7.9258	-0.9022	1.2607	0.7435	0.7025	0.4031	220.59	221.76	221.51
-6	0.7242	0.7048	0.4047	3.6613	0.1268	-7.8280	-6.9258	-0.9022	1.2607	0.7435	0.7025	0.4031	211.62	212.76	212.51
-5	0.7241	0.7048	0.4047	3.6615	0.1268	-6.8280	-5.9258	-0.9022	1.2606	0.7435	0.7025	0.4031	202.66	203.76	203.51
-4	0.7242	0.7048	0.4047	3.6609	0.1268	-5.8280	-4.9258	-0.9022	1.2607	0.7435	0.7025	0.4031	193.72	194.76	194.51
-3	0.7242	0.7047	0.4047	3.6643	0.1267	-4.8281	-3.9259	-0.9022	1.2606	0.7435	0.7024	0.4030	184.82	185.76	185.51
-2	0.7241	0.7048	0.4047	3.6610	0.1268	-3.8280	-2.9258	-0.9022	1.2607	0.7435	0.7026	0.4031	175.99	176.76	176.51
-1	0.7242	0.7048	0.4047	3.6613	0.1268	-2.8280	-1.9258	-0.9022	1.2607	0.7435	0.7025	0.4031	167.34	167.76	167.51
0	0.7242	0.7048	0.4047	3.6476	0.1270	-1.8277	-0.9255	-0.9022	1.2606	0.7432	0.7026	0.4031	159.18	158.76	158.51
1	0.7241	0.7049	0.4047	3.6614	0.1268	-0.8280	0.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	151.37	149.76	149.51
2	0.7242	0.7048	0.4047	3.6617	0.1268	0.1720	1.0741	-0.9022	1.2607	0.7435	0.7025	0.4030	140.64	140.76	140.51
3	0.7242	0.7048	0.4047	3.6614	0.1268	1.1720	2.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	130.74	131.76	131.51
4	0.7242	0.7048	0.4047	3.6612	0.1268	2.1720	3.0741	-0.9021	1.2606	0.7435	0.7026	0.4031	121.49	122.76	122.51
5	0.7242	0.7048	0.4047	3.6613	0.1268	3.1720	4.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	112.41	113.76	113.51
6	0.7242	0.7048	0.4047	3.6611	0.1268	4.1721	5.0742	-0.9021	1.2607	0.7435	0.7025	0.4031	103.38	104.76	104.51
7	0.7242	0.7048	0.4047	3.6608	0.1268	5.1720	6.0743	-0.9022	1.2607	0.7435	0.7026	0.4031	94.37	95.76	95.51
8	0.7242	0.7048	0.4047	3.6614	0.1268	6.1720	7.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	85.37	86.76	86.51
9	0.7242	0.7048	0.4047	3.6613	0.1268	7.1720	8.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	76.37	77.76	77.51
10	0.7239	0.7048	0.4047	3.6612	0.1268	8.1720	9.0742	-0.9022	1.2607	0.7435	0.7025	0.4031	67.37	68.76	68.51

(Note: MinBC, MinNP and MinMNP mean the minimum values of negative log-likelihood functions with respect to BC, NP and MNP families respectively.)

For the transformed data, we perform regression analyses on three families $\text{BC}\left(\lambda, \tau, \xi_{ijk}\right)$, $\text{NP}\left(\delta, \sigma, \eta_{ijk}\right)$ and $\text{MNP}\left(\lambda, \sigma, \eta_{ijk}\right)$. Results are given in Table 5, where estimates of a contrast $\alpha_1 - \alpha_2$ are included. Clearly a change of θ_i with respect to *Index* violates invariance of $\hat{\lambda}_{\text{BC}}$, and produces variations of signs in $\hat{\lambda}_{\text{BC}}$. Small values of $\hat{\lambda}_{\text{BC}}$ suggest that Logarithmic transformation is appropriate, but $\hat{\lambda}_{\text{MNP}}$ is entirely different from $\hat{\lambda}_{\text{BC}}$, and $\hat{\delta}$ indicates another transformation, that is, the Weibull distribution.

The other estimates with $\text{BC}\left(\lambda, \tau, \xi_{ijk}\right)$ change greatly according to θ_i. Furthermore, as for the estimate of contrast $\alpha_1 - \alpha_2$, $\hat{\alpha}_1^{\text{BC}} - \hat{\alpha}_2^{\text{BC}}$ varies greatly according to θ_i, but $\hat{\alpha}_1^{\text{NP}} - \hat{\alpha}_2^{\text{NP}}$ and $\hat{\alpha}_1^{\text{MNP}} - \hat{\alpha}_2^{\text{MNP}}$ are invariant. Also, estimates corresponding to factors B and C with $\text{NP}\left(\delta, \sigma, \eta_{ijk}\right)$ and $\text{MNP}\left(\delta, \sigma, \eta_{ijk}\right)$ are invariant. The ratio $\hat{\tau} / \left\{\sum_{i,j,k=1}^{3}\left(1 + \hat{\lambda}_{\text{BC}}\hat{\xi}_{ijk}\right)^2 / 27\right\}^{1/2}$ with $\hat{\xi}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k$ seems to be constant as expected.

VI. Scale transformations and Tuning Design Matrices

A scale-invariant property of the maximum likelihood estimates $\hat{\delta}$ and $\hat{\sigma}$ for the power parameter δ and σ of NP family does not hold generally in the general regression situation, where there are continuous explanatory variables in design matrices. However, we can examine behaviors of estimates $\hat{\delta}$ and $\hat{\sigma}$ in a different way. In this section we first show that effects of scale transformations with random variables can be attained by tuning design matrices in the current regression problem.

We shall consider M populations $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i = 1,\cdots, M)$, where δ and σ are common, and $\log\eta_i$ has a regression structure such as

$$\log\eta_i = \beta_0 + \beta_1 z_i \quad (i=1,\cdots,M)$$

with population parameters β_0, $\beta_1 (\neq 0)$ and a continuous explanatory variable z.

For BC families $\mathrm{BC}(\lambda,\tau,\xi_i)$ $(i=1,\cdots,M)$, we suppose the following linear regression structure

$$\xi_i = \psi_0 + \psi_1 z_i \quad (i=1,\cdots,M)$$

with population parameters ψ_0 and ψ_1. For NP, MNP families $\mathrm{NP}(\delta,\sigma,\eta_i)$ and $\mathrm{MNP}(\delta,\sigma,\eta_i)$ $(i=1,\cdots,M)$, we suppose the same regression structure of $\mathrm{GF}(\delta,\sigma,\eta_i)$.

Here we note that under a scale transformation $X_i \to \theta_i X_i$ $(\theta_i > 0)$, η_i of $\mathrm{GF}(\delta,\sigma,\eta_i)$ is transformed to $\theta_i\eta_i$, from which we know that $\log\eta_i \to \log(\theta_i\eta_i) = \beta_0 + \beta_1 z_i^*$ $\left(z_i^* = z_i + \log\theta_i / \beta_1\right)$. The effect of a scale transformation $X_i \to \theta_i X_i$ is attained by tuning the value of our explanatory variable z. In the following example, instead of using scale transformations $X_i \to \theta_i X_i$ $(i=1,\cdots,M)$, we vary the value of z_i in $\log\eta_i = \beta_0 + \beta_1 z_i$ $(i=1,\cdots,M)$.

Every time we change z_i in $\log\eta_i = \beta_0 + \beta_1 z_i$ $(i=1,\cdots,M)$, we draw M random samples of sizes n_i from $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i=1,\cdots,M)$ and examine behaviors of estimates $\{\hat{\lambda}_{\mathrm{BC}},\hat{\tau},\hat{\psi}_0,\hat{\psi}_1\}$ of $\prod_{i=1}^{M}\mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$, $\{\hat{\delta},\hat{\sigma}_{\mathrm{NP}},\hat{\beta}_0^{\mathrm{NP}},\hat{\beta}_1^{\mathrm{NP}}\}$ of $\prod_{i=1}^{M}\mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$ and $\{\hat{\lambda}_{\mathrm{MNP}},\hat{\sigma}_{\mathrm{MNP}},\hat{\beta}_0^{\mathrm{MNP}},\hat{\beta}_1^{\mathrm{MNP}}\}$ of $\prod_{i=1}^{M}\mathrm{MNP}(\lambda,\sigma,\eta_i)^{n_i}$.

We give an example on regression analysis, and examine behaviors of $\hat{\lambda}_{\mathrm{BC}}$. Let us consider five populations $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i=1,2,3,4,5)$, of which $\log\eta_i$ has a regression structure such as

$$\begin{pmatrix} \log\eta_1 \\ \log\eta_2 \\ \log\eta_3 \\ \log\eta_4 \\ \log\eta_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta_0 + \begin{pmatrix} 0.001 \\ 0.002 \\ 0.003 \\ 0.004 \\ 0.005 \times Index \end{pmatrix} \beta_1 ,$$

$$\beta_0 = 5, \beta_1 = 0.5$$

with a proportional factor $Index$. The factor $Index$ takes values $\{0,50,100,200,300,400,500\}$.

For the case $\delta = -0.3$, $\sigma = 0.1$ and fixed $Index$, we draw (n+1)-quantiles of size $n = 6$ from each of $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i = 1,2,3,4,5)$.

According to each $Index$, using five samples from $\mathrm{GF}(\delta,\sigma,\eta_i)$ $(i = 1,2,3,4,5)$ mentioned above, we perform regression analyses on families $\prod_{i=1}^{5} \mathrm{NP}(\delta,\sigma,\eta_i)^{n_i}$, $\prod_{i=1}^{5} \mathrm{MNP}(\delta,\sigma,\eta_i)^{n_i}$ and $\prod_{i=1}^{5} \mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$ $(n_1 = n_2 = \cdots = n_5 = 6)$. Here we suppose that ξ_i has a regression structure such as

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \psi_0 + \begin{pmatrix} 0.001 \\ 0.002 \\ 0.003 \\ 0.004 \\ 0.005 \times Index \end{pmatrix} \psi_1 .$$

Results are given in Table 6.

Table 6. Example on behaviors of power parameter estimates and the other estimates in BC, NP and NMP families for the case of tuning design matrices.

Index	$\hat{\lambda}_{BC}$	$\hat{\tau}$	$\hat{\psi}_0$	$\hat{\psi}_1$	$\hat{\delta}$	$\hat{\sigma}_{NP}$	$\hat{\beta}_0^{NP}$	$\hat{\beta}_1^{NP}$	$\hat{\lambda}_{MNP}$	$\hat{\sigma}_{MNP}$	$\hat{\beta}_0^{MNP}$	$\hat{\beta}_1^{MNP}$	MinBC	MinNP	MinMNP
0	-2.142	1.64E-06	0.467	1.04E-05	-0.165	0.0769	5.022	0.498	-2.158	0.0790	5.034	0.482	118.13	118.13	118.13
50	-1.187	0.0002	0.841	0.0012	-0.165	0.0769	5.022	0.500	-2.141	0.0790	5.034	0.500	119.06	118.88	118.88
100	-0.500	0.0064	1.840	0.0379	-0.165	0.0769	5.022	0.500	-2.142	0.0790	5.034	0.500	119.93	119.63	119.63
200	-0.147	0.0382	3.564	0.2304	-0.165	0.0769	5.022	0.500	-2.143	0.0790	5.034	0.500	121.49	121.13	121.13
300	-0.067	0.0574	4.281	0.3485	-0.165	0.0769	5.022	0.500	-2.140	0.0790	5.034	0.500	123.00	122.63	122.63
400	-0.038	0.0666	4.591	0.4057	-0.165	0.0769	5.022	0.500	-2.145	0.0790	5.034	0.500	124.51	124.13	124.13
500	-0.024	0.0714	4.746	0.4360	-0.165	0.0769	5.022	0.500	-2.142	0.0790	5.034	0.500	126.01	125.63	125.63

(Note: The true densities are $\prod_{i=1}^{5} \mathrm{GF}(\delta,\sigma,\eta_i)^{n_i}$ with $n_1 = n_2 = \cdots = n_5 = 6$, $\delta = -0.3$, $\sigma = 0.1$, $\log\eta_i = \beta_0 + \beta_1 z_i$ $(i = 1,2,\cdots,5), \beta_0 = 5, \beta_1 = 0.5,$ $(z_1,z_2,z_3,z_4,z_5) = (0.001,0.002,0.003,0.004,0.005 \times Index)$ $(Index = 0,50,100,200,300,400,500)$, and $\xi_i = \psi_0 + \psi_1 z_i$ $(i = 1,2,\cdots,5)$ for $\prod_{i=1}^{5} \mathrm{BC}(\lambda,\tau,\xi_i)^{n_i}$.)

Clearly a change of *Index* causes large fluctuations of $\hat{\lambda}_{\mathrm{BC}}$. Values of $\hat{\lambda}_{\mathrm{BC}}$ corresponding to large *Index* suggest that Logarithmic transformation is appropriate. Estimates $\hat{\psi}_0$ and $\hat{\psi}_1$ are seriously unstable. On the other hand, $\hat{\beta}_0{}^{\mathrm{NP}}$ and $\hat{\beta}_1{}^{\mathrm{NP}}$ seem to be stable.

An implication of this example is serious, because this example shows that a change of design matrices in regression analysis brings entirely different estimation results with respect to BC family. On the other hand, NP and MNP families always produce stable estimation results regardless of a change of design matrices.

VII. Conclusions

Through many concrete examples we have examined behaviors of the maximum likelihood estimate $\hat{\lambda}_{\mathrm{BC}}$ for the power parameter λ of the Box-Cox power normal family under scale transformations of the relevant random variables. Our summary is as follows.

(1) One-sample problem. The maximum likelihood estimate $\hat{\lambda}_{\mathrm{BC}}$ is invariant under a scale transformation of the random variable. However, there is some functional relationship among parameters, including λ, of the Box-Cox power normal family, which causes not only large fluctuations of the other estimates except for $\hat{\lambda}_{\mathrm{BC}}$ but also a large prediction error with BC family in the meaning of bootstrap generalized information criteria.

(2) Two-sample problem. The maximum likelihood estimate $\hat{\lambda}_{\mathrm{BC}}$ is not invariant under individual scale transformations of each random variable. A tendency of Logarithmic transformation in the Box-Cox power normal family often appears when there are large differences between variances of two populations.

(3) Multi-sample problem, especially, multifactor designs. The maximum likelihood estimate $\hat{\lambda}_{\mathrm{BC}}$ is not invariant under individual scale transformations of each random variable. Estimates of parameters, including λ, of the Box-Cox power normal family move violently under the above scale transformations.

(4) The general regression situation. Changing a design matrix causes large fluctuations of estimates of λ and the other parameters in the Box-Cox power normal family. An implication of this result is serious, because it means that alterations to a design matrix bring us entirely different estimates under the same experimental situation.

On the other hand, the maximum likelihood estimates $\hat{\delta}$ and $\hat{\sigma}$ for δ and σ of our new power normal family are invariant under scale transformations of the relevant random variables in cases of one-sample, two-sample and multi-sample problems.

In the general regression situation changes of a design matrix do not give any critical effect to estimate $\hat{\delta}$ and $\hat{\sigma}$ of our new power normal family. Estimates $\hat{\delta}$ and $\hat{\sigma}$ seem to be stable.

Both parameters δ and σ have a direct relationship $\lambda = \delta / \sigma$ for the power parameter λ of the Box-Cox power normal family. Thus, when we are interested in λ of the Box-Cox power normal family, our modified new power normal family is useful to estimate λ, because the maximum likelihood estimate $\hat{\lambda}_{\mathrm{MNP}}$ for the power parameter λ of our modified new power normal family is invariant under scale transformations of the relevant random variables in cases of one-sample, two-sample and multi-sample problems. In the general regression situation $\hat{\lambda}_{\mathrm{MNP}}$ also remains stable under alterations to a design matrix in the same experiment.

References

[1] G.E.P. Box and D.R. Cox, *An analysis of transformations*, J. R. Stat. Soc. B 26(1964), pp.211-252.

[2] N.R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed., Wiley-Interscience, NewYork, 1998.

[3] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, 1993.

[4] T. Isogai, *Power transformation of the F distribution and a power normal family*, J. Appl. Stat. 26 (1999), pp.355-371.

[5] T. Isogai, *Applications of power transformation formulas of the F distribution and a power normal family,* Journal of the Japanese Society for Quality Control, 31(2001), pp.89-104, (in Japanese).

[6] T. Isogai, *Applications of a new power normal family,* J. Appl. Stat. 32 (2005), pp.421-436.

[7] T. Isogai, H. Uchida, S. Miyama and S. Nishiyama, *Statistical modeling of enamel rater value data*, J. Appl. Stat. 35(2008), pp.515-535.

[8] T. Isogai, *Applications of a New Power Normal Family – A Case Study on* C_{pk} *Index* -, Rev. Grad. School Mar. Sci., Kobe Univ. 7(2010), pp.39-52.

[9] N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous univariate distributions*, Vol.2, 2nd ed., Wiley-Interscience, New York, 1995.

[10] S. Konishi and G. Kitagawa, *Information Criteria and Statistical Modeling*, Springer Series in Statistics, 2007.

[11] S. Kullback, *Information Theory and Statistics*, Dover, 1968.

[12] R. L. Prentice, *Discrimination among some parametric models*, Biometrika 62(1975), pp.607-614.