THE BIVARIATE POWER-NORMAL DISTRIBUTION

Goto, Masashi
Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University

Hamasaki, Toshimitsu
Biostatistics, Pfizer Global Research and Development, Tokyo Laboratories, Pfizer Pharmaceuticals Inc.,

https://doi.org/10.5109/13507

出版情報：Bulletin of informatics and cybernetics. 34 (1), pp.29-49, 2002-10. Research Association of Statistical Sciences
バージョン：

Bulletin of Informatics and Cybernetics, Vol. 34, No. 1, 2002

THE BIVARIATE POWER-NORMAL DISTRIBUTION

By

Masashi GOTO* and Toshimitsu HAMASAKI†

Abstract

In this paper, the earlier work on so-called "power-normal distribution" is extended to a multivariate case, especially focusing on a bivariate one. The power-normal distribution is a family of distributions including the truncated normal and the lognormal. The present work introduces the moments and other related properties of the bivariate power-normal distribution. The numerical illustrations are provided to demonstrate the elements and the applications of the distribution.

Key Words and Phrases: Power-normal distribution; Truncated bivariate normal distribution; Multivariate skewness; Multivariate kurtosis.

1. Introduction

"Power-normal distribution" is a parametric class of probability distributions which includes the truncated normal and the lognormal as a special case. The power-normal distribution is on the basis of the Box and Cox power-transformation which is defined by, for a positive random variable X

$$X^{(\lambda)} = \begin{cases} \dfrac{X^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log X, & \lambda = 0 \end{cases} \tag{1.1}$$

where λ is the shape parameter (or the transformation parameter) and is chosen as a power-transformed variable $X^{(\lambda)}$ has the normal (Box and Cox, 1964). Unfortunately $X^{(\lambda)}$ lies in lower or upper bounded region according to $\lambda > 0$ or $\lambda < 0$. Therefore, $X^{(\lambda)}$ has the truncated normal except for $\lambda = 0$. X is then said to have the power-normal distribution, written $X \sim \mathrm{PN}(\lambda, \mu, \sigma^2)$ if $X^{(\lambda)}$ has the truncated normal distribution with mean μ and variance σ^2 (Goto Matsubara and Tsuchiya, 1983: Johnson, Kotz and Balakrishnan, 1994). Its probability density function (pdf) is given by

$$g(x : \lambda, \mu, \sigma) = \frac{x^{\lambda-1}}{\sigma A(\lambda, \mu, \sigma)} \phi\left(\frac{x^{(\lambda)} - \mu}{\sigma}\right), \quad x > 0 \tag{1.2}$$

* Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University, 1–3 Machikaneyama, Osaka 560-8531 Japan. E-mail: gotoo@sigmath.es.osaka-u.ac.jp

† Biostatistics, Pfizer Global Research and Development, Tokyo Laboratories, Pfizer Pharmaceuticals Inc., P.O.Box 226 Mitsui Bldg., 2-1-1 Nishi-Shinjuku, Shinjuku, Tokyo 163-0461, Japan. E-mail: Toshimitsu.Hamasaki@japan.pfizer.com

where $\phi(\cdot)$ denotes the pdf of standard normal distribution $\mathrm{N}(0,1)$ and

$$A(\lambda,\mu,\Sigma)=\begin{cases}\Phi(k), & \lambda>0,\\ 1, & \lambda=0,\\ \Phi(-k), & \lambda<0\end{cases} \tag{1.3}$$

where $\Phi(\cdot)$ denotes the distribution function of standard normal distribution and k is the standardized truncation point of the truncated normal for the power-transformed variable $X^{(\lambda)}$, which is given by $k=(\lambda\mu+1)/\lambda\sigma$. With the shape parameter λ , the power-normal distribution coincides with the truncated normal for $\lambda=1$ and with the lognormal if $\lambda=0$. The systematic developments of this distribution have been given by Goto, Uesaka and Inoue (1979), Goto and Inoue (1980), Uesaka and Goto (1980, 1982), Goto *et al.* (1983), Goto, Inoue and Tsuchiya (1984) and Goto, Yamamoto and Inoue (1991).

The purpose of the present paper is to introduce a multivariate version of the power-normal distribution, especially focus on a bivariate case. Such an extension is potentially relevant for practical applications since in the multivariate case there are far fewer distributions available for dealing non-normal data than the univariate case. For example, in medical fields, to evaluate whether there are any effects of treatment on blood pressures for patients with hypertension, the two measure of blood pressures, systolic and diastolic blood pressures are usually observed. Then, the systolic blood pressure is said to be a lognormal, and for the diastolic, it has a normal. Usually the transformation is performed on each component separately, and achievement of joint normality is expected. However, in such a situation, the joint transformation may be more suitable to describe the data, and then the joint distribution which can deal the non-normal data allowing the correlation between two measures should be considered.

The paper is structured as follows: In Section 2, definition, basic properties of the bivariate power-normal distribution and its moments are given. In addition, some properties of the bivariate power-transformed distribution are discussed. In Section 3, the computational algorithm for estimating parameters of the distribution is described. In Section 4, the numerical illustrations are provided to demonstrate the elements and the applications of the distribution. Finally, in Section 5, a multivariate power-normal distribution is introduced, and further developments and applications to practical fields are considered.

2. The Bivariate Power-Normal Distribution

2.1. Definition

In this section, for the two-dimensional extension of (1.2), we consider a positive random variable $X=(X_1,X_2)^{\mathrm{T}}$, where $X_1,X_2>0$.

Let a power-transformed variable $X^{(\lambda)}=(X_1^{(\lambda_1)},X_2^{(\lambda_2)})^{\mathrm{T}}$ of $X=(X_1,X_2)^{\mathrm{T}}$ be the truncated bivariate normal distribution with mean vector $\mu=(\mu_1,\mu_2)^{\mathrm{T}}$ and variance covariance matrix

$$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$$

where ρ is the correlation coefficient between $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$. $X=(X_1,X_2)^{\mathrm{T}}$ is then said to have the bivariate power-normal distribution if the joint pdf is

$$g(x_1, x_2) = \frac{x_1^{\lambda_1 - 1} x_2^{\lambda_2 - 1}}{A(\lambda, \mu, \Sigma)} f(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}), \quad x_1, x_2 > 0 \tag{2.1}$$

where

$$f(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)})}{2}\right\}$$

and

$$Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) = \frac{1}{1-\rho^2} \times \left\{ \left(\frac{x_1^{(\lambda_1)} - \mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1^{(\lambda_1)} - \mu_1}{\sigma_1}\right)\left(\frac{x_2^{(\lambda_2)} - \mu_2}{\sigma_2}\right) + \left(\frac{x_2^{(\lambda_2)} - \mu_2}{\sigma_2}\right)^2 \right\}$$

where the truncated proportional constant term $A(\lambda, \mu, \Sigma)$ is given by

$$A(\lambda, \mu, \Sigma) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} \phi_2(x_1, x_2) dx_1 dx_2 \tag{2.2}$$

in terms of the joint pdf of the bivariate standard normal distribution [1]

$$\phi_2(x_1, x_2 : \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho)^2}\right\}$$

with the values of a_j and b_j given in Table 1, and the standardized truncation point k_j is given by

$$k_j = \frac{\lambda_j \mu_j + 1}{\lambda_j \sigma_j}, \quad j = 1, 2 \ .$$

Table 1. The values of λ_1, λ_2, a_1, b_1, a_2 and b_2

λ_1	λ_2	a_1	b_1	a_2	b_2
	$\lambda_2 < 0$			$-\infty$	$-k_2$
$\lambda_1 < 0$	$\lambda_2 = 0$	$-\infty$	$-k_1$	$-\infty$	∞
	$\lambda_2 > 0$			$-k_2$	∞
	$\lambda_2 < 0$			$-\infty$	$-k_2$
$\lambda_1 = 0$	$\lambda_2 = 0$	$-\infty$	∞	$-\infty$	∞
	$\lambda_2 > 0$			$-k_2$	∞
	$\lambda_2 < 0$			$-\infty$	$-k_2$
$\lambda_1 > 0$	$\lambda_2 = 0$	$-k_1$	$-\infty$	$-\infty$	∞
	$\lambda_2 > 0$			$-k_2$	∞

[1] In general, a bivariate standard distribution is $N_2(0, I)$ with variance-covariance matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

but in this paper it is $N_2(0, I)$ with variance-covariance matrix

$$I^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Then, the magnitude of $A(\lambda, \mu, \Sigma)$ can be evaluated by using the terms of the bivariate standard normal distribution function $\Phi_2(x_1, x_2 : \rho)$ and the univariate standard normal distribution function, as shown in Table 2.

For a univariate case, the power-normal distribution has the six typical shapes corresponding to the value of λ, i.e., $\lambda > 1$(J-shape distribution), $\lambda = 0$ (truncated normal distribution), $c < \lambda < 1$, $\lambda = c$ (exponential-shape distribution), $0 < \lambda < c$ (exponential-shape distribution) and $\lambda < 0$ (L-shape distribution), where $c = 4/(k^2 + 4)$ (Goto *et al.*, 1983). From the analogy of a univariate case, the bivariate power-normal distribution may include various shapes of distributions with the combinations of λ_1 and λ_2. Figures 1(a) and 1(b) provide contour plots of the various bivariate densities with the combinations of λ_1 and λ_2, where $k_1 = k_2 = 3$, $\tau_1 = \tau_2 = 2$ and ρ is equal to 0.3 and 0.9, respectively, where τ_1 and τ_2 are the coefficients of variation for $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$, respectively.

Table 2. The relationships among $\Phi_2(x_1, x_2 : \rho)$, $\Phi(x)$ and $A(\lambda, \mu, \Sigma)$

λ_1	λ_2	$A(\lambda, \mu, \Sigma)$
	$\lambda_2 < 0$	$\Phi_2(-k_1, -k_2 : \rho)$
$\lambda_1 < 0$	$\lambda_2 = 0$	$\Phi(-k_1)$
	$\lambda_2 > 0$	$\Phi(-k_1) - \Phi_2(-k_1, -k_2 : \rho)$
	$\lambda_2 < 0$	$1 - \Phi(k_2)$
$\lambda_1 = 0$	$\lambda_2 = 0$	1
	$\lambda_2 > 0$	$\Phi(k_2)$
	$\lambda_2 < 0$	$\Phi(k_1) - \Phi_2(k_1, k_2 : \rho)$
$\lambda_1 > 0$	$\lambda_2 = 0$	$\Phi(k_1)$
	$\lambda_2 > 0$	$\Phi_2(k_1, k_2 : \rho)$

2.2. The Marginal and the Conditional Distribution

In this section, we discuss the marginal and the conditional distributions of the bivariate power-normal distribution.

Let $g_j(x_j)$ denote the pdf of the univariate power-normal distribution for each $X_j (j = 1, 2)$. $g_j(x_j)$ is then

$$g_j(x_j : \lambda_j, \mu_j, \sigma_j) = \frac{x_j^{\lambda_j - 1}}{\sigma_j A_j(\lambda_j, \mu_j, \sigma_j)} \phi\left(\frac{x_j^{(\lambda_j)} - \mu_j}{\sigma_j}\right) \tag{2.3}$$

where

$$A_j(\lambda_j, \mu_j, \sigma_j) = \begin{cases} \Phi(k_j), & \lambda > 0, \\ 1, & \lambda = 0, \\ \Phi(-k_j), & \lambda < 0 \end{cases}$$

if $\rho = 0$, the density function (2.1) can be written by $g(x_1, x_2) = g_1(x_1) g_2(x_2)$ as $A(\lambda, \mu, \Sigma)$ is $A(\lambda, \mu, \Sigma) = A_1(\lambda_1, \mu_1, \sigma_1) A_2(\lambda_2, \mu_2, \sigma_2)$. By the definition of the bivariate power-normal distribution (4) for (X_1, X_2), after some simple algebra, the pdf of the marginal distribution for X_1 is given by

$$g_1(x_1) = \begin{cases} \dfrac{x_1^{\lambda_1-1}}{A(\lambda,\mu,\Sigma)}\phi\left(\dfrac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right)\Phi\left[\dfrac{\operatorname{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\dfrac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}+k_2\right)\right], & \lambda_1 \neq 0, \lambda_2 \neq 0 \\ \dfrac{x_1^{\lambda_1-1}}{A(\lambda,\mu,\Sigma)}\phi\left(\dfrac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right), & \lambda_1 \neq 0, \lambda_2 = 0 \\ \dfrac{1}{x_1}\phi\left(\dfrac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right), & \lambda_1 = 0, \lambda_2 = 0. \end{cases} \tag{2.4}$$

Therefore, by comparing (2.3) with (2.4), it is clear that the density of the marginal distribution for the bivariate power-normal distribution is not consistent with that for the univariate power-normal. However, only if $A(\lambda, \mu, \Sigma) = 1$, the pdfs (2.3) and (2.4) have the same form of density. Similarly, the pdf of the marginal distribution for is given by X_2 is given by

$$g_2(x_2) = \begin{cases} \dfrac{x_2^{\lambda_2-1}}{A(\lambda,\mu,\Sigma)}\phi\left(\dfrac{x_2^{(\lambda_2)}-\mu_2}{\sigma_2}\right)\Phi\left[\dfrac{\operatorname{sgn}(\lambda_1)}{\sqrt{1-\rho^2}}\left(\rho\dfrac{x_2^{(\lambda_2)}-\mu_2}{\sigma_2}+k_1\right)\right], & \lambda_2 \neq 0, \lambda_1 \neq 0 \\ \dfrac{x_2^{\lambda_2-1}}{A(\lambda,\mu,\Sigma)}\phi\left(\dfrac{x_2^{(\lambda_2)}-\mu_2}{\sigma_2}\right), & \lambda_2 \neq 0, \lambda_1 = 0 \\ \dfrac{1}{x_2}\phi\left(\dfrac{x_2^{(\lambda_2)}-\mu_2}{\sigma_2}\right), & \lambda_2 = 0, \lambda_1 = 0. \end{cases}$$

Next we consider the conditional distribution and regression of the bivariate power-normal distribution. By (2.1) and (2.4), for some λ_1, the pdf of conditional distribution of X_2 given $X_1 = x_1$, is give by

$$g(x_2\,|x_1) = \begin{cases} \dfrac{x_2^{\lambda_2-1} f(x_2^{(\lambda_2)}\big| x_1^{(\lambda_1)})}{\Phi\left[\dfrac{\operatorname{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\dfrac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}+k_2\right)\right]}, & \lambda_2 \neq 0 \\ \dfrac{1}{x_2} f(x_2^{(\lambda_2)}\big| x_1^{(\lambda_1)}), & \lambda_2 = 0 \end{cases} \tag{2.5}$$

where $f(x_2^{(\lambda_2)}\big| x_1^{(\lambda_1)})$ is the conditional pdf of $X_2^{(\lambda_2)}$ given $X_1^{(\lambda_1)} = x_1^{(\lambda_1)}$ in which $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$ has the bivariate normal distribution, that is

$$f(x_2^{(\lambda_2)}\big| x_1^{(\lambda_1)}) = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\sigma_2^2(1-\rho^2)}\left\{x_2^{(\lambda_2)}-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x_1^{(\lambda_1)}-\mu_1)\right\}^2\right].$$

Therefore, for $\lambda_2 \neq 0$, the conditional expectation of X_2 given $X_1 = x_1$ is given by

$$\mathrm{E}[X_2 \,|X_1 = x_1] = \frac{\int_0^\infty x_2^{\lambda_2} f(x_2^{(\lambda_2)} \big| x_1^{(\lambda_1)}) dx_2}{\Phi\left[\frac{\mathrm{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}+k_2\right)\right]}. \tag{2.6}$$

In particular, for $\lambda_2 > 0$, it becomes

$$\begin{aligned}&\mathrm{E}[X_2 \,|X_1 = x_1] \\ &= C_0 \sum_{v=0}^{\infty} \frac{(\sqrt{2})^{p+v-1}}{v!}(1-\rho^2)^{(p-v)/2}\left(k_2+\rho\frac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right)\Gamma\left(\frac{p+v+1}{2}\right)\end{aligned} \tag{2.7}$$

where $p = 1/\lambda_2$ and

$$\begin{aligned}C_0 = \frac{(\lambda_2\sigma_2)^p}{\sqrt{2}} &\exp\left\{-\frac{1}{2(1-\rho^2)}\left(k_2+\rho\frac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right)\right\} \\ &\times\left\{\Phi\left[\frac{1}{\sqrt{1-\rho^2}}\left(k_2+\rho\frac{x_1^{(\lambda_1)}-\mu_1}{\sigma_1}\right)\right]\right\}^{-1}.\end{aligned}$$

Then, the conditional expectation (2.7) provides the regression function of X_2 on X_1 which (X_1, X_2) has the bivariate power-normal distribution.

2.3. The Moments and Other Properties

Here we further discuss the moments and other properties of the bivariate power-normal distribution.

For $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$, the joint moment about the origin of order (m_1, m_2) of variable (X_1, X_2) is defined by

$$\mathrm{E}[X_1^{m_1}X_2^{m_2}] = \int_0^\infty \int_0^\infty \frac{x_1^{m_1+\lambda_1-1}x_2^{m_2+\lambda_2-1}}{A(\lambda,\mu,\Sigma)} f(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) dx_1 dx_2. \tag{2.8}$$

For $\lambda_1 > 0$ and $\lambda_2 > 0$, the joint moment (2.8) can be written in another form

$$\mathrm{E}[X_1^{m_1}X_2^{m_2}] = C(m_1, m_2) \sum_{\nu=0}^{\infty} \frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(p_1, a_1 : \rho) S_\nu(p_2, a_2 : \rho) \tag{2.9}$$

where

$$C(m_1, m_2) = \frac{(\lambda_1\sigma_1)^{p_1}}{(\lambda_2\sigma_2)^{p_2}}\phi_2(k_1, k_2 : \rho),$$

$$S_\nu(p_j, a_j : \rho) = \int_0^\infty \exp\left\{\frac{\nu^2 - 2a_j\nu}{2(1-\rho^2)}\right\} d\nu, \quad j = 1, 2,$$

$p_j = m_j/\lambda_j$, $a_1 = k_1 - \rho k_2$ and $a_2 = k_2 - \rho k_1$. From (2.9), the variance, covariance and correlation coefficient between X_1 and X_2 of the bivariate power-normal distribution can be numerically assessed. In addition, the numerical values for the skewness and kurtosis of the bivariate power-normal distribution can be obtained. When actually assessing the numerical values of the above statistics, (2.9) should be broken down into the following three terms

$$\begin{aligned} &\mathrm{E}\left[X_1^{m_1} X_2^{m_2}\right] = C(m_1, m_2) \\ &\times \left\{ S_0(p_1, a_1 : \rho) S_0(p_2, a_2 : \rho) + \sum_{l=1}^{\infty} \frac{1}{(2l)!} \left(\frac{\rho}{1-\rho^2}\right)^{2l} S_{2l}(p_1, a_1 : \rho) S_{2l}(p_2, a_2 : \rho) \right. \\ &\left. + \sum_{l=1}^{\infty} \frac{1}{(2l+1)!} \left(\frac{\rho}{1-\rho^2}\right)^{2l+1} S_{2l+1}(p_1, a_1 : \rho) S_{2l+1}(p_2, a_2 : \rho) \right\} \end{aligned} \tag{2.10}$$

and then some numerical calculations are needed for each the three terms. See Appendix 1 for the details.

For $\lambda_1 < 0$ and $\lambda_2 > 0$, the joint moment about the origin of order (m_1, m_2) of variable (X_1, X_2) can be written by

$$\mathrm{E}\left[X_1^{m_1} X_2^{m_2}\right] = \frac{(\eta_1\sigma_1)^{-q_1}(\eta_2\sigma_2)^{-q_2}}{A(\lambda, \mu, \Sigma)} \int_{-\infty}^{k_2}\int_{-\infty}^{k_1} \frac{\phi_2(x_1, x_2 : \rho)}{(k_1 - x_1)^{q_1}(k_2 - x_2)^{q_2}} dx_1 dx_2$$

where $\eta_j = -\lambda_j$ and $q_j = -p_j (j = 1, 2)$. In particular, if $q_j > 0$, the following inequality hold.

$$\int_{-\infty}^{k_2}\int_{-\infty}^{k_1} \frac{\phi_2(x_1, x_2 : \rho)}{(k_1 - x_1)^{q_1}(k_2 - x_2)^{q_2}} dx_1 dx_2 \le \int_{-\infty}^{k_2}\int_{-\infty}^{k_1} \frac{dx_1 dx_2}{(k_1 - x_1)^{q_1}(k_2 - x_2)^{q_2}} \tag{2.11}$$

Then, all joint moments exist as the right side of inequality (2.11) converges absolutely if $q_1 > 0$ and $q_2 > 0$. However, except for $q_1 > 0$ and $q_2 > 0$ it diverges. Hence, the joint moment with degree less than $|\lambda_1|$ and $|\lambda_2|$ when $\lambda_1 < 0$ or $\lambda_2 < 0$.

2.4. Some Properties for the Bivariate Power-Transformed Distribution

As described in Section 2.1, each power-transformed variable $X_j^{(\lambda_j)} (j = 1, 2)$ lies in $-1/\lambda_j < X_j^{(\lambda_j)} < \infty$ if $\lambda_j > 0$, otherwise X_j lies in $-\infty < X_j^{(\lambda_j)} < -1/\lambda_j$ if $\lambda_j < 0$ under the condition of $X_j > 0$. Then, the bivariate power-transformed variable $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$ has a truncated bivariate normal distribution, and its joint pfd of $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$ is defined by

$$h(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) = \frac{f(x_1^{(\lambda_1)}, x_2^{(\lambda_2)})}{A(\lambda, \mu, \Sigma)} \tag{2.12}$$

and the marginal pfds for each $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$ are given by

$$h_1(x_1^{(\lambda_1)}) = \frac{f_1(x_1^{(\lambda_1)})}{A(\lambda, \mu, \Sigma)} \Phi\left[\frac{\mathrm{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_1^{(\lambda_1)} - \mu_1}{\sigma_1} + k_2\right)\right]$$

and

$$h_2(x_2^{(\lambda_2)}) = \frac{f_2(x_2^{(\lambda_2)})}{A(\lambda, \mu, \Sigma)} \Phi\left[\frac{\mathrm{sgn}(\lambda_1)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_2^{(\lambda_2)} - \mu_2}{\sigma_2} + k_1\right)\right]$$

respectively. For $\lambda_2 \neq 0$, the conditional pfd of $X_2^{(\lambda_2)}$ given $X_1^{(\lambda_1)} = x_1^{(\lambda_1)}$ is obtained by

$$h_2(x_2^{(\lambda_2)} \Big| X_1^{(\lambda_1)} = x_1^{(\lambda_1)}) = \frac{f_2(x_2^{(\lambda_2)} \Big| x_1^{(\lambda_1)})}{\Phi\left[\frac{\mathrm{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_1^{(\lambda_1)}-\mu_2}{\sigma_1}+k_2\right)\right]}. \tag{2.13}$$

Form the conditional pdf (2.13), for $\lambda_2 \neq 0$, the conditional expectation of $X_2^{(\lambda_2)}$ given $X_1^{(\lambda_1)} = x_1^{(\lambda_1)}$ can be written by

$$\begin{aligned} \mathrm{E}\left[X_2^{(\lambda_2)} \Big| X_1^{(\lambda_1)} = x_1^{(\lambda_1)}\right] &= \mu_2 + \rho\frac{\sigma_1}{\sigma_2}(x_1^{(\lambda_1)} - \mu_1) \\ &+ \frac{\sigma_2\sqrt{1-\rho^2}\phi\left[\frac{\mathrm{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_1^{(\lambda_1)}-\mu_2}{\sigma_1}+k_2\right)\right]}{\Phi\left[\frac{\mathrm{sgn}(\lambda_2)}{\sqrt{1-\rho^2}}\left(\rho\frac{x_1^{(\lambda_1)}-\mu_2}{\sigma_1}+k_2\right)\right]}. \end{aligned} \tag{2.14}$$

Then, the conditional expectation (2.14) provides the regression function of $X_2^{(\lambda_2)}$ on $X_1^{(\lambda_1)}$ in which bivariate power-transformed variable $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$ has the truncated bivariate normal. Similarly as Section 2.3, the joint moment about the origin of order (m_1, m_2) of $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$ can be written by

$$\begin{aligned} \mathrm{E}\left[\left(X_1^{(\lambda_1)}\right)^{m_1}\left(X_2^{(\lambda_2)}\right)^{m_2}\right] &= C'(m_1, m_2) \\ &\times \sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(m_1, \theta_1 : \rho)S_\nu(m_2, \theta_2 : \rho) \end{aligned} \tag{2.15}$$

where

$$C'(m_1, m_2) = \frac{\sigma_1^{m_1}\sigma_2^{m_2}}{2\pi\sqrt{1-\rho^2}A(\lambda, \mu, \Sigma)}\phi_2\left(\frac{\mu_1}{\sigma_1}, \frac{\mu_2}{\sigma_2} : \rho\right),$$

$$S_\nu(m_1, \theta_1 : \rho) = \int_{a_1'}^{b_1'} v_1^{m_1+\nu}\exp\left\{\frac{v_1^2 - 2\theta_2 v_1}{2(1-\rho^2)}\right\}dv_1,$$

$$S_\nu(m_2, \theta_2 : \rho) = \int_{a_2'}^{b_2'} v_2^{m_2+\nu}\exp\left\{\frac{v_2^2 - 2\theta_2 v_2}{2(1-\rho^2)}\right\}dv_2, \text{and}$$

$$\theta_1 = \frac{\mu_1}{\sigma_1} - \rho\frac{\mu_2}{\sigma_2}, \quad \theta_2 = \frac{\mu_2}{\sigma_2} - \rho\frac{\mu_1}{\sigma_1}$$

where a_1', b_1', a_2' and b_2' are given in Table 3. Also, c_1 and c_2 in Table 3 are $c_1 = |1/\lambda_1|$ and $c_2 = |1/\lambda_2|$ respectively. Actually, the complete and incomplete gamma functions are needed to calculate the values of $S_j(m_j, \theta_j : \rho)$. See Appendix 2 for the details.

Table 3. The values of a_1', b_1', a_2' and b_2'

λ_1	λ_2	a_1	b_1	a_2	b_2
$\lambda_1 < 0$	$\lambda_2 < 0$	$-\infty$	c_1/σ_1	$-\infty$	c_2/σ_2
	$\lambda_2 > 0$	$-\infty$	c_1/σ_1	c_2/σ_2	∞
$\lambda_1 > 0$	$\lambda_2 < 0$	$-c_1/\sigma_1$	∞	$-\infty$	c_2/σ_2
	$\lambda_2 > 0$	$-c_1/\sigma_1$	∞	c_2/σ_2	∞

As defined in Section 2.1, ρ is the correlation coefficient of the bivariate power-transformed variable $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$. However, since $(X_1^{(\lambda_1)}, X_2^{(\lambda_2)})$ has the truncated bivariate normal distribution, if allowing the truncation, the form of ρ becomes more complicated. If we denote the correlation coefficient allowing the truncation by ρ^*, ρ^* is defined by

$$\rho^* = \frac{\mathrm{cov}\left[X_1^{(\lambda_1)}, X_2^{(\lambda_2)}\right]}{\sqrt{\mathrm{var}\left[X_1^{(\lambda_1)}\right]}\sqrt{\mathrm{var}\left[X_2^{(\lambda_2)}\right]}} = \frac{\mu'_{11} - \mu'_{10}\mu'_{01}}{\sqrt{\mu'_{20} - \mu'^2_{10}}\sqrt{\mu'_{02} - \mu'^2_{01}}} \tag{2.16}$$

where

$$\mu'_{11} = C_0(1,1)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(1,\theta_1,\rho)S_\nu(1,\theta_2,\rho),$$

$$\mu'_{10} = C_0(1,0)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(1,\theta_1,\rho)S_\nu(0,\theta_2,\rho),$$

$$\mu'_{01} = C_0(0,1)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(0,\theta_1,\rho)S_\nu(1,\theta_2,\rho),$$

$$\mu'_{20} = C_0(2,0)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(2,\theta_1,\rho)S_\nu(0,\theta_2,\rho),$$

$$\mu'_{02} = C_0(0,2)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(0,\theta_1,\rho)S_\nu(2,\theta_2,\rho), \text{and}$$

$$C_0(m_1,m_2) = \frac{\sigma_1^{m_1}\sigma_2^{m_2}}{2\pi A(\lambda,\mu,\Sigma)\sqrt{1-\rho^2}}.$$

Therefore ρ^* is slightly different from the correlation coefficient between $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$, that is ρ. Figure 2 shows the relationships between ρ^* and ρ for the various shapes of $\lambda_1 = \lambda_2$ when $k_1 = k_2 = 1$ and $\tau_1 = \tau_2 = 2, 4, 16$. It is clear form the figure that ρ^* has a smaller value compared with that of ρ.

If the measures proposed by Mardia (1970) are used to assess a bivariate normality of the bivariate power-transformed distribution, written by $X^{(\lambda)} = (X_1^{(\lambda_1)}, X_2^{(\lambda_2)})^{\mathrm{T}}$, the bivariate skewness for the distribution, β^*_{12} is given by

$$\begin{aligned}\beta^*_{12} &= \mathrm{E}\left[(X^{(\lambda)}-\mu)^{\mathrm{T}}\Sigma^{-1}(X^{(\lambda)}-\mu)\right]^3 \\ &= (1-\rho^2)^{-3}\left[\mu'^2_{30} + \mu'^2_{03} + 3(1+2\rho^2)(\mu'^2_{12} + \mu'^2_{21}) - 2\rho^3\mu'_{30}\mu'_{03}\right. \\ &\quad \left. +6\rho\left\{\mu'_{30}(\rho\mu'_{12} - \mu'_{21}) - (2+\rho^3)\mu'_{12}\mu'_{21}\right\}\right]\end{aligned} \tag{2.17}$$

where

$$\mu'_{30} = C_0(3,0)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(3,\theta_1:\rho)S_\nu(0,\theta_2:\rho),$$

$$\mu'_{03} = C_0(0,3)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(0,\theta_1:\rho)S_\nu(3,\theta_2:\rho),$$

$$\mu'_{12} = C_0(1,2)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(1,\theta_1:\rho)S_\nu(2,\theta_2:\rho), \text{and}$$

$$\mu'_{21} = C_0(2,1)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(2,\theta_1:\rho)S_\nu(1,\theta_2:\rho).$$

Similarly, the bivariate kurtosis β^*_{22} is also given by

$$\begin{aligned}\beta^*_{22} &= \mathrm{E}\left[(X^{(\lambda)}-\mu)^{\mathrm{T}}\Sigma^{-1}(X^{(\lambda)}-\mu)\right]^2 \\ &= \frac{\mu'_{40}+\mu'_{04}+2\mu'_{22}+4\rho(\rho\mu'_{22}-\mu'_{13}-\mu'_{31})}{(1-\rho^2)^2} \end{aligned} \quad (2.18)$$

where

$$\mu'_{40} = C_0(4,0)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(4,\theta_1:\rho)S_\nu(0,\theta_2:\rho),$$

$$\mu'_{04} = C_0(0,4)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(0,\theta_1:\rho)S_\nu(4,\theta_2:\rho),$$

$$\mu'_{13} = C_0(1,3)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(1,\theta_1:\rho)S_\nu(3,\theta_2:\rho),$$

$$\mu'_{31} = C_0(3,1)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(3,\theta_1:\rho)S_\nu(1,\theta_2:\rho), \text{and}$$

$$\mu'_{22} = C_0(2,2)\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\frac{\rho}{1-\rho^2}\right)^{\nu} S_\nu(2,\theta_1:\rho)S_\nu(2,\theta_2:\rho).$$

Figures 3(a) and 3(b) show the relationships between β^*_{12} and ρ, β^*_{22} and ρ for the various shapes of $\lambda = \lambda_1 = \lambda_2$ when $k_1 = k_2 = 1$ and $\tau_1 = \tau_2 = 2, 4, 16$, respectively. The figures suggest that both β^*_{12} and β^*_{22} are larger as ρ increases toward one.

3. Parameter Estimation

In this section, we discuss the computational algorithm for estimating parameters from the bivariate power-normal distribution.

Let $X_1 = (X_{11}, X_{21})^{\mathrm{T}}, \cdots, X_n = (X_{1n}, X_{2n})^{\mathrm{T}}$ be the vector of observations has the bivariate power-normal distribution. The likelihood function for the sample size n is given by

$$L(x_1, x_2) = \prod_{i=1}^{n}\frac{x_1^{\lambda_1-1}x_2^{\lambda_2-1}}{A(\lambda,\mu,\sum)}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)})}{2}\right\}$$

and then the log-likelihood function becomes

$$
\begin{aligned}
l(x_1, x_2) &= \log L(x_1, x_2) \\
&= -n\log(2\pi) - \tfrac{n}{2}\left\{\log\sigma_1^2 + \log\sigma_1^2 + \log(1-\rho^2)\right\} - \frac{1}{2}\sum_{i=1}^{n} Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) \\
&\quad +(\lambda_1 - 1)\sum_{i=1}^{n}\log x_{1i} + (\lambda_1 - 1)\sum_{i=1}^{n}\log x_{2i} - n\log A(\lambda, \mu, \Sigma).
\end{aligned}
\tag{3.1}
$$

The maximum likelihood estimates of λ_1, λ_2, μ_1, μ_2, σ_1^2, σ_2^2 and ρ can be obtained by maximizing the log-likelihood function.

There are the two approaches to dealing with (λ_1, λ_2) when estimating μ_1, μ_2, σ_1^2, σ_2^2 and ρ (Bickel and Doksum, 1981; Hinkley and Runger, 1984). The first is that the estimation of (λ_1, λ_2) is performed separated from those of μ_1, μ_2, σ_1^2, σ_2^2 and ρ. Namely, (λ_1, λ_2) is chosen $(\hat{\lambda}_1, \hat{\lambda}_2)$ depending on the bivariate power-transformed scale and then for fixed $(\lambda_1, \lambda_2) = (\hat{\lambda}_1, \hat{\lambda}_2)$, μ_1, μ_2, σ_1^2, σ_2^2 and ρ are estimated. The second is that the estimation of (λ_1, λ_2) is performed simultaneously with those of μ_1, μ_2, σ_1^2, σ_2^2 and ρ. Namely, (λ_1, λ_2) is not chosen $(\hat{\lambda}_1, \hat{\lambda}_2)$ depending on the bivariate power-transformed scale and then (λ_1, λ_2) is not estimated as a nuisance parameter, but together with μ_1, μ_2, σ_1^2, σ_2^2 and ρ. In this paper, we use the first approach.

As pointed out in Goto *et al.* (1984), it seems to be difficult that the estimation allowing the truncation is performed in practical use. Though the estimates when allowing the truncation may provide more precise values than the estimates when ignoring the truncation if $\lambda_j > 0$ and small k_j, the influence of ignoring the truncation on the estimates would be smaller as sample size n is increased (Hamasaki and Goto, 2002). Then, we follow the procedure in Box and Cox (1964), that $X_1^{(\lambda)} = (X_{11}^{(\lambda_1)}, X_{21}^{(\lambda_2)}), \cdots, X_n^{(\lambda)} = (X_{1n}^{(\lambda_1)}, X_{2n}^{(\lambda_2)})$ has the bivariate normal distribution without the truncation, assuming that truncation can be ignored. If $A(\lambda, \mu, \Sigma) \approx 1$, for fixed (λ_1, λ_2), the maximum likelihood estimates of μ_1, μ_2, σ_1^2, σ_2^2 and ρ are given by

$$
\hat{\mu}_1(\lambda_1) = \sum_{i=1}^{n} \frac{x_{1i}^{(\lambda_1)}}{n}, \quad \hat{\mu}_2(\lambda_2) = \sum_{i=1}^{n} \frac{x_{2i}^{(\lambda_2)}}{n},
$$

$$
\hat{\sigma}_1^2(\lambda_1) = \sum_{i=1}^{n} \frac{\left\{x_{1i}^{(\lambda_1)} - \hat{\mu}_1(\lambda_1)\right\}^2}{n}, \quad \hat{\sigma}_2^2(\lambda_2) = \sum_{i=1}^{n} \frac{\left\{x_{2i}^{(\lambda_2)} - \hat{\mu}_2(\lambda_2)\right\}^2}{n}, \quad \text{and}
$$

$$
\hat{\rho}(\lambda_1, \lambda_2) = \sum_{i=1}^{n} \left(\frac{x_{1i}^{(\lambda_1)} - \hat{\mu}_1(\lambda_1)}{\sigma_1}\right)\left(\frac{x_{2i}^{(\lambda_2)} - \hat{\mu}_2(\lambda_2)}{\sigma_2}\right)
$$

respectively. Substitution of the maximum likelihood estimates $\hat{\mu}_1(\lambda_1)$, $\hat{\mu}_2(\lambda_2)$, $\hat{\sigma}_1^2(\lambda_1)$, $\hat{\sigma}_2^2(\lambda_2)$, and $\hat{\rho}(\lambda_1, \lambda_2)$ into the log-likelihood function (2.19) yields

$$
\begin{aligned}
l_{\max}(x_1, x_2) = -n\log(2\pi) - \frac{n}{2}\left\{\log\hat{\sigma}_1^2(\lambda_1) + \log\hat{\sigma}_2^2(\lambda_2) + \log(1 - \hat{\rho}^2(\lambda_1, \lambda_2))\right\} \\
+(\lambda_1 - 1)\sum_{i=1}^{n}\log x_{1i} + (\lambda_1 - 1)\sum_{i=1}^{n}\log x_{2i}
\end{aligned}
$$

apart from constant. The maximum likelihood estimates $(\hat{\lambda}_1, \hat{\lambda}_2)$ of (λ_1, λ_2) are the values of the transformation parameters (λ_1, λ_2) in which the maximized log-likelihood function is a maximum. Furthermore, substitution of $(\hat{\lambda}_1, \hat{\lambda}_2)$ into $\hat{\mu}_1(\lambda_1)$, $\hat{\mu}_2(\lambda_2)$, $\hat{\sigma}_1^2(\lambda_1)$, $\hat{\sigma}_2^2(\lambda_2)$, and $\hat{\rho}(\lambda_1, \lambda_2)$ yields the maximum likelihood estimates $\hat{\mu}_1(\hat{\lambda}_1)$, $\hat{\mu}_2(\hat{\lambda}_2)$, $\hat{\sigma}_1^2(\hat{\lambda}_1)$, $\hat{\sigma}_2^2(\hat{\lambda}_2)$ and $\hat{\rho}(\hat{\lambda}_1, \hat{\lambda}_2)$. In practice, they can be solved by using Newton-Raphson's method. See Hamasaki and Goto (2002) for the detailed discussions.

4. Some Examples with Real Data

In this section, some numerical illustrations are provided to demonstrate the materials and the applications of the bivariate power-normal distribution. We shall make use of a data set collected by the Australian Institute of Sport and reported by Cook and Weisberg (1999), containing several variables measured on 202 Australian athletes. Azzalini and Valle (1996) and Azzalini and Capitanio (1999) have considered the data to illustrate the application of their proposed distribution, "the skew-normal distribution" which includes a normal as a special case. We shall consider the same pairs of variables as considered in Azzalini and Valle (1996) and Azzalini and Capitanio (1999), that is (Height, Weight) and (LBM, BMI), where the meaning of names is: LBM, lean body mass; BMI, body mass index = Weight/(Height)^2.

Table 4. Parameter estimates of (Height, Weight) from the fitted bivariate power-normal distribution

Distributions	Estimates	Height	Weight
Bivariate Normal	μ	180.104	75.008
	σ^2	9.710	13.891
	ρ	0.781	
	β_{12}	1.688	
	β_{22}	10.810	
Bivariate Power-Normal	$l_{\max}$	-887.068	
	λ	0.200	0.000
	μ	9.125	4.300
	σ^2	0.154	0.190
	ρ	0.808	
	ρ^*	0.795	
	β_{12}	2.089	
	β_{22}	11.958	
Bivariate Power-Transformed	β_{12}^*	0.899	
	β_{22}^*	9.106	

Tables 4 and 5 show the parameter estimates of (Height, Weight) and (LBM, BMI) from the fitted bivariate power-normal distribution, respectively. For (Height, Weight), the optimized values of shape parameter (0.200, 0.00) suggest that both are close to zero and (Height, Weight) has a bivariate lognormal distribution. While for (LBM, BMI), the optimized values of shape parameter (0.001, -1.200) suggest that LBM has a lognormal and BMI has a more log tailed distribution than a lognormal (L-shape distribution).

For the bivariate normality, in the pair of data (Height, Weight), the values of bivariate skewness β_{12} and kurtosis β_{22} are 2.089 and 11.958, respectively, and both are larger compared with those obtained when bivariate normal distribution is fitted. While, in the pair of data (LBM, BMI), the values of β_{12} and β_{12} are 2.235 and 10.221, respectively, and both are larger compared with those obtained from the fit of the bivariate normal distribution. In addition, the estimates of correlation for bivariate power-normal distribution are smaller than those of the bivariate power-transformed data. Figures 4(a) and 4(b) display the scatter plot of (Height, Weight) and (LBM, BMI) with contours of the fitted bivariate power-normal distribution, respectively. For both the plots, the observed points and the fitted density exhibit moderate skewness for each of the components and the bivariate power-normal distribution may well-describe the data.

For the power-transformed variables, in the both pair of the data, values of β^*_{12} and β^*_{22} are smaller compared with those obtained form of the bivariate normal and power-normal distributions, especially β^*_{12} are very close to zero. In addition, the estimates of correlation for bivariate power-transformed data are greater than those obtaied from the fit of bivariate normal distribution. Figures 5(a) and 5(b) display the scatter plot of the bivariate power-transformed data (Height, Weight) and (LBM, BMI) with contours of the fitted bivariate normal distribution, respectively. Both plots suggest that the bivariate power-transformed data may be close to the bivariate normal.

Table 5. Parameter estimates of (LBI, BMI) from the fitted bivariate power-normal distribution

Distribution	Estimates	LBI	BMI
Bivariate Normal	μ	64.874	22.956
	σ^2	13.038	2.857
	ρ	0.714	
	β_{12}	1.540	
	β_{22}	9.741	
Bivariate	$l_{\max}$	-639.225	
Power-Norma	λ	0.001	-1.200
	μ	4.155	0.184
	σ^2	0.203	0.003
	ρ	0.737	
	ρ^*	0.721	
	β_{12}	2.235	
	β_{22}	10.221	
Bivariate	β^*_{12}	0.182	
Power-Transformed	β^*_{22}	7.670	

5. Conclusions

A multivariate version of the Box and Cox power-transformation has been discussed by Andrews, Gnanadesikan and Warner (1971) and Gnanadesikan (1977) who have focused on the formal extension, but have not given much attention on the properties of the distributions before/and after a multivariate power-normal transformation.

In this paper, the earlier work on so-called power-normal distribution has been extended to a bivariate case and various issues related to the bivariate power-normal distribution have been discussed. Why we have focused on the bivariate case is that the bivariate power-normal distribution provides the bases of the extension to a multivariate case of the power-normal distribution and its structure can be directly derided form the univariate case. However many other issues related to multivariate power-normal distribution are pending. For a positive random variable $X = (X_1, \cdots, X_p)^{\mathrm{T}}$, the pfd of the multivariate power-normal distribution is given by

$$g(x_1,\cdots,x_p)=(2\pi)^{-p/2}\left|\Sigma^{-1}\right|^{1/2}\frac{\prod_{j=1}^{p}x_j^{\lambda_p-1}}{A(\lambda,\mu,\Sigma)}\exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right\}$$

where $\lambda=(\lambda_1,\cdots,\lambda_p)^{\mathrm{T}}$, $\mu=(\mu_1,\cdots,\mu_p)^{\mathrm{T}}$ and

$$\Sigma=\begin{pmatrix}\sigma_1^2 & \cdots & \rho\sigma_1\sigma_p\\ \vdots & \ddots & \vdots\\ \rho\sigma_1\sigma_p & \cdots & \sigma_p^2\end{pmatrix}.$$

The multivariate power-normal distribution would have the potential applications in multivariate analysis such as discriminant analysis, regression analysis and graphical models and so on. Also, in medical application, it would be helpful to analyze multivariate data such as a laboratory.

Acknowledgement

The authors would like to express their appreciation for Professor Takashi Yanagawa of Kyushu University and two referees who provided several useful suggestions and comments.

Appendix 1

Here we consider the detailed calculation for the joint moment about the origin of order (m_1,m_2) of variable (X_1,X_2). Let the first, the second and third terms in (12) be set as T_0, T_1 and T_2, respectively. Then, T_0, T_1 and T_2 can be expanded as follows:

$$T_0=d_0\left\{\Gamma\left(\frac{p_1+1}{2}\right)\Gamma\left(\frac{p_2+1}{2}\right)V_{11}^{(0)}+\frac{\sqrt{2}a_2}{\sqrt{1-\rho^2}}\left(\frac{p_1+1}{2}\right)\left(\frac{p_2+2}{2}\right)V_{12}^{(0)}\right.$$
$$\left.+\frac{\sqrt{2}a_1}{\sqrt{1-\rho^2}}\left(\frac{p_1+2}{2}\right)\left(\frac{p_2+1}{2}\right)V_{21}^{(0)}+\frac{2a_1a_2}{1-\rho^2}\left(\frac{p_1+2}{2}\right)\left(\frac{p_2+2}{2}\right)V_{22}^{(0)}\right\},$$

$$T_1=d_0\left\{\Gamma\left(\frac{p_1+1}{2}\right)\Gamma\left(\frac{p_2+1}{2}\right)V_{11}^{(1)}+\frac{\sqrt{2}a_2}{\sqrt{1-\rho^2}}\left(\frac{p_1+1}{2}\right)\left(\frac{p_2}{2}\right)V_{12}^{(1)}\right.$$
$$\left.+\frac{\sqrt{2}a_1}{\sqrt{1-\rho^2}}\left(\frac{p_1}{2}\right)\left(\frac{p_2+1}{2}\right)V_{21}^{(1)}+\frac{2a_1a_2}{1-\rho^2}\left(\frac{p_1}{2}\right)\left(\frac{p_2}{2}\right)V_{22}^{(1)}\right\},$$

$$T_2=d_1\left\{\Gamma\left(\frac{p_1}{2}\right)\Gamma\left(\frac{p_2}{2}\right)V_{11}^{(2)}+\frac{\sqrt{2}a_2}{\sqrt{1-\rho^2}}\left(\frac{p_1}{2}\right)\left(\frac{p_2+2}{2}\right)V_{12}^{(2)}\right.$$
$$\left.+\frac{\sqrt{2}a_1}{\sqrt{1-\rho^2}}\left(\frac{p_1+1}{2}\right)\left(\frac{p_2}{2}\right)V_{21}^{(2)}+\frac{2a_1a_2}{1-\rho^2}\left(\frac{p_1+1}{2}\right)\left(\frac{p_2+1}{2}\right)V_{22}^{(2)}\right\}$$

respectively, where

$$d_0 = \frac{1-\rho^2}{2}\left\{2(1-\rho^2)\right\}^{(p_1+p_2)/2}, \quad d_1 = \frac{\rho}{2}\left\{2(1-\rho^2)\right\}^{(p_1+p_2+1)/2},$$

$$V_{11}^{(1)} = \sum_{l=1}^{\infty} J_l^{(1)}(1,1)H_1^{(1)}(p_1,a_1,2l)H_1^{(2)}(p_2,a_2,2l),$$

$$V_{12}^{(1)} = \frac{p_2}{2}\sum_{l=1}^{\infty} J_l^{(1)}(1,1)H_1^{(1)}(p_1,a_1,2l)H_2^{(2)}(p_2,a_2,2l),$$

$$V_{21}^{(1)} = \frac{p_1}{2}\sum_{l=1}^{\infty} J_l^{(1)}(2,1)H_2^{(1)}(p_1,a_1,2l)H_1^{(2)}(p_2,a_2,2l),$$

$$V_{22}^{(1)} = \frac{p_1p_2}{2}\sum_{l=1}^{\infty} J_l^{(1)}(2,2)H_2^{(1)}(p_1,a_1,2l)H_2^{(2)}(p_2,a_2,2l),$$

$$V_{11}^{(2)} = \frac{p_1p_2}{4}\sum_{l=1}^{\infty} J_l^{(2)}(1,1)H_2^{(1)}(p_1,a_1,2l+1)H_1^{(2)}(p_2,a_2,2l+1),$$

$$V_{12}^{(2)} = \frac{p_1(p_2+1)}{2}\sum_{l=1}^{\infty} J_l^{(2)}(2,1)H_1^{(1)}(p_1,a_1,2l+1)H_2^{(2)}(p_2,a_2,2l+1),$$

$$V_{21}^{(2)} = \frac{(p_1+1)p_2}{2}\sum_{l=1}^{\infty} J_l^{(2)}(2,1)H_2^{(1)}(p_1,a_1,2l+1)H_1^{(2)}(p_2,a_2,2l+1),$$

and

$$V_{22}^{(2)} = \frac{(p_1+1)(p_2+1)}{2}\sum_{l=1}^{\infty} J_l^{(2)}(2,2)H_2^{(1)}(p_1,a_1,2l+1)H_2^{(2)}(p_2,a_2,2l+1).$$

Also

$$J_l^{(1)}(1,1) = J_{l-1}^{(1)}(1,1)\frac{(p_1+2l-1)(p_2+2l-1)\rho^2}{2l(2l-1)},$$

$$J_l^{(1)}(1,2) = J_{l-1}^{(1)}(1,2)\frac{(p_1+2l-1)(p_2+2l)\rho^2}{2l(2l-1)},$$

$$J_l^{(1)}(2,1) = J_{l-1}^{(1)}(2,1)\frac{(p_1+2l)(p_2+2l-1)\rho^2}{2l(2l-1)},$$

$$J_l^{(1)}(2,2) = J_{l-1}^{(1)}(2,2)\frac{(p_1+2l)(p_2+2l)\rho^2}{2l(2l-1)},$$

$$J_0^{(2)}(1,1) = J_0^{(2)}(1,2) = J_0^{(2)}(2,1) = J_0^{(2)}(2,2) = 1,$$

$$H_1^{(1)}(p_1,a_1,v) = \sum_{\xi=1}^{\infty} M_\xi(1,1), \quad H_1^{(2)}(p_2,a_2,v) = \sum_{\xi=1}^{\infty} M_\xi(2,1),$$

$$H_2^{(1)}(p_1,a_1,v) = \sum_{\xi=1}^{\infty} M_\xi(1,2), \quad H_2^{(2)}(p_2,a_2,v) = \sum_{\xi=1}^{\infty} M_\xi(2,2),$$

$$M_\xi(1,1) = M_{\xi-1}(1,1)\frac{p_1+v+2\xi-1}{2\xi(2\xi-1)}\left(\frac{a_1^2}{1-\rho^2}\right),$$

$$M_\xi(1,2) = M_{\xi-1}(1,2)\frac{p_1+v+2\xi}{2\xi(2\xi-1)}\left(\frac{a_1^2}{1-\rho^2}\right),$$

$$M_\xi(2,1) = M_{\xi-1}(2,1)\frac{p_2+v+2\xi-1}{2\xi(2\xi-1)}\left(\frac{a_2^2}{1-\rho^2}\right),$$

$$M_\xi(2,2) = M_{\xi-1}(2,2)\frac{p_2+v+2\xi}{2\xi(2\xi-1)}\left(\frac{a_2^2}{1-\rho^2}\right),$$

and

$$M_0(1,1) = M_0(1,2) = M_0(2,1) = M_0(2,2) = 1.$$

Appendix 2

$S_\nu(m_1, \theta_1 : \rho)$ in (18) can be written by

$$S_\nu(m_1,\theta_1:\rho) = \sum_{\xi=0}^{\infty}\frac{1}{\xi!}\left(\frac{\theta_1}{1-\rho}\right)^\xi \int_{c_1/\sigma_1}^{\infty} \exp\left\{-\frac{v_1^2}{2(1-\rho^2)}\right\}dv_1.$$

The integral for the right hand can be represented in terms of the complete gamma function and the incomplete gamma function as follows; for $c_1 \geq 0$, it become

$$\sqrt{\frac{1-\rho^2}{2}}\left\{2(1-\rho^2)\right\}^{(m_1+\nu+\xi)/2}\Gamma\left(\frac{m_1+\nu+\xi+1}{2}, \frac{c_1^2}{2(1-\rho^2)\sigma_1^2}\right)$$

and for $c_1 < 0$

$$\sqrt{\frac{1-\rho^2}{2}}\left\{2(1-\rho^2)\right\}^{(m_1+\nu+\xi)/2}$$
$$\times\left\{\Delta\Gamma\left(\frac{m_1+\nu+\xi+1}{2}\right) - \Gamma\left(\frac{m_1+\nu+\xi+1}{2}, \frac{c_1^2}{2(1-\rho^2)\sigma_1^2}\right)\right\}$$

where $\Delta = 1 + (-1)^{m_1+\nu+\xi}$. $S_\nu(m_2, \theta_2 : \rho)$ can be calculated in the same way described in that of $S_\nu(m_1, \theta_1 : \rho)$.

References

Andrews, D.F., Gnanadesikan, R. and Warner, J.L. (1971). Transformation of multivariate data. *Biometrics* **27**, 825-840.

Azzalini, A. and Valle, D.A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715-726.

Azzalini, A and Capitanio, A. (1999). Statistical application of the multivariate skew normal distribution. *Journal of the Royal Statistical Society* **B61**, 579-602.

Box, G.P.E. and Cox, D.R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society* **B26**, 211-252.

Bickel, P.J. and Doksum, K.A. (1981). An analysis of transformations: revisited. *Journal of the American Statistical Association* **76**, 296-311.

Cook, R.D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics.* John Wiley & Sons.

Hinkley D.V. and Runger, G. (1984). The analysis of transformed data. *Journal of the American Statistical Association* **79**, 302-320.

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations.* John Wiley & Sons.

Goto, M. and Inoue, T. (1980). Some properties of power-normal distribution. *Japanese Journal of Biometrics* **1**, 28-54.

Goto, M., Inoue, T. and Tsuchiya, Y. (1984). On estimation of parameters in power-normal distribution. *Bulletin of Informatics and Cybernetics* **21**(1-2), 41-53.

Goto, M., Matsubara, Y. and Tsuchiya, Y. (1983). Power-normal distribution and its applications. *Reports of Statistical Application Research* **30**(3), 8-28.

Goto, M., Uesaka, H. and Inoue, T. (1979). Some linear models for power transformed data. *Invited Presentation at the Tenth International Biometric Conference at Brazil,* August, 6-10.

Goto, M., Yamamoto, Y. and Inoue. T. (1991). Parameter estimation for power-normal distribution: Asymptotic behavior of the estimators (in Japanese). *Bulletin of Computational Statistics of Japan* **4**, 45-60.

Hamasaki, T. and Goto, M. (2002). On parameter estimation of the bivariate power-normal distribution (in Japanese). *Submitted.*

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions.* John Wiley & Sons.

Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57**, 519-530.

Uesaka, H. and Goto, M. (1980). An analysis of clinical chemistry data based on power-normal distribution (in Japanese). *Japanese Journal of Applied Statistics* **9**(1), 23-33.

Uesaka, H. and Goto, M. (1982). Some properties of Box-Cox estimates of transforming parameters in power-normal distributions. *Japanese Journal of Biometrics* **3**, 1-22.

Received April 5, 2002
Revised August 13, 2002

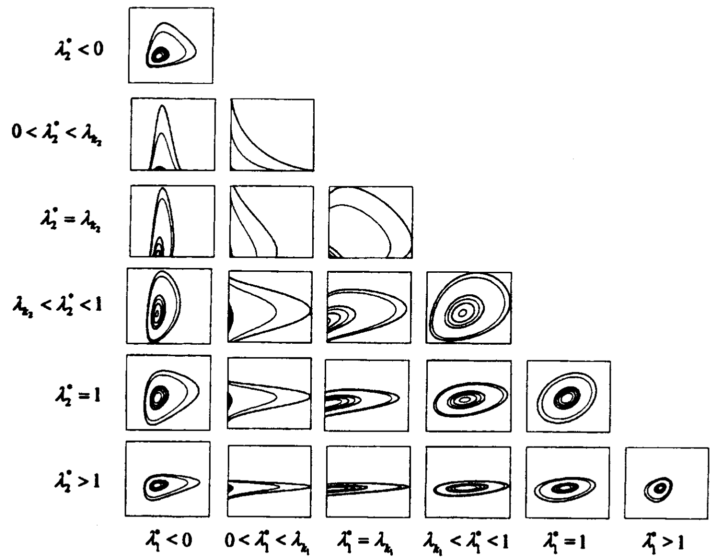

Figure 1(a). Contour plot of the various bivariate power-normal distribution with combinations of λ_1 and λ_2 for $k_1 = k_2 = 1$, $\tau_1 = \tau_2 = 2$ and $\rho = 0.3$

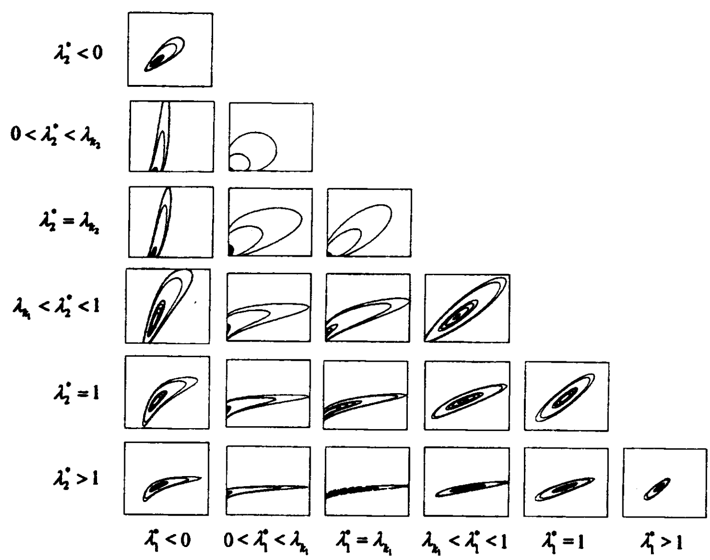

Figure 1(b). Contour plot of the various bivariate power-normal distribution with combinations of λ_1 and λ_2 for $k_1 = k_2 = 1$, $\tau_1 = \tau_2 = 2$ and $\rho = 0.9$

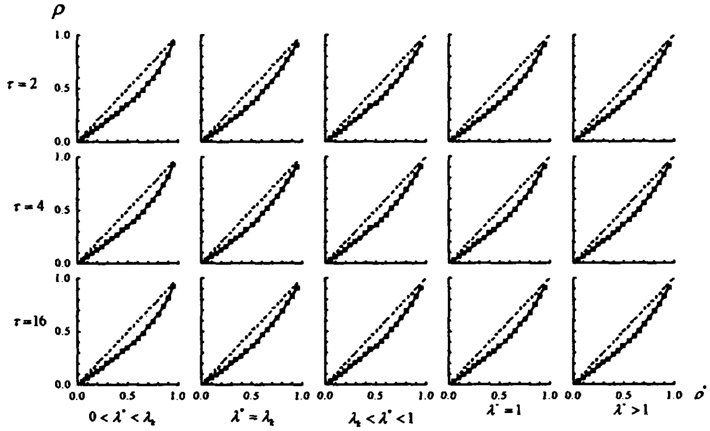

Figure 2. Scatter plot of ρ^* and ρ for the various shapes of $\lambda=\lambda_1=\lambda_2$, when $k_1=k_2=1$ and $\tau(=\tau_1=\tau_2)=2,4,16$

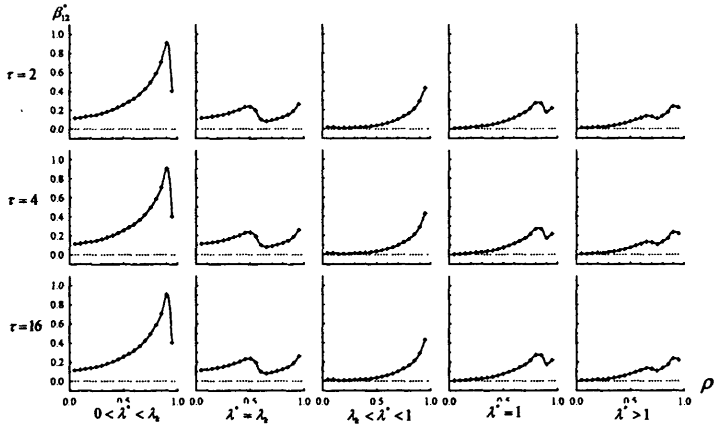

Figure 3(a). Scatter plot of β_{12}^* and ρ for the various shapes of $\lambda=\lambda_1=\lambda_2$ when $k_1=k_2=1$ and $\tau_1=\tau_2=2,4,16$

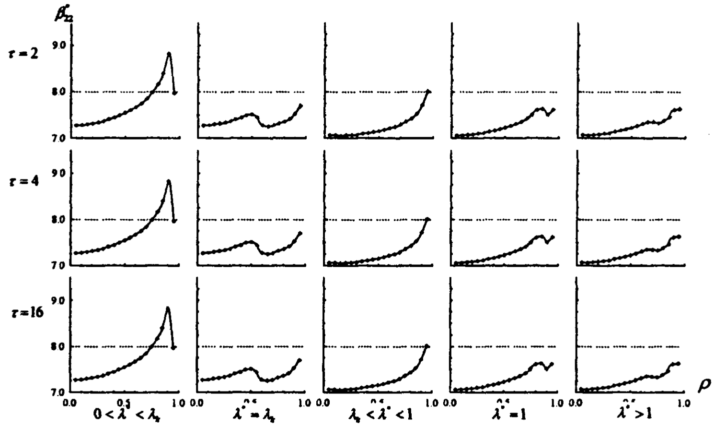

Figure 3(b). Scatter plot of β_{22}^{*} and ρ for the various shape of $\lambda = \lambda_1 = \lambda_2$ when $k_1 = k_2 = 1$ and $\tau_1 = \tau_2 = 2,4,16$

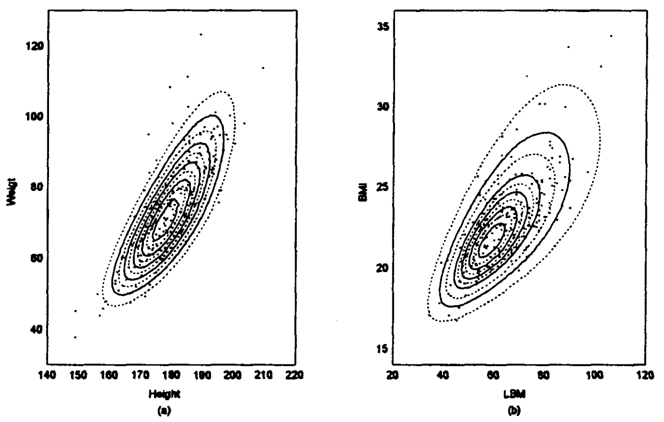

Figure 4. Scatter plots of (a) (Height, Weight) and (b) (LBM, BMI), and levels of the fitted bivariate power-normal distribution

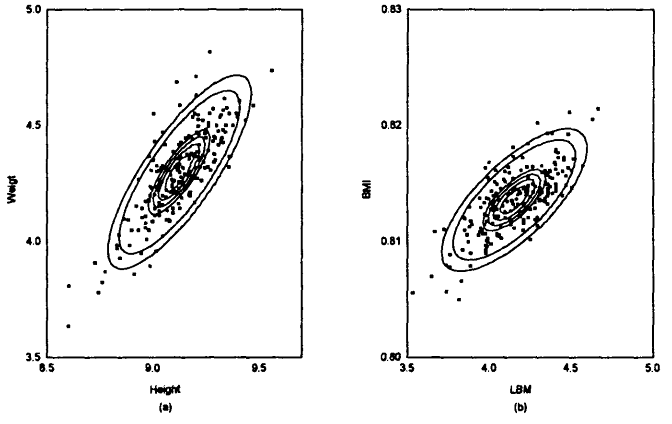

Figure 5. Scatter plots of the power-transformed (a) (Height, Weight) and (b) (LBM, BMI), and levels of the fitted bivariate normal distribution