Research Report

KSTS/RR-04/011
Oct. 22, 2004

Dependent models on two tori, cylinders and unit discs

by

Shogo Kato
Kunio Shimizu

Shogo Kato
Keio University
Kunio Shimizu
Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522 Japan

KSTS/RR-04/011
October 22, 2004

Dependent models on two tori, cylinders and unit discs

Shogo Kato[a,*] , Kunio Shimizu[b,†]

[a] *School of Fundamental Science and Technology, Keio University*
[b] *Department of Mathematics, Keio University*

3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, JAPAN

Abstract

A note on some maximum entropy distributions on the cylinder is provided. Then the paper proposes new distributions on subsets of $\mathbb{R}^4$ such as two tori, cylinders and unit discs with the specified marginal distributions. A related Markov process on subset of $\mathbb{R}^2$ and models for incomplete observations are constructed.

Key words: directional statistics, Markov process, maximum entropy, von Mises distribution.

1 Introduction

In directional statistics various methods have been proposed in the literature to obtain distributions on the manifolds such as real and complex spheres, cylinder, unit disc, and torus. A class of distributions on the cylinder with specified marginal distributions were proposed by Johnson and Wehrly (1978). Wehrly and Johnson (1980) also proposed distributions on the torus in a similar manner and studied a related Markov process on the circle and statistical inference for its process. Shimizu and Iida (2002) proposed a Pearson type VII distribution on arbitrary dimensional spheres by using the scale mixture of the multivariate normal distributions. Jones (2002) proposed distributions on the unit disc with a single specified marginal distribution. Kato and Shimizu (2004) obtained some symmetric or asymmetric t-distributions on the circle, complex sphere and cylinder by the method of the scale mixture or the conditional distribution of the multivariate normal distribution. For other methods to obtain circular distributions, see Mardia and Jupp (1999) or Jammalamadaka and SenGupta (2001).

*Corresponding author. Tel: +81-45-566-1641
Email addresses: shogo@stat.math.keio.ac.jp (Shogo Kato)
shimizu@math.keio.ac.jp (Kunio Shimizu)

†This work was supported in part by the Ministry of Education, Culture, Sport, Science and Technology in Japan under a Grant in Aid of the 21st Century Center of Excellence for Integrative Mathematical Sciences: Progress in Mathematics Motivated by Social and Natural Sciences.

In this paper we propose new distributions with specified marginal distributions on subset of $\mathbb{R}^4$ such as two tori, cylinders and unit discs. These distributions provide models for observations which are represented as points on two bivariate manifolds. For example, in meteorology, two pairs of wind directions measured at two locations at two points in time are observations on two tori. Another example is given for two pairs of wind direction and speed, i.e., observations on two cylinders. Or observations may lie on two unit discs. We also study a related Markov process on the subset of $\mathbb{R}^2$. Then models for incomplete observations are constructed. They can be applied to two observations such as one on the circle and the other on the torus.

Subsequent sections are organized as follows. In Section 2 we give a note on Johnson and Wehrly's paper (1978). We specify the normalizing constant of the distribution they proposed in the paper. Then new maximum entropy distributions are proposed, which can be viewed as an extension of their distributions. Some properties are given for the proposed distributions. In Section 3 a theorem (Theorem 4) is shown, which constructs families of distributions on subset of $\mathbb{R}^4$ with the specified marginal distributions. Then we give examples of the distributions obtained by applying Theorem 4. In Section 4 a related Markov process is constructed. It is obtained by using Theorem 4, too. In Section 5 another theorem is shown, which provides models for two observations such as one on the circle and the other on the cylinder. In Section 6 we discuss distributions on the unit disc which are related to Theorem 4. It is shown that Theorem 4 can be applied to construct distributions on the unit disc. They contain the bivariate spherically symmetric beta distributions on the unit disc.

2 Note on Johnson and Wehrly's paper

Johnson and Wehrly (1978) proposed some distributions on the cylinder which maximize the entropy subject to constraints of certain moments. In Theorem 2 they introduced a distribution on the cylinder with probability density function (p.d.f.)

$$f(\theta, x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\kappa^2/(4\sigma^2)} C_1 \exp\left\{ -\frac{(x-\lambda)^2}{2\sigma^2} + \frac{\kappa x}{\sigma^2}\cos(\theta-\mu) \right\}, \tag{1}$$

where $0 \le \theta < 2\pi,\ -\infty < x < \infty,\ -\infty < \lambda < \infty,\ \kappa > 0$, and $0 \le \mu < 2\pi$. In the paper they did not specify the constant C_1. However, it can be specified by using the Fourier expansion.

Theorem 1 *Let (Θ, X) have the joint p.d.f. (1). Then the constant C_1 can be expressed as*

$$C_1^{-1} = 2\pi \left\{ I_0\left(\frac{\kappa\lambda}{\sigma^2}\right) I_0\left(\frac{\kappa^2}{4\sigma^2}\right) + 2\sum_{j=1}^{\infty} I_j\left(\frac{\kappa^2}{4\sigma^2}\right) I_{2j}\left(\frac{\kappa\lambda}{\sigma^2}\right) \right\}, \tag{2}$$

where I_j denotes the modified Bessel function of the first kind and order j which is given by

$$I_j(z) = \frac{1}{2\pi}\int_0^{2\pi} \cos(j\theta)\exp\{z\cos\theta\}\,d\theta = \sum_{r=0}^{\infty} \frac{1}{\Gamma(r+j+1)\ r!}\left(\frac{z}{2}\right)^{2r+j}, \quad z \in \mathbb{C}.$$

Proof. By using the Fourier expansion: $\exp\{\kappa\cos\theta\} = I_0(\kappa) + 2\sum_{j=1}^{\infty} I_j(\kappa)\cos(j\theta)$ (see Abramowitz and Stegun, 1972, p. 376), the constant C_1 can be specified as

$$C_1^{-1} = \int_0^{2\pi}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\kappa^2/(4\sigma^2)} \exp\left\{ -\frac{(x-\lambda)^2}{2\sigma^2} + \frac{\kappa x}{\sigma^2}\cos(\theta-\mu) \right\} dx d\theta$$

$$
\begin{aligned}
&= \int_0^{2\pi} \exp\left\{ \frac{\kappa\lambda}{\sigma^2}\cos(\theta-\mu) + \frac{\kappa^2}{4\sigma^2}\cos\left(2(\theta-\mu)\right) \right\} d\theta \\
&= 2\pi \left\{ I_0\left(\frac{\kappa\lambda}{\sigma^2}\right) I_0\left(\frac{\kappa^2}{4\sigma^2}\right) + 2\sum_{j=1}^{\infty} I_j\left(\frac{\kappa^2}{4\sigma^2}\right) I_{2j}\left(\frac{\kappa\lambda}{\sigma^2}\right) \right\}.
\end{aligned}
$$

□

The marginal p.d.f. of Θ can be expressed as

$$
f(\theta) = C_1 \exp\left\{ \frac{\kappa\lambda}{\sigma^2}\cos(\theta-\mu) + \frac{\kappa^2}{4\sigma^2}\cos\left(2(\theta-\mu)\right) \right\}, \quad 0 \le \theta < 2\pi.
$$

This marginal p.d.f. has the same form as the p.d.f. (1.2.5) in the book by Watson (1983). Or it can be viewed as a special case of Rukhin's generalized von Mises or a skew von Mises distribution below (Rukhin, 1972) .

The marginal p.d.f. of X is given by

$$
f(x) = C_2 \exp\left\{ -\frac{(x-\lambda)^2}{2\sigma^2} \right\} I_0\left(\frac{\kappa x}{\sigma^2}\right), \quad -\infty < x < \infty, \tag{3}
$$

where C_2 can be expressed as

$$
C_2 = \frac{\sqrt{2\pi}}{\sigma} \mathrm{e}^{-\kappa^2/(4\sigma^2)} C_1,
$$

by using C_1 given in (2).

The distribution (3) provides a skew normal distribution different from that by Azzalini (1985). When $\kappa \to 0$, it reduces to the normal distribution with mean λ and variance σ^2.

It is remarked that we can obtain the distribution on the circle by wrapping the distribution (3) in the same way as Pewsey (2000) obtained the wrapped skew-normal distribution by wrapping a skew normal distribution by Azzalini (1985). Its p.d.f. is given by

$$
f(\theta) = C_2 \sum_{n=-\infty}^{\infty} \exp\left\{ -\frac{(\theta+2\pi n-\lambda)^2}{2\sigma^2} \right\} I_0\left(\frac{\kappa(\theta+2\pi n)}{\sigma^2}\right), \quad 0 \le \theta < 2\pi.
$$

The trigonometric moment for this p.d.f. is obtainable because it is equal to the characteristic function of X. (See Jammalamadaka and SenGupta, 2001, Proposition 2.1.) The characteristic function of X is given by

$$
\begin{aligned}
E[\mathrm{e}^{itX}] &= 2\pi C_1 \exp\left[i\lambda t - \frac{\sigma^2 t^2}{2} \right] \left\{ I_0\left(\frac{\kappa\lambda}{\sigma^2} + i\kappa t\right) I_0\left(\frac{\kappa^2}{4\sigma^2}\right) \right. \\
&\quad \left. + 2\sum_{j=1}^{\infty} I_j\left(\frac{\kappa^2}{4\sigma^2}\right) I_{2j}\left(\frac{\kappa\lambda}{\sigma^2} + i\kappa t\right) \right\}, \quad i^2 = -1.
\end{aligned}
$$

When $\lambda = 0$, the p.d.f. (3) can be expressed as

$$
f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\kappa^2/(4\sigma^2)} I_0^{-1}\left(\frac{\kappa^2}{4\sigma^2}\right) \exp\left\{ -\frac{x^2}{2\sigma^2} \right\} I_0\left(\frac{\kappa x}{\sigma^2}\right), \quad -\infty < x < \infty. \tag{4}
$$

When $\kappa^2 > 2\sigma^2$, the distribution (4) becomes bimodal, taking maximum value at $x = \pm(\sigma^2/\kappa)B^{-1}(1 - 2\sigma^2/\kappa^2)$, where $B(x) = I_2(x)/I_0(x),\ x \ge 0$, and minimal value at $x = 0$.

When $\kappa^2 \le 2\sigma^2$, the distribution (4) becomes unimodal, taking maximum value at $x = 0$. See Appendix A.

As Theorem 3 Johnson and Wehrly (1978) proposed a maximum entropy p.d.f. subject to $E[X]$, $E[\cos\Theta]$, $E[\sin\Theta]$, $E[X\cos\Theta]$, and $E[X\sin\Theta]$ taking specified values consistent with expectation. Now we propose another maximum entropy distribution on the cylinder which is viewed as an extension of the distribution they proposed in Theorem 3. It is defined as follows. The proof is similar to that of Theorem 3 by Johnson and Wehrly (1978) and therefore omitted.

Theorem 2 *Let (Θ, X) have the joint p.d.f.*

$$\begin{aligned} f(\theta, x) &= C_3 \exp\{-\lambda x + \kappa x \cos(\theta - \mu_1) + \nu_1 \cos(\theta - \mu_2) \\ &\quad + \nu_2 \cos(2(\theta - \mu_3))\}, \qquad 0 \le \theta < 2\pi,\ x > 0, \end{aligned} \tag{5}$$

where $\lambda > \kappa > 0$, $\nu_1, \nu_2 > 0$, $0 \le \mu_1, \mu_2, \mu_3 < 2\pi$. The normalizing constant C_3 can be expressed as

$$\begin{aligned} C_3^{-1} &= \frac{2\pi}{\sqrt{\lambda^2 - \kappa^2}} \Bigg[I_0(\nu_1) I_0(\nu_2) + 2 \sum_{n=1}^{\infty} I_0(\nu_2) I_n(\nu_1) \rho^n \cos(n(\mu_1 - \mu_2)) \\ &\quad + 2 \sum_{n=1}^{\infty} I_0(\nu_1) I_{2n}(\nu_2) \rho^{2n} \cos(2n(\mu_1 - \mu_3)) \\ &\quad + 2 \sum_{m,n=1}^{\infty} I_m(\nu_1) I_{2n}(\nu_2) \Big\{ \rho^{m+2n} \cos(m(\mu_1 - \mu_2) + 2n(\mu_1 - \mu_3)) \\ &\quad + \rho^{|m-2n|} \cos(m(\mu_1 - \mu_2) - 2n(\mu_1 - \mu_3)) \Big\} \Bigg], \end{aligned}$$

where $\rho = (\lambda - \sqrt{\lambda^2 - \kappa^2})/\kappa$. Then the p.d.f. (5) is a maximum entropy p.d.f. subject to $E[X]$, $E[\cos\Theta]$, $E[\sin\Theta]$, $E[\cos 2\Theta]$, $E[\sin 2\Theta]$, $E[X\cos\Theta]$, and $E[X\sin\Theta]$ taking specified values consistent with expectation.

The marginal p.d.f. of Θ is given by

$$f(\theta) = C_3 \frac{\exp\{\nu_1 \cos(\theta - \mu_2) + \nu_2 \cos(2(\theta - \mu_3))\}}{\lambda - \kappa \cos(\theta - \mu_1)}, \quad 0 \le \theta < 2\pi.$$

The marginal p.d.f. of X is given by

$$\begin{aligned} f(x) &= 2\pi C_3\, \mathrm{e}^{-\lambda x} \Big\{ I_0(\nu^*(x)) I_0(\nu_2) \\ &\quad + 2 \sum_{j=1}^{\infty} I_j(\nu_2) I_{2j}(\nu^*(x)) \cos\{2j(\mu^*(x) - \mu_3)\} \Big\}, \quad x > 0, \end{aligned}$$

where $\nu^*(x)$ (≥ 0) and $\mu^*(x)$ $(0 \le \mu^*(x) < 2\pi)$ are the constants which satisfy

$$\nu^*(x) \cos(\mu^*(x)) = \kappa x \cos\mu_1 + \nu_1 \cos\mu_2, \tag{6}$$

$$\nu^*(x) \sin(\mu^*(x)) = \kappa x \sin\mu_1 + \nu_1 \sin\mu_2. \tag{7}$$

The conditional distribution of Θ given $X = x$ is given by the p.d.f.

$$f_1(\theta|x) = C_4 \exp\{\nu^* \cos(\theta - \mu^*) + \nu_2 \cos(2(\theta - \mu_3))\}, \quad 0 \le \theta < 2\pi, \tag{8}$$

where ν^* $(= \nu^*(x))$ and μ^* $(= \mu^*(x))$ are the constants which satisfy (6) and (7). The normalizing constant C_4 can be expressed as

$$C_4^{-1} = 2\pi \left\{ I_0(\nu^*) I_0(\nu_2) + 2\sum_{j=1}^{\infty} I_j(\nu_2) I_{2j}(\nu^*) \cos(2j(\mu^* - \mu_3)) \right\}.$$

The conditional p.d.f. of X given $\Theta = \theta$ is expressed as

$$f_2(x|\theta) = (\lambda - \kappa\cos(\theta - \mu_1)) \exp\{-(\lambda - \kappa\cos(\theta - \mu_1))x\}, \quad x > 0. \tag{9}$$

The conditional distributions (8) and (9) are Rukhin's generalized von Mises $\mathrm{GVM}(\nu^*, \nu_2, \mu^*, \mu)$ and exponential p.d.f.s, respectively.

Johnson and Wehrly (1978) discussed the distribution on the cylinder proposed by Mardia and Sutton (1978). It is obtained as a conditional distribution of a trivariate normal distribution or a maximum entropy distribution subject to constraints of certain moments. The following theorem provides a distribution on the cylinder which includes the distribution by Mardia and Sutton (1978) as a special case. The proof is straightforward and omitted.

Theorem 3 *Let (Θ, X) have the joint p.d.f.*

$$f(\theta, x) = C_5 \exp\left[-\frac{(x - \mu(\theta))^2}{2\sigma^2} + \kappa_1\cos(\theta - \mu_1) + \kappa_2\cos(2(\theta - \mu_2))\right], \tag{10}$$

where $0 \le \theta < 2\pi$, $-\infty < x < \infty$, $\sigma > 0$, $\kappa_1, \kappa_2 > 0$, $0 \le \mu_1, \mu_2 < 2\pi$, $\mu(\theta) = \mu' + \lambda\cos(\theta - \nu)$, $-\infty < \mu' < \infty$, $\lambda > 0$, and $0 \le \nu < 2\pi$. The normalizing constant C_5 is given by

$$C_5^{-1} = (2\pi)^{3/2}\sigma \left\{ I_0(\kappa_1)I_0(\kappa_2) + 2\sum_{j=1}^{\infty} I_j(\kappa_2)I_{2j}(\kappa_1)\cos(2j(\mu_1 - \mu_2)) \right\}.$$

Then $f(\theta, x)$ is the maximum entropy p.d.f. on the cylinder subject to $E[X^2]$, $E[X]$, $E[X\cos\Theta]$, $E[X\sin\Theta]$, $E[\cos\Theta], E[\sin\Theta]$, $E[\cos 2\Theta]$, and $E[\sin 2\Theta]$ taking specified values consistent with expectation.

The distribution (10) is also obtainable as a conditional distribution of a trivariate normal distribution without any constraints of mean vector and covariance matrix. See Appendix B for details.

This distribution has the property that the conditional distribution of X given $\Theta = \theta$ is a normal distribution $N(\mu(\theta), \sigma^2)$ and the marginal distribution of Θ is Rukhin's generalized von Mises distribution $\mathrm{GVM}(\kappa_1, \kappa_2, \mu_1, \mu_2)$.

3 Distributions with specified marginal distributions

The following theorem provides classes of continuous distributions on subsets of $\mathbb{R}^4$ with specified marginal distributions. For instance, distributions on the two tori, cylinders and unit discs are constructed.

Theorem 4 *Let $f_1(x_1, x_2)$ be a specified p.d.f. on the support $M_1 \subset \mathbb{R}^2$ and $f_2(y_1, y_2)$ a specifed p.d.f. on the support $M_2 \subset \mathbb{R}^2$, and $F_1(x_1, x_2)$ and $F_2(y_1, y_2)$ be their distribution functions, respectively. Suppose that $f_{11}(x_1)$ $(f_{21}(y_1))$ is the marginal p.d.f. of X_1 (Y_1) and*

$F_{11}(x_1)$ ($F_{21}(y_1)$) its distribution function. Let $g(\cdot)$ be a p.d.f. on the circle and $h(\cdot,\cdot)$ a p.d.f. on the torus. Then

$$\begin{aligned} f_1(x_1,x_2,y_1,y_2) &= 2\pi f_1(x_1,x_2) f_2(y_1,y_2) \\ &\quad \times g\left[2\pi\left\{\frac{1}{f_{11}(x_1)}\frac{\partial F_1(x_1,x_2)}{\partial x_1} \pm \frac{1}{f_{21}(y_1)}\frac{\partial F_2(y_1,y_2)}{\partial y_1}\right\}\right], \end{aligned} \tag{11}$$

$$\begin{aligned} f_2(x_1,x_2,y_1,y_2) &= 4\pi^2 f_1(x_1,x_2) f_2(y_1,y_2) \\ &\quad \times h\Bigg[2\pi\left\{F_{11}(x_1) \pm F_{21}(y_1)\right\}, \\ &\qquad 2\pi\left\{\frac{1}{f_{11}(x_1)}\frac{\partial F_1(x_1,x_2)}{\partial x_1} \pm \frac{1}{f_{21}(y_1)}\frac{\partial F_2(y_1,y_2)}{\partial y_1}\right\}\Bigg], \end{aligned} \tag{12}$$

where $(x_1,x_2) \in M_1,\ (y_1,y_2) \in M_2$, are p.d.f.s on $M_1 \times M_2$. Both have the marginal distributions of (X_1,X_2) $((Y_1,Y_2))$ with the p.d.f. $f_1(x_1,x_2)$ $(f_2(y_1,y_2))$.

Proof. It is clear that $f_1(x_1,x_2,y_1,y_2) \geq 0$. We show

$$\int_{M_2}\int_{M_1} f_1(x_1,x_2,y_1,y_2)dx_1dx_2dy_1dy_2 = 1.$$

Consider the integral

$$\begin{aligned} &\int_{M_2}\int_{M_1} f_1(x_1,x_2,y_1,y_2)dx_1dx_2dy_1dy_2 \\ &= 2\pi\int_{M_2} f_2(y_1,y_2)\int_{M_1} f_1(x_1,x_2) \\ &\quad \times g\left[2\pi\left\{\frac{1}{f_{11}(x_1)}\frac{\partial F_1(x_1,x_2)}{\partial x_1} \pm \frac{1}{f_{21}(y_1)}\frac{\partial F_2(y_1,y_2)}{\partial y_1}\right\}\right] dx_1dx_2dy_1dy_2. \end{aligned} \tag{13}$$

Making the change of the variable $t = t(x_2) = 2\pi\{\partial F_1(x_1,x_2)/\partial x_1\}/\{f_{11}(x_1)\}$, (13) is calculated as

$$\begin{aligned} &\int_{M_2}\int_{M_1} f_1(x_1,x_2,y_1,y_2)dx_1dx_2dy_1dy_2 \\ &= \int_{M_2} f_2(y_1,y_2)\int_{\mathbb{R}}\int_0^{2\pi} g\left[t \pm \frac{2\pi}{f_{21}(y_1)}\frac{\partial F_2(y_1,y_2)}{\partial y_1}\right] f_{11}(x_1)dtdx_1dy_1dy_2 \\ &= \int_{M_2} f_2(y_1,y_2)\int_{\mathbb{R}} f_{11}(x_1)dx_1dy_1dy_2 \\ &= \int_{M_2} f_2(y_1,y_2)dy_1dy_2 \\ &= 1. \end{aligned}$$

¿From this result, it is obvious that the marginal p.d.f. of (Y_1,Y_2) is $f_2(y_1,y_2)$. Similarly, we can show that the marginal p.d.f. of (X_1,X_2) is $f_1(x_1,x_2)$.

By making the change of variables $t_1 = t_1(x_2) = 2\pi[\{\partial F_1(x_1,x_2)/\partial x_1\}/f_{11}(x_1)]$ and $t_2 = t_2(x_1) = 2\pi\{F_{11}(x_1) \pm F_{21}(y_1)\}$, we can also show that

$$\int_{M_2}\int_{M_1} f_2(x_1,x_2,y_1,y_2)dx_1dx_2dy_1dy_2 = 1.$$

The marginal p.d.f. is also obtained by proceeding that process.

□

Examples

When we put $M_1 = M_2 = [0, 2\pi) \times \mathbb{R}$ in distributions (11) and (12), classes of distributions on the two cylinders are constructed. For example, by taking the distribution (1) or (5) as a marginal distribution, we can get distributions on the two cylinders. When $M_1 = M_2 = [0, 2\pi)^2$, classes of distributions on the two tori are constructed. For the examples of distributions on the torus, see Mardia (1975) or Breckling (1989). Distributions on the two tori or unit discs are made in a similar manner.

It is not necessary that M_1 and M_2 are the same manifolds. Actually, by putting M_1 is a torus and M_2 a cylinder, we can get families of distributions on the direct product of the torus and cylinder. Similarly, distributions on the direct product of the torus and the unit disc are constructed.

4 A related Markov process

In a similar way as in Wehrly and Johnson (1980), families of distributions for Markov process on subsets of $\mathbb{R}^2$ are constructed through Theorem 4. They are constructed as follows.

Let $X_0, X_1, \ldots$ $(X_i = (X_{i1}, X_{i2})',\ i = 0, 1, 2, \ldots)$ be random variables taking values on $M \subset \mathbb{R}^2$ such that

$$\begin{aligned} p(x_0) &= f(x_{01}, x_{02}), \\ p(x_n | x_0, x_1, \ldots, x_{n-1}) &= p(x_n | x_{n-1}) \\ &= 2\pi f(x_{n1}, x_{n2})\ g\left[2\pi\left\{\frac{1}{f_1(x_{n1})}\frac{\partial F(x_{n1}, x_{n2})}{\partial x_{n1}}\right.\right. \\ &\qquad \left.\left.\pm \frac{1}{f_1(x_{n-1,1})}\frac{\partial F(x_{n-1,1}, x_{n-1,2})}{\partial x_{n-1,1}}\right\}\right], \end{aligned}$$

where $f(\cdot, \cdot)$ denotes a p.d.f. on M and $F(\cdot, \cdot)$ its distribution function. Then $X_0, X_1, \ldots$ is a Markov process on M with the initial distribution $p(x_0)$ and the stationary transition p.d.f. $p(x_n | x_{n-1})$.

Similarly, we can construct another Markov process by using the distribution (12) as the stationary transition p.d.f.

5 Models on other manifolds

The following theorem provides models for two observations such as one on the circle and the other on the torus, for example, a model for incomplete pairs of wind directions observed at two locations at two points in time in meteorology and in environmetal science. The proof is omitted.

Theorem 5 *Let $f_1(x_1, x_2)$ be a specified p.d.f. on the support $M_1 \subset \mathbb{R}^2$ and $f_2(y)$ a specified p.d.f. on the support $M_2 \subset \mathbb{R}$, and $F_1(x_1, x_2)$ and $F_2(y)$ be their distribution functions. Suppose that $f_{11}(x_1)$ is the marginal p.d.f. of X_1. Let $g(\cdot)$ be a p.d.f. on the circle . Then*

$$\begin{aligned} f(x_1, x_2, y) &= 2\pi f_1(x_1, x_2) f_2(y) \\ &\quad \times g\left[2\pi\left\{\frac{1}{f_{11}(x_1)}\frac{\partial F_1(x_1, x_2)}{\partial x_1} \pm F_2(y)\right\}\right], \end{aligned}$$

where $(x_1, x_2) \in M_1,\ y \in M_2$, *is a p.d.f. on* $M_1 \times M_2$. *It has the marginal distribution of* (X_1, X_2) (Y) *with the p.d.f.* $f_1(x_1, x_2)$ $(f_2(y))$.

Note that Theorem 5 can be considered to be a special case of Theorem 4. Actually, by putting $M_2 = [0, 1) \times M,\ M \subset \mathbb{R}$, and $f_2(y_1, y_2) = f_2(y_2),\ y_2 \in M$ in Theorem 4, we can get Theorem 5.

6 Distribution on the unit disc

Jones (2002) proposed a class of distributions on the unit disc with a single specified marginal distribution. It is obtained by specifying a single marginal distribution with p.d.f. $f_X(x)$ and a conditional distribution with p.d.f. $f(y|x)$. Its p.d.f. is given by $f(x, y) = f_X(x)f(y|x)$.

In this section we propose a class of distributions on the unit disc with two specified marginal distributions. The process is very similar to that of Wehrly and Johnson (1980). Its p.d.f. is defined as follows.

Let $f_1(r)$ be a specified p.d.f. on $[0, 1)$ and $f_2(\theta)$ a specified p.d.f. on $[0, 2\pi)$, and $F_1(r)$ and $F_2(\theta)$ be their distribution functions, respectively. Let $g(\cdot)$ be another p.d.f. on the circle. Then

$$f(r, \theta) = 2\pi f_1(r)f_2(\theta)g\left[2\pi\left\{F_1(r) \pm F_2(\theta)\right\}\right], \quad 0 \le r < 1,\ 0 \le \theta < 2\pi, \tag{14}$$

is a p.d.f. on the unit disc with the marginal p.d.f.s $f_1(r)$ and $f_2(\theta)$.

This p.d.f. can be obtained by applying Theorem 4. Actually, by putting $M_1 = [0, 1)^2$, $M_2 = [0, 1) \times [0, 2\pi),\ f_1(x_1, x_2) = f_1(x_2),\ 0 \le x_1, x_2 < 1$ and $f_2(y_1, y_2) = f_2(y_2),\ 0 \le y_1 < 1,\ 0 \le y_2 < 2\pi$, we get distributions with the p.d.f. (14). Distributions on the unit disc in this manner contain the bivariate spherically symmetric beta (or Pearson type II) distributions with p.d.f.

$$f(r, \theta) = \frac{\alpha}{\pi} r(1 - r^2)^{\alpha-1}, \quad 0 \le r < 1,\ 0 \le \theta < 2\pi,$$

where $\alpha > 0$. It is obtained by putting $f_1(r) = 2\alpha r(1-r^2)^{\alpha-1},\ f_2(\theta) = 1/2\pi$, and $g(t) = 1/2\pi$ in (14).

Appendix

A Unimodality

The distribution (4) becomes unimodal or bimodal, depending on the choice of κ and σ.

That is shown by differentiating the p.d.f. (4) with respect to x, and equating to zero,

$$\begin{aligned}\frac{d}{dx}f(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\kappa^2/(4\sigma^2)} I_0^{-1}\left(\frac{\kappa^2}{4\sigma^2}\right) \exp\left(-\frac{x^2}{2\sigma^2}\right) \\ &\quad \times \left\{-\frac{x}{\sigma^2} I_0\left(\frac{\kappa x}{\sigma^2}\right) + \frac{\kappa}{\sigma^2} I_1\left(\frac{\kappa x}{\sigma^2}\right)\right\} = 0. \end{aligned} \tag{15}$$

Using the fact that

$$I_0(x) - I_2(x) = \frac{2}{x} I_1(x),$$

(15) can be calculated as

$$I_2\left(\frac{\kappa x}{\sigma^2}\right) \Big/ I_0\left(\frac{\kappa x}{\sigma^2}\right) = 1 - \frac{2\sigma^2}{\kappa^2}.$$

Let $B(x)$ denote the ratio of the Bessel functions

$$B(x) = \frac{I_2(x)}{I_0(x)}, \quad x \geq 0.$$

Then $B(x)$ has the following properties:

(a) $0 \leq B(x) \leq 1, \quad x \geq 0$

(b) $\lim_{x \to +0} B(x) = 0, \quad \lim_{x \to \infty} B(x) = 1$

(c) $\frac{dB(x)}{dx} > 0$

The proof is as follows.
(a) It is obvious that $B(x) \geq 0$. Using the fact that $A(x)(= I_1(x)/I_0(x)) \geq 0,\ x \geq 0$, we have

$$\begin{aligned} B(x) &= \frac{I_0(x) - (2/x)I_1(x)}{I_0(x)} \\ &= 1 - \frac{2}{x}A(x) \\ &\leq 1, \quad x > 0. \end{aligned}$$

(b) Clearly $\lim_{x \to +0} B(x) = 0$. By using the fact that $\lim_{x \to \infty} A(x) = 1$ (See Jammalamadaka and SenGupta (2001)), $\lim_{x \to \infty} B(x) = 1$ is shown as follows

$$\lim_{x \to \infty} B(x) = \lim_{x \to \infty} \left(1 - \frac{2}{x}A(x) \right) = 1.$$

(c) The p variate von Mises-Fisher distribution on the unit sphere Ω^p in $\mathbb{R}^p$ is defined by p.d.f.

$$f(\boldsymbol{x}) = \left(\frac{\kappa}{2} \right)^{p/2-1} \frac{1}{\Gamma(p/2) I_{p/2-1}(\kappa)} \exp\left(\kappa \boldsymbol{\mu}' \boldsymbol{x} \right), \quad \boldsymbol{x} \in \Omega^p,$$

where $\kappa \geq 0,\ \boldsymbol{\mu} \in \Omega^p$. The Fisher information of maximum likelihood estimator of κ (Mardia and Jupp, 1999, p. 199) is given by

$$-E\left[\frac{\partial^2}{\partial \kappa^2} \log f(\boldsymbol{X}) \right] = A_p'(\kappa),$$

where $A_p(z) = I_{p/2}(z)/I_{p/2-1}(z)$. From the fact that the Fisher information is positive, we have

$$\begin{aligned} \frac{d}{dz}\left\{ \frac{I_2(z)}{I_0(z)} \right\} &= \frac{d}{dz}\left\{ \frac{I_1(z)}{I_0(z)} \cdot \frac{I_2(z)}{I_1(z)} \right\} = \frac{d}{dz}\left\{ A_2(z) A_4(z) \right\} \\ &= A_2'(z) A_4(z) + A_2(z) A_4'(z) \\ &> 0, \quad z > 0. \end{aligned}$$

By using the properties of $B(x)$, it is shown that the distribution (4) becomes unimodal when $\kappa^2 > 2\sigma^2$ and bimodal otherwise.

B Derivation of the distribution (10)

We noted that the distribution (10) can be obtained by conditioning a trivariate normal distribution without any constraints of mean vector and covariance matrix. It is obtained as follows.

Let $\boldsymbol{X}$ be a random vector which obeys the trivariate normal distribution with mean vector $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)'$ and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^3 \end{pmatrix},$$

where $-\infty < \eta_i < \infty$, $\sigma_i > 0$ $(i = 1, 2, 3)$ and $-1 < \rho_{12}, \rho_{13}, \rho_{23} < 1$. We transform the trivariate random vector $\boldsymbol{X} = (X, X_1, X_2)' = (X, R\cos\Theta, R\sin\Theta)'$ where $R > 0$, $0 \le \Theta < 2\pi$. Then the conditional distribution of (Θ, X) given $R = r$ is given by the p.d.f. (10) by defining new parameters as

$$\mu(\theta) = \eta_1 + d\eta_2 + e\eta_3 + r^* \cos(\theta - \eta^*), \quad \sigma^2 = \frac{1}{a},$$

$$\kappa_1 \cos\mu_1 = r\left\{\eta_2(b - ad^2) + \eta_3 f\right\}, \quad \kappa_1 \sin\mu_1 = r\left\{\eta_3(c - ae^2) + \eta_2 f\right\},$$

$$\kappa_2 \cos 2\mu_2 = \frac{r^2}{4}(c - b + ad^2 - ae^2), \quad \kappa_2 \sin 2\mu_2 = -\frac{r^2 f}{2},$$

where

$$r^* \cos\eta^* = rd, \quad r^* \sin\eta^* = re, \quad a = \frac{\rho(1-\rho_{23}^2)}{\sigma_1^2}, \quad b = \frac{\rho(1-\rho_{13}^2)}{\sigma_2^2},$$

$$c = \frac{\rho(1-\rho_{12}^2)}{\sigma_3^2}, \quad d = -\frac{\rho(\rho_{12} - \rho_{13}\rho_{23})}{\sigma_1\sigma_2}, \quad e = -\frac{\rho(\rho_{13} - \rho_{12}\rho_{23})}{\sigma_1\sigma_3},$$

$$f = -\frac{\rho(\rho_{23} - \rho_{12}\rho_{13})}{\sigma_2\sigma_3}, \quad \rho^{-1} = 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23}.$$

References

Abramowitz, M., Stegun, I., 1972, Handbook of Mathematical Functions, Dover, New York.

Azzalini, A., 1985, A class of distributions which includes the normal ones, Scand. J. Statist., 12, 171–178.

Breckling, J., 1989, Analysis of Directional Time Series: Applications to Wind Speed and Direction, Springer Verlag, New York.

Jammalamadaka, S.R., SenGupta, A., 2001, Topics in Circular Statistics, World Scientific Publ., New Jersey.

Johnson, R.A., Wehrly, T.E., 1978, Some angular-linear distributions and related regression models, J. Amer. Statist. Assoc., 73, 602–606.

Jones, M.C., 2002, Marginal replacement in multivariate densities, with application to skewing spherically symmetric distributions, J. of Multivariate Anal., 81, 85–99.

Kato, S., Shimizu, K., 2004, A further study of t-distributions on spheres, manuscript submitted for publication.

Mardia, K.V., 1975, Statistics of directional data (with discussion), J. Roy. Statist. Soc. Ser. B, 37, 349–393.

Mardia, K.V., Jupp, P.E., 1999, Directional Statistics, Wiley, Chichester.

Mardia, K.V., Sutton, N.W., 1978, A model for cylindrical variables with applications, J. Roy. Statist. Soc. Ser. B, 40, 229–233.

Pewsey, A., 2000, The wrapped skew-normal distribution on the circle., Commun. Statist.– Theory Meth., 29, 2459–2472.

Rukhin, A.L., 1972, Some statistical decisions about distributions on a circle for large samples, Sankhyā Ser. A, 34, 243–250.

Shimizu, K., Iida, K., 2002, Pearson type VII distributions on spheres, Commun. Statist.– Theory Meth., 31, 513–526.

Watson, G.S., 1983, Statistics on Spheres, Wiley, New York.

Wehrly, T.E., Johnson, R.A., 1980, Bivariate models for dependence of angular observations and a related Markov process, Biometrika, 67, 255–256.