Extensions of Probability Distributions on Torus, Cylinder and Disc

August 2013

Minzhen Wang

主　論　文　要　旨

報告番号	㊙ 乙 第　　　号	氏 名	王 敏真

主 論 文 題 目：

Extensions of Probability Distributions on Torus, Cylinder and Disc
（トーラス、シリンダー及びディスク上の確率分布の拡張）

（内容の要旨）

角度観測値を含むデータの統計学は方向統計学として知られている。二つの時点において観測される風向のような二変量角度データのモデル化の際には二変量角度分布すなわちトーラス上の分布を考えることになる。本論文では、方向統計学の枠組みの中でトーラス上の分布と共にシリンダー及びディスク上の分布の拡張を提案している。

円周上の分布に関する予備的な結果として、円周上のハート型分布に従う確率変数を Möbius 変換して得られる分布の分布関数、三角モーメント、単峰性と対称性の条件について調べた。トーラス上の分布については、Möbius 変換によって潜在構造が定められるモデル、言い換えれば三変量から二変量の分布を生成する変量減少法の一つから導かれるトーラス上の二変量ハート型分布を提案し、結合確率密度関数、結合三角モーメント及び角度変数間相関係数の具体的な式を与えた。また、二変量ハート型分布を利用して、朝 6 時と正午に継続的に観測された風向データの解析例を示した。

次に、シリンダー上の分布を、von Mises 分布と変換された Kumaraswamy 分布の組み合わせから生成し、周辺分布と条件付き分布を与えた。この提案分布は Johnson と Wehrly による分布の拡張である。また、正弦関数によって歪められた巻き込み Cauchy 分布と Weibull 分布の組み合わせから生成した分布も提案した。その他、周辺分布が指定されているときにシリンダー上の分布を生成する方法を用いることによって、ハート型分布と変換された Kumaraswamy 分布を周辺に持ち、von Mises 分布をリンクとする分布の提案を行っている。

最後に、長さを保つ修正メビウス変換を用いてディスク上の Möbius 分布の拡張としての非対称分布を生成し、周辺分布などの諸性質を調べた。

SUMMARY OF Ph.D. DISSERTATION

School	Student Identification Number	SURNAME, First name
Fundamental Science and Technology	81047518	WANG, Minzhen

Title

Extensions of Probability Distributions on Torus, Cylinder and Disc

Abstract

Statistics for data which include angular observations is known as directional statistics. Bivariate circular data such as wind directions measured at two points in time are modeled by using bivariate circular distributions or distributions on the torus. In this thesis, we propose some extensions of distributions on the torus and also the cylinder and the disc in the framework of directional statistics.

For the univariate circular case as a preliminary result, the distribution function, trigonometric moments, and conditions for unimodality and symmetry are studied when the Möbius transformation is applied to a univariate cardioid random variable. For the bivariate circular case, we propose a bivariate cardioid distribution which is generated from a circular-circular structural model linked with Möbius transformation or a method of trivariate reduction. The joint probability density function, trigonometric moments and circular-circular correlation coefficient are explicitly expressed. An illustration is given for wind direction data at 6 a.m. and noon as an application of the bivariate cardioid distribution.

Next, we propose new distributions on the cylinder. A distribution generated from a combination of von Mises and transformed Kumaraswamy distributions is an extension of the Johnson and Wehrly model. The marginal and conditional distributions of the proposed distribution are given. Another model is a combination of sine-skewed wrapped Cauchy and Weibull distributions. A distribution using the method of generating a cylindrical distribution with specified cardioid and transformed Kumaraswamy marginals and von Mises link is also proposed.

Finally, we generate skew or asymmetric distributions on the disc by using a modified Möbius transformation as extensions of the Möbius distribution. Some properties such as the marginal density functions of the proposed distributions are obtained.

A Thesis for the Degree of Ph.D. in Science

Extensions of Probability Distributions on Torus, Cylinder and Disc

August 2013

Graduate School of Science and Technology
Keio University

Minzhen Wang

Contents

1 Introduction and preliminaries **1**

1.1 Introduction 1

1.2 Preliminary notions for circular data 3

1.2.1 Mean direction, mean resultant length and circular variance 3

1.2.2 Trigonometric moments, circular skewness and kurtosis . . 4

1.2.3 Distribution and probability density functions 5

1.2.4 Characteristic function 6

1.3 Some existing distributions on the circle 8

1.3.1 Uniform distribution 8

1.3.2 Cardioid distribution 9

1.3.3 von Mises distribution 9

1.3.4 Wrapped Cauchy distribution 10

1.3.5 Generalized cardioid distribution of Jones & Pewsey 12

1.4 A new family of distributions on the circle 13

1.4.1 Möbius transformation 13

1.4.2 Möbius transformation of a cardioid random variable . . . 14

2 Distributions on the torus **22**

2.1 Introduction 22

2.2 Some existing distributions on the torus 23

2.3 A family of bivariate cardioid distributions 25

2.3.1 Definition and joint probability density function 25
2.3.2 Trigonometric moments 26
2.3.3 Correlation coefficient 28
2.4 Illustrative example 29

3 Distributions on the cylinder **34**
3.1 Introduction 34
3.2 Existing distributions on the cylinder 35
3.3 New distributions on the cylinder 38
3.4 Illustrative examples 42
3.5 Circular-linear regression models 51

4 Distributions on the disc **53**
4.1 Introduction 53
4.2 Möbius distribution 54
4.3 A modified Möbius distribution 55
4.3.1 Kumaraswamy's distribution 55
4.3.2 A modified Möbius distribution 57
4.3.3 Some other models 61
4.4 Illustrative example 62

Appendix **67**

Acknowledgement **76**

Bibliography **77**

Chapter 1

Introduction and preliminaries

1.1 Introduction

Statistics for data which include angular observations is known as directional statistics. Bivariate circular data such as wind directions measured at two points in time are modeled by using bivariate circular distributions or distributions on the torus. Likewise circular-linear data are modeled by using distributions on the cylinder and disc. In this thesis, we propose some extensions of distributions on the torus, cylinder and disc in the framework of directional statistics.

In Chapter 1, we introduce some preliminary notions and existing models for circular data. A new circular distribution (Wang and Shimizu, 2012) is also introduced, which is obtained by applying the Möbius transformation to a univariate cardioid random variable. The distribution function, trigonometric moments, and conditions for unimodality and symmetry are studied. Kato and Jones (2010) study a family of distributions which is obtained by applying the Möbius transformation to a von Mises random variable, and we discuss the relationship between our model and the Kato–Jones model. The bivariate circular case (Wang and Shimizu, 2012) will be treated in Chapter 2. We propose a bivariate cardioid distribution which is generated from a circular-circular structural model linked with Möbius transformation or a method of trivariate reduction. The joint probability density function, trigonometric moments and circular-circular correlation coefficient are explicitly expressed. An illustration is given for wind direction data at 6 a.m. and noon as an application of the bivariate cardioid distribution.

Chapters 3 and 4 are based on the paper by Wang, Shimizu and Uesu (to appear). In Chapter 3, we propose distributions on the cylinder. A distribution generated from a combination of von Mises and transformed Kumaraswamy distributions is proposed as an extension of the Johnson and Wehrly (1978) model. The marginal and conditional distributions of the proposed distribution are given. A distribution using the method of generating a cylindrical distribution with specified marginals is also proposed. Two illustrative examples are given for the magnitudes and the angles calculated from the epicenters of earthquakes and the movements of blue periwinkles. In addition, it is remarked that the circular-linear regression models by SenGupta and Ugwuowo (2006) are derived from the conditional distributions of cylindrical distributions. An application of the local influence method will be treated in Appendix.

In Chapter 4, we generate skew or asymmetric distributions on the disc by using the Möbius transformation and modified Möbius transformations as extensions of the Möbius distribution proposed by Jones (2004). The new distributions called the modified Möbius distributions have six parameters. They can be reduced to the Möbius and uniform distributions as special cases, but many members of the family are skew distributions for both the linear and the angular random variables. Some properties such as the joint probability and marginal density functions of the proposed distributions are obtained. As an illustrative example, we fit these models to data of Johnson and Wehrly (1977), which consist of values of ozone concentration and wind direction.

1.2 Preliminary notions for circular data

1.2.1 Mean direction, mean resultant length and circular variance

Directions in the plane can be regarded as unit vectors $\boldsymbol{x}$. Suppose that unit vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ are given corresponding with angles θ_j, i.e. $\boldsymbol{x}_j = (\cos\theta_j, \sin\theta_j)$ for $j = 1, \dots, n$. The sample mean direction $\bar{\theta}$ of $\theta_1, \dots, \theta_n$ is the direction of the resultant vector of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. The Cartesian coordinates of the center of mass are $(\bar{C}, \bar{S})$, where

$$\bar{C} = \frac{1}{n}\sum_{j=1}^{n} \cos\theta_j, \quad \bar{S} = \frac{1}{n}\sum_{j=1}^{n} \sin\theta_j.$$

Thus $\bar{\theta}$ satisfies the equations

$$\bar{C} = \bar{R}\cos\bar{\theta}, \quad \bar{S} = \bar{R}\sin\bar{\theta},$$

where $\bar{R}$ is the mean resultant length which is given by

$$\bar{R} = \left(\bar{C}^2 + \bar{S}^2\right)^{1/2}$$

and expressible as

$$\bar{R} = \frac{1}{n}\sum_{j=1}^{n} \cos(\theta_j - \bar{\theta}).$$

The range of the mean resultant length $\bar{R}$ is $[0, 1]$ and note that $\bar{\theta}$ is not defined when $\bar{R} = 0$, and can be explicitly represented as

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S}/\bar{C}), & \bar{C} > 0,\ \bar{S} \geq 0, \\ \pi/2, & \bar{C} = 0,\ \bar{S} > 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + \pi, & \bar{C} < 0, \\ 3\pi/2, & \bar{C} = 0,\ \bar{S} < 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + 2\pi, & \bar{C} > 0,\ \bar{S} < 0. \end{cases}$$

The sample mean resultant length $\bar{R}$ is a statistic which measures the concentration of the data. In general, $\bar{R}$ is more important than other measures of dispersion in the field of directional statistics. However, similarly to the sample variance for data in the real line, there is a concept about the sample circular variance, which is defined as

$$V = 1 - \bar{R},$$

whose range is $[0, 1]$. Note that Batschelet (1981) and Jammalamadaka and SenGupta (2001) refer to $2(1-\bar{R})$, because it is approximated by $\sum_{j=1}^{n}(\theta_j - \bar{\theta})^2/n$ when n is large. Here, we give a simple proof for this approximation:

$$2(1-\bar{R}) = 2\left\{1 - \frac{1}{n}\sum_{j=1}^{n}\cos(\theta_j - \bar{\theta})\right\} = 4\,\frac{1}{n}\sum_{j=1}^{n}\sin^2\left(\frac{\theta_j - \bar{\theta}}{2}\right) \approx \frac{1}{n}\sum_{j=1}^{n}(\theta_j - \bar{\theta})^2.$$

1.2.2 Trigonometric moments, circular skewness and kurtosis

The sample pth trigonometic moment about zero mean direction for $p = 0, 1, 2, \ldots$ is defined as

$$m_p = a_p + \mathrm{i}\, b_p, \quad \mathrm{i} = \sqrt{-1}$$

where

$$a_p = \frac{1}{n}\sum_{j=1}^{n} \cos p\theta_j, \quad b_p = \frac{1}{n}\sum_{j=1}^{n} \sin p\theta_j.$$

Then

$$m_p = \bar{R}_p\, \mathrm{e}^{\mathrm{i}\bar{\theta}_p},$$

where $\bar{\theta}_p$ and $\bar{R}_p$ denote the sample mean direction and sample mean resultant length of $p\theta_1, \ldots, p\theta_n$. The sample circular skewness and kurtosis are defined by

$$\hat{s} = \frac{\bar{R}_2 \sin(\bar{\theta}_2 - 2\bar{\theta})}{(1-\bar{R})^{3/2}}$$

and

$$\hat{k} = \frac{\bar{R}_2 \cos(\bar{\theta}_2 - 2\bar{\theta}) - \bar{R}^4}{(1-\bar{R})^{3/2}}$$

respectively.

1.2.3 Distribution and probability density functions

Suppose that an initial direction and an orientation of the unit circle have been chosen. Then the distribution function F of a random angle Θ is defined as

$$F(x) = \Pr(-\pi < \Theta \leq x), \quad -\pi \leq x < \pi,$$

and it is extended to the function on the whole real line given by

$$F(x+2\pi) - F(x) = 1, \quad -\infty < x < \infty.$$

The second equation means that any arc of length 2π on the circle has probability 1. For $\alpha - \pi \leq \beta \leq \alpha + \pi$,

$$\Pr(\alpha < \Theta \leq \beta) = F(\beta) - F(\alpha) = \int_\alpha^\beta dF(x).$$

By definition

$$F(-\pi) = 0, \quad F(\pi) = 1,$$

If the distribution function F is absolutely continuous then it has a probability density function f such that

$$\int_\alpha^\beta f(\theta) d\theta = F(\beta) - F(\alpha), \quad -\infty < \alpha \leq \beta < \infty.$$

A function f is the probability density function of an absolutely continuous distribution function if and only if

(i) $f(\theta) \geq 0$ almost everywhere on $[-\pi, \pi)$,

(ii) $\int_{-\pi}^{\pi} f(\theta) d\theta = 1$,

(iii) $f(\theta + 2\pi) = f(\theta)$ almost everywhere on $[-\pi, \pi)$.

1.2.4 Characteristic function

The characteristic function of a linear random variable X is defined as $\phi_t = \mathrm{E}\left(\mathrm{e}^{\mathrm{i}tX}\right)$. For angular case, since Θ and $\Theta + 2\pi$ represent the same direction, it holds that

$$\mathrm{E}\left(\mathrm{e}^{\mathrm{i}t\Theta}\right) = \mathrm{E}\left\{\mathrm{e}^{\mathrm{i}t(\Theta+2\pi)}\right\} = (\cos 2\pi t + \mathrm{i} \sin 2\pi t)\mathrm{E}\left(\mathrm{e}^{\mathrm{i}t\Theta}\right),$$

and it is necessary to restrict t to integer values. The characteristic function of a random angle Θ is given by

$$\phi_p = \mathrm{E}\left(\mathrm{e}^{\mathrm{i}p\Theta}\right) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}p\theta} dF(\theta), \quad p = 0, \pm 1, \pm 2, \ldots.$$

Thus,

$$\phi_0 = 1, \quad \bar{\phi}_p = \phi_{-p}, \quad |\phi_p| \leq 1,$$

where $\bar{\phi}_p$ denotes the complex conjugate of ϕ_p. If we write

$$\phi_p = \alpha_p + \mathrm{i}\,\beta_p,$$

where

$$\alpha_p = \mathrm{E}(\cos p\Theta) = \int_{-\pi}^{\pi} \cos p\theta dF(\theta)$$

and

$$\beta_p = \mathrm{E}(\sin p\Theta) = \int_{-\pi}^{\pi} \sin p\theta dF(\theta),$$

then

$$\alpha_{-p} = \alpha_p, \quad |\alpha_p| \leq 1, \quad \beta_{-p} = -\beta_p, \quad |\beta_p| \leq 1,$$

and ϕ_p is called the pth trigonometric moment, and α_p and β_p the pth cosine and sine moments respectively. Furthermore, if $\sum_{p=1}^{\infty}(\alpha_p^2 + \beta_p^2)$ is convergent then the variable Θ has a density f which is defined almost everywhere by a Fourier series

$$f(\theta) = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \phi_p \mathrm{e}^{-\mathrm{i}p\theta} = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{p=1}^{\infty} (\alpha_p \cos p\theta + \beta_p \sin p\theta) \right\}.$$

1.3 Some existing distributions on the circle

In this section we introduce the most fundamental continuous distributions on the unit circle. Except the uniform distribution, all of the distributions mentioned in this section are symmetric and unimodal.

1.3.1 Uniform distribution

The most basic distribution on the circle is the uniform distribution which has probability density function

$$f(\theta) = \frac{1}{2\pi}, \quad -\pi \leq \theta < \pi.$$

It is often used as the null model. The trigonometric moment of the uniform distribution is

$$\phi_p = \begin{cases} 1, & p = 0, \\ 0, & p \neq 0. \end{cases}$$

Thus, the mean resultant length ρ is 0. This means that there is no concentration about any particular direction. Let $\Theta_1, \ldots, \Theta_n$ be independent and identically distributed as the circular uniform distribution with common characteristic function ϕ_p. then the characteristic function of the sum $S_n = \Theta_1 + \cdots + \Theta_n$ is

$$\phi(n, p) = \phi_p^n = \begin{cases} 1, & p = 0, \\ 0, & p \neq 0, \end{cases}$$

which is the characteristic function of the circular uniform distribution. It follows from the uniqueness property that S_n is uniformly distributed on the circle. Furthermore, the sum of indenpendent circular uniform and any other distributions is the circular uniform distribution.

1.3.2 Cardioid distribution

The cardioid distribution $\mathrm{C}(\mu,\rho)$ with parameters μ $(-\pi \leq \mu < \pi)$ and ρ $(0 \leq \rho \leq 1/2)$ has probability density function

$$f(\theta) = \frac{1}{2\pi}\{1 + 2\rho\,\cos(\theta - \mu)\}, \quad -\pi \leq \theta < \pi.$$

This form appears in Jeffreys (1961, p. 328), and the name is taken from the fact that the shape of $r = f(\theta)$ in polar coordinates resembles a heart. In the literature, it is often assumed that $|\rho| \leq 1/2$, but we restrict ρ to be non-negative to avoid problems of non-identifiability. The distribution reduces to a circular uniform distribution if $\rho = 0$.

The probability density function is symmetric about μ and unimodal. If a random variable Θ is distributed as $\mathrm{C}(\mu,\rho)$, the characteristic function or pth trigonometric moment is $\mathrm{E}(\mathrm{e}^{\mathrm{i}p\theta}) = 1$ if $p = 0$, $\rho\,\mathrm{e}^{\mathrm{i}\mu}$ if $p = 1$ and 0 if $p \geq 2$, where $\mathrm{i} = \sqrt{-1}$, so that the mean direction of Θ is μ, the mean resultant length is ρ and the circular variance is $1 - \rho$. The cardioid distribution never degenerates to a distribution concentrated at only one point whatever the parameters be chosen. If Θ_1 is distributed as a cardioid distribution with mean direction μ_1 and mean resultant length ρ_1, and Θ_2 (not necessarily distributed as a cardioid) is independent of Θ_1 and has mean direction μ_2 and mean resultant length ρ_2, then the sum $\Theta_1 + \Theta_2$ is distributed as a cardioid distribution with mean direction $\mu_1 + \mu_2$ and mean resultant length $\rho_1\rho_2$.

1.3.3 von Mises distribution

The von Mises distribution $\mathrm{VM}(\mu,\kappa)$ with parameters μ $(-\pi \leq \mu < \pi)$ and κ $(\kappa \geq 0)$ has probability density function

$$f(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)}\,\mathrm{e}^{\kappa\cos(\theta-\mu)}, \quad -\pi \leq \theta < \pi,$$

where I_p denotes the modified Bessel function of the first kind and order p, which can be defined by

$$I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos{(p\theta)}\, \mathrm{e}^{\kappa \cos\theta} d\theta.$$

The modified Bessel function $I_p(\kappa)$ has an infinite power series

$$I_p(\kappa) = \sum_{r=0}^{\infty} \frac{1}{\Gamma(p+r+1)r!} \left(\frac{\kappa}{2}\right)^{2r+p}.$$

The parameter μ is the mean direction and the parameter κ is the concentration parameter. The von Mises distribution reduces to a circular uniform distribution if $\kappa = 0$ and concentrates into a point when $\kappa \to \infty$. Since $\mathrm{VM}(\mu, \kappa)$ is symmetric about $\theta = \mu$, the pth sine moment about μ and the pth cosine moment about μ are respectively

$$\begin{aligned}
\bar{\beta}_p &= \mathrm{E}[\sin p(\Theta - \mu)] = 0, \\
\bar{\alpha}_p &= \mathrm{E}[\cos p(\Theta - \mu)] = \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} \cos p(\theta - \mu) \mathrm{e}^{\kappa \cos{(\theta-\mu)}} d\theta = \frac{I_p(\kappa)}{I_0(\kappa)},
\end{aligned}$$

and the mean resultant length ρ is $A(\kappa) = I_1(\kappa)/I_0(\kappa)$. The von Mises distribution $\mathrm{VM}(\mu, \kappa)$ has mode and antimode at μ and $\mu + \pi$ respectively. In particular, the maximum entropy distribution on the circle with given mean direction μ and given mean resultant length ρ is the von Mises distribution $\mathrm{VM}(\mu, \kappa)$, with $\kappa = A^{-1}(\rho)$.

1.3.4 Wrapped Cauchy distribution

Wrapping a distribution on the real line makes the corresponding distribution on the circle. More precisely, if X is a random variable on the line, the corresponding random variable X_w of the wrapped distribution is given by

$$X_w = X \pmod{2\pi}.$$

If X has distribution function F, then the distribution function F_w of X_w is

$$F_w(\theta) = \sum_{k=-\infty}^{\infty} \{F[\theta + (2k-1)\pi] - F[(2k-1)\pi]\}, \quad -\pi \le \theta < \pi.$$

In particular, if X has a probability density function f, then the corresponding probability density function f_w of X_w is

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} f[\theta + (2k-1)\pi].$$

Some properties of the wrapped distribution are:

(a) $(X+Y)_w = X_w + Y_w$.

(b) If the characteristic function of X is $\phi = \phi(p)$, then the characteristic function ϕ_p, $p = 0, \pm 1, \pm 2, \ldots$, of X_w is given by $\phi_p = \phi(p)$.

(c) If ϕ is integrable, then X has a density expressible as

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + (2k-1)\pi) = \frac{1}{2\pi}\left\{1 + 2\sum_{p=1}^{\infty}(\alpha_p \cos p\theta + \beta_p \sin p\theta)\right\}$$

with $\phi_p = \alpha_p + \mathrm{i}\beta_p$.

If a random variable X is distributed as a Cauchy distribution with location μ and scale $\gamma > 0$, then the wrapped Cauchy random variable $\Theta = X \pmod{2\pi}$ has probability density function

$$f(\theta) = \frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2 - 2\rho\cos(\theta-\mu)}, \quad -\pi \le \theta < \pi,$$

where $-\pi \leq \mu < \pi$, $\rho = \mathrm{e}^{-\gamma}$, $0 \leq \rho < 1$. The characteristic function is $\phi_p = \rho^{|p|}\mathrm{e}^{\mathrm{i}p\mu}$. In particular, the wrapped Cauchy distribution, denoted by $\mathrm{WC}(\mu, \rho)$, is symmetric about μ, which is the mean direction of this distribution, and the mean resultant length is ρ. As $\rho \to 0$ it tends to the circular uniform distribution, and as $\rho \to 1$, it is concentrated at point μ. The convolution of wrapped Cauchy distributions $\mathrm{WC}(\mu_1, \rho_1)$ and $\mathrm{WC}(\mu_2, \rho_2)$ is the wrapped Cauchy distribution $\mathrm{WC}(\mu_1 + \mu_2, \rho_1\rho_2)$.

1.3.5 Generalized cardioid distribution of Jones & Pewsey

Jones and Pewsey (2005) proposed a general family of symmetric unimodal distributions on the circle that incorporates all of the aforementioned distributions. The probability density function is

$$f_\psi(\theta) = \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi)\cos(\theta - \mu)\}^{1/\psi}}{2\pi P_{1/\psi}\left(\cosh(\kappa\psi)\right)}, \quad -\pi \leq \theta < \pi,$$

where $-\pi \leq \mu < \pi$, $\kappa \geq 0$, $-\infty < \psi < \infty$, and $P_{1/\psi}$ is the associated Legendre function of the first kind of degree $1/\psi$ and order 0 which is defined by

$$P_{1/\psi}(z) = \frac{1}{\pi}\int_0^\pi \left(z + \sqrt{z^2 - 1}\cos x\right)^{1/\psi} dx.$$

As special cases, the generalized cardioid distribution reduces to

(1) circular uniform distribution ($\kappa = 0$; $\psi \to \pm\infty$) ,

(2) von Mises distribution ($\psi \to 0$),

(3) wrapped Cauchy distribution ($\psi = -1$),

(4) cardioid distribution ($\psi = 1$),

(5) Cartwright's (1963) power-of-cosine distribution ($\psi = 1/n, \kappa \to \infty$),

(6) Shimizu and Iida's (2002) circular t-distribution $(-1 < \psi < 0)$.

1.4 A new family of distributions on the circle

Kato and Jones (2010) study a family of distributions which is derived from the Möbius, or linear fractional, transformation of a von Mises random variable. Likewise, Wang and Shimizu (2012) propose a new family of distributions which is derived from the Möbius transformation of a cardioid random variable. We give the results by Wang and Shimizu (2012) in this section.

1.4.1 Möbius transformation

The Möbius transformation from the unit circle onto itself is defined by

$$M(\xi) = \arg\left\{\beta \frac{\mathrm{e}^{\mathrm{i}\xi} + \alpha}{1 + \overline{\alpha}\mathrm{e}^{\mathrm{i}\xi}}\right\} = 2\ \arctan\left\{\left(\frac{1-\rho_\alpha}{1+\rho_\alpha}\right)\tan\frac{1}{2}(\xi - \mu_\alpha)\right\} + (\mu_\alpha + \mu_\beta), \tag{1.1}$$

where $\alpha = \rho_\alpha \mathrm{e}^{\mathrm{i}\mu_\alpha}$ $(\rho_\alpha \geq 0,\ \rho_\alpha \neq 1,\ -\pi \leq \mu_\alpha < \pi)$, $\beta = \mathrm{e}^{\mathrm{i}\mu_\beta}$ $(-\pi \leq \mu_\beta < \pi)$ and $\overline{\alpha}$ denotes the complex conjugate of α. In (1.1), μ_β is a rotation parameter and, under the assumption that $\beta = 1$, α $(\neq 0)$ is the parameter that attracts the points on the unit circle toward a point $\alpha/|\alpha|$ on the unit circle, except the point $\xi = -\alpha/|\alpha|$, which is invariant under the transformation. Some properties of the Möbius transformation are discussed in Kato, Shimizu and Shieh (2008) in the context of circular-circular regression.

1.4.2 Möbius transformation of a cardioid random variable

Probability density function

When a random variable ξ is distributed as $\mathrm{C}(\mu, \rho)$, we consider the distribution of a random variable $\eta = M(\xi)$ transformed by the Möbius transformation (1.1). Since $\eta = M(\xi)$ is a one-to-one mapping and

$$\frac{\mathrm{d}}{\mathrm{d}\eta} M^{-1}(\eta) = \frac{1-\rho_\alpha^2}{1+\rho_\alpha^2 - 2\rho_\alpha \cos(\eta - \mu_\alpha - \mu_\beta)},$$

the probability density function of η is

$$g(\eta) = \frac{|1-\rho_\alpha^2| h(\eta)}{2\pi\{1+\rho_\alpha^2 - 2\rho_\alpha \cos(\eta - \gamma_2)\}}, \tag{1.2}$$

where

$$h(\eta) = 1 + 2\rho \left\{ \frac{\cos(\eta - \gamma_1 - \gamma_2) - 2\rho_\alpha \cos\gamma_1 + \rho_\alpha^2 \cos(\eta + \gamma_1 - \gamma_2)}{1 + \rho_\alpha^2 - 2\rho_\alpha \cos(\eta - \gamma_2)} \right\},$$

with four parameters γ_1 $(= \mu - \mu_\alpha)$, γ_2 $(= \mu_\alpha + \mu_\beta)$, ρ and ρ_α. The distribution is denoted by $\mathrm{MC}_1(\gamma_1, \gamma_2, \rho, \rho_\alpha)$ or MC_1 for short. Note that γ_2 is a location parameter. Figure 1 gives density plots for $\gamma_2 = 0$ and various combinations of the other three parameters γ_1, ρ and ρ_α. It illustrates the flexibility of this family. The distribution with density (1.2) reduces to a wrapped Cauchy distribution if $\rho = 0$.

Kato and Jones (2010) generated a family of distributions by applying the Möbius transformation to a von Mises distribution with mean direction μ and concentration parameter κ. The cardioid distribution is approximated by the von Mises distribution with small κ; an approximation $\mathrm{e}^x \approx 1 + x$ leads to equation (1.2) from the density (2) of Kato and Jones (2010). The above distribution by

Kato and Jones (2010) and MC_1 are special cases of the five-parameter family introduced in Section 8 of Kato and Jones (2010), an asymmetric extension of the family of Jones and Pewsey (2005). For MC_1, the distribution function is obtainable in an explicit form as can be seen in the following. The distribution function, cosine and sine moments, conditions for unimodality and symmetry of the resulting distribution are studied.

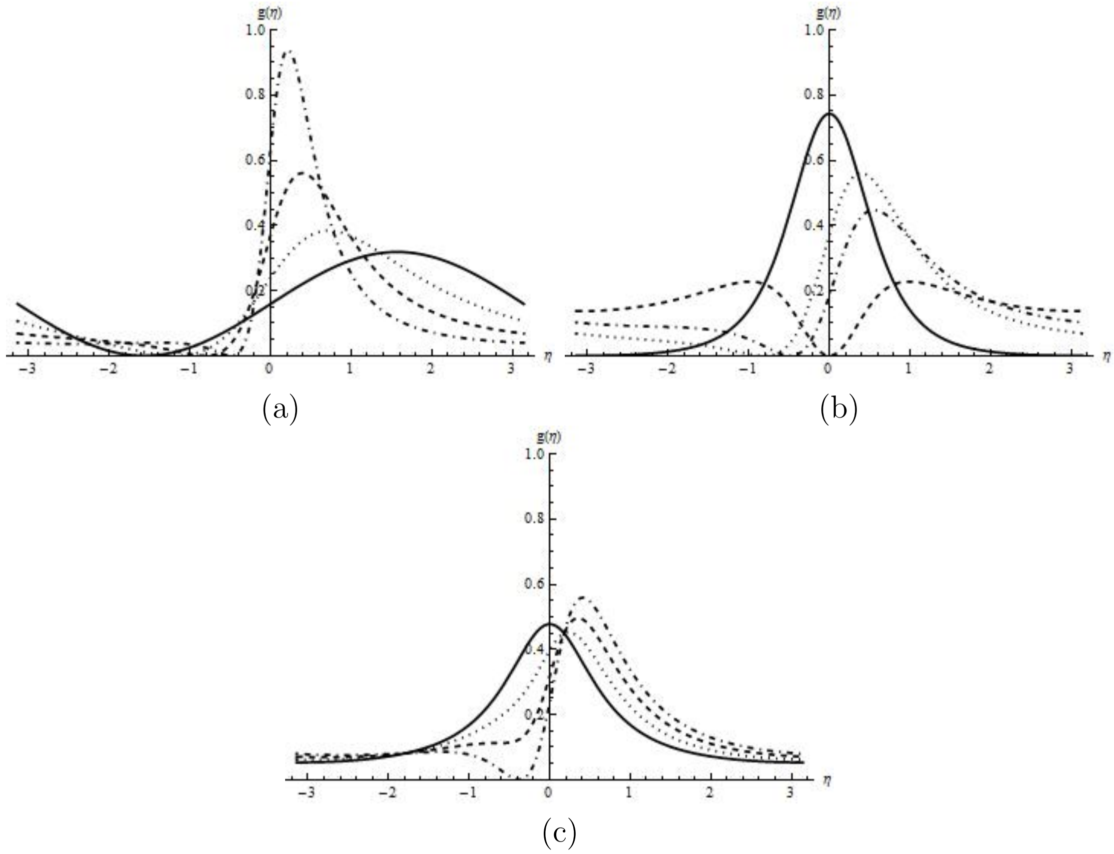


Figure 1.1: Density plots of (1.2) for $\gamma_2 = 0$, and (a) $\rho = 1/2$, $\gamma_1 = \pi/2$, and $\rho_\alpha = 0$ (solid), 1/5 (dotted), 2/5 (dashed), 3/5 (dot-dashed); (b) $\rho = 1/2$, $\rho_\alpha = 2/5$, and $\gamma_1 = 0$ (solid), $\pi/2$ (dotted), π (dashed), $3\pi/2$ (dot-dashed); and (c) $\rho_\alpha = 1/2$, $\gamma_1 = 2\pi/3$, and $\rho = 0$ (solid), 1/6 (dotted), 1/3 (dashed), 1/2 (dot-dashed). Unimodal/bimodal and symmetry/asymmetry depend upon the parameters chosen.

Distribution function

For $-\pi \leq \gamma_2 < 0$, the distribution function is

$$G(\eta) = \begin{cases} \frac{1}{2} + \frac{1}{\pi}\arctan\left\{\left|\frac{1+\rho_\alpha}{1-\rho_\alpha}\right| \tan\frac{\eta-\gamma_2}{2}\right\} + G_1, & -\pi \leq \eta < 0, \\ 1 + \frac{1}{\pi}\arctan\left\{\left|\frac{1-\rho_\alpha}{1+\rho_\alpha}\right| \cot\frac{-\eta+\gamma_2}{2}\right\} + G_1, & 0 \leq \eta < \pi, \end{cases}$$

and for $0 \leq \gamma_2 < \pi$, it is

$$G(\eta) = \begin{cases} \frac{1}{\pi}\arctan\left\{\left|\frac{1-\rho_\alpha}{1+\rho_\alpha}\right| \cot\frac{-\eta+\gamma_2}{2}\right\} + G_1, & -\pi \leq \eta < 0, \\ \frac{1}{2} + \frac{1}{\pi}\arctan\left\{\left|\frac{1+\rho_\alpha}{1-\rho_\alpha}\right| \tan\frac{\eta-\gamma_2}{2}\right\} + G_1, & 0 \leq \eta < \pi, \end{cases}$$

where

$$\begin{aligned} G_1 &= \frac{1}{\pi}\arctan\left\{\left|\frac{1-\rho_\alpha}{1+\rho_\alpha}\right| \tan\left(\gamma_2/2\right)\right\} \\ &\quad + \frac{2\rho\,|1-\rho_\alpha^2|\cos\left(\eta/2\right)}{\pi}\left\{\rho_\alpha \sin\left(\frac{\eta}{2}+\gamma_1\right) + \rho_\alpha \sin\left(\frac{\eta}{2}-\gamma_1\right)\right. \\ &\quad \left. + \rho_\alpha^2 \sin\left(\frac{\eta}{2}+\gamma_1-\gamma_2\right) + \sin\left(\frac{\eta}{2}-\gamma_1-\gamma_2\right)\right\} \\ &\quad / \left[\left\{1+\rho_\alpha^2 - 2\rho_\alpha\cos\left(\eta-\gamma_2\right)\right\}\left\{1+\rho_\alpha^2+2\rho_\alpha\cos\gamma_2\right\}\right]. \end{aligned}$$

If $\rho = 0$ and $\gamma_2 = 0$, then

$$G(\eta) = 1/2 + (1/\pi)\arctan\left\{|(1+\rho_\alpha)/(1-\rho_\alpha)|\tan(\eta/2)\right\}$$

for $-\pi \leq \eta < \pi$, which is the distribution function of the wrapped Cauchy distribution with zero mean direction.

Moments and other characteristics

When $0 \leq \rho_\alpha < 1$, the cosine and sine moments of MC_1 are given by

$$\mathrm{E}(\cos p\eta) = p\rho\rho_\alpha^{p-1}(1-\rho_\alpha^2)\cos(p\gamma_2+\gamma_1) + \rho_\alpha^p \cos p\gamma_2 \tag{1.3}$$

and

$$\mathrm{E}(\sin p\eta) = p\rho\rho_\alpha^{p-1}(1-\rho_\alpha^2)\sin(p\gamma_2+\gamma_1) + \rho_\alpha^p \sin p\gamma_2 \tag{1.4}$$

for $p = 0, \pm 1, \ldots$. Thus, the mean direction of MC_1 is

$$\mu_{\mathrm{MC}} = \gamma_2 + \arctan\left\{\frac{\rho\,(1-\rho_\alpha^2)\sin\gamma_1}{\rho_\alpha + \rho\,|1-\rho_\alpha^2|\cos\gamma_1}\right\}, \tag{1.5}$$

and the mean resultant length is

$$\rho_{\mathrm{MC}} = \sqrt{\rho^2(1-\rho_\alpha^2)^2 + \rho_\alpha^2 + 2\rho\rho_\alpha(1-\rho_\alpha^2)\cos\gamma_1}. \tag{1.6}$$

It holds that $\rho_{\mathrm{MC}} \leq \rho(1-\rho_\alpha^2) + \rho_\alpha < 1$.

For the case $\rho_\alpha > 1$, we only need to convert the ρ_α to $1/\rho_\alpha$ and convert the γ_1 to $-\gamma_1$ in the formulas (1.3), (1.4), (1.5) and (1.6). The cardioid distribution is closed under convolution, but generally MC_1 does not have this property.

Conditions for unimodality

We discuss conditions for unimodality of the probability density function of MC_1. The procedure for proof is similar to that of Kato and Jones (2010).

Let $a_1 = \rho\rho_\alpha(1+\rho_\alpha^2)\cos\gamma_1 - \rho_\alpha^2$, $a_2 = 2\rho\rho_\alpha(1-\rho_\alpha^2)\sin\gamma_1$, $a_3 = \rho(1+\rho_\alpha^4 - 6\rho_\alpha^2)\cos\gamma_1 + \rho_\alpha(1+\rho_\alpha^2)$, $a_4 = -\rho(1-\rho_\alpha^4)\sin\gamma_1$ and $a_5 = 2\rho\rho_\alpha(1-\rho_\alpha^2)\sin\gamma_1$. Then the derivative of $g(\eta)$ with respect to η is

$$g'(\eta) = -\frac{|1-\rho_\alpha^2|}{\pi\{1+\rho_\alpha^2 - 2\rho_\alpha\cos(\eta-\gamma_2)\}^3}\frac{(b_0+b_1t+b_2t^2+b_3t^3+b_4t^4)}{(1+t^2)^2},$$

where $t = \tan(\eta/2)$ and the coefficients b_0, b_1, b_2, b_3 and b_4 are given by

$$
\begin{aligned}
b_0 &= a_5 + a_4 \cos\gamma_2 - a_3 \sin\gamma_2 + a_2\{\sin\gamma_2\}^2 - a_1 \sin 2\gamma_2,\\
b_1 &= 4a_1 \cos 2\gamma_2 + 2a_3 \cos\gamma_2 + 2a_4 \sin\gamma_2 - 2a_2 \sin 2\gamma_2,\\
b_2 &= 2a_5 + 4a_2 - 6a_2\{\sin\gamma_2\}^2 + 6a_1 \sin 2\gamma_2,\\
b_3 &= 2a_3 \cos\gamma_2 - 4a_1 \cos 2\gamma_2 + 2a_4 \sin\gamma_2 + 2a_2 \sin 2\gamma_2,\\
b_4 &= a_5 - a_4 \cos\gamma_2 + a_3 \sin\gamma_2 + a_2\{\sin\gamma_2\}^2 - a_1 \sin 2\gamma_2.
\end{aligned}
$$

Thus, the modality of MC_1 to be solved reduces to find the root of the quartic equation $b_0 + b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 = 0$. It follows from Ferrari's method that the equation has two real roots and two complex roots if the discriminant D is negative and then the probability density function is unimodal, while it has four real roots or four complex roots if D is positive and then the density is bimodal. Here D is

$$
\begin{aligned}
D = {} & b_1^2\, b_2^2\, b_3^2 - 4\, b_1^3\, b_3^3 - 4\, b_1^2\, b_2^3\, b_4 + 18\, b_1^3\, b_2\, b_3\, b_4 - 27\, b_1^4\, b_4^2 + 256\, b_0^3\, b_4^3 \\
& + b_0\, (-4\, b_2^3\, b_3^2 + 18\, b_1\, b_2\, b_3^3 + 16\, b_2^4\, b_4 - 80\, b_1\, b_2^2\, b_3\, b_4 - 6\, b_1^2\, b_3^2\, b_4 \\
& + 144\, b_1^2\, b_2\, b_4^2) + b_0^2\, (-27\, b_3^4 + 144\, b_2\, b_3^2\, b_4 - 128\, b_2^2\, b_4^2 - 192\, b_1\, b_3\, b_4^2)
\end{aligned}
\tag{1.7}
$$

as already given in Kato and Jones (2010). Figure 2 shows areas of positive D (bimodality for density) and negative D (unimodality for density) for two pairs of parameters. For example, Figure 2(a) suggests that the density (1.2) with $\gamma_2 = 0$ and $\rho = 1/2$ can be bimodal for any γ_1 when ρ_α is close to 1 and when approximately $\rho_\alpha > 0.2$ for $\gamma_1 = \pi$. The case $\gamma_2 = 0,\ \rho = 1/2,\ \rho_\alpha = 2/5,\ \gamma_1 = \pi$ in Figure 1(b) gives an example of bimodal density.

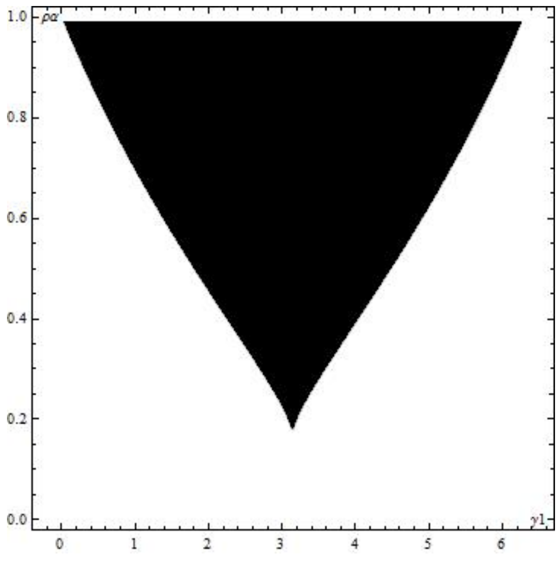
In a special case when $\gamma_1 = 0$, the density and its derivative are of the forms

$$g(\eta) = \frac{|1-\rho_\alpha^2|}{2\pi\left\{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)\right\}}\left\{1+2\rho\frac{(1+\rho_\alpha^2)\cos(\eta-\gamma_2)-2\rho_\alpha}{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)}\right\}$$

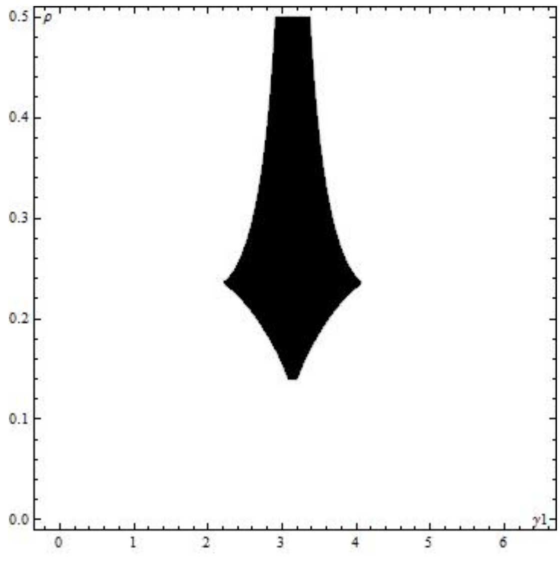
and

$$g^{'}(\eta) = -\frac{|1-\rho_\alpha^2|\left\{2a_1\cos(\eta-\gamma_2)+a_3\right\}}{2\pi\left\{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)\right\}}\sin(\eta-\gamma_2).$$

Since $2a_1\cos(\eta-\gamma_2)+a_3>0$ for all η, we find stationary points which satisfy $\sin(\eta-\gamma_2)=0$, i.e. $\eta=\gamma_2$ and $\eta=\gamma_2\pm\pi$ (mod 2π). Since $g''(\gamma_2)=-\{\rho_\alpha+\rho(1+4\rho_\alpha+\rho_\alpha^2)\}|1-\rho_\alpha^2|/\{\pi(1-\rho_\alpha)^4\}\le 0$ and $g''(\gamma_2\pm\pi)=\{\rho_\alpha+\rho(1-4\rho_\alpha+\rho_\alpha^2)\}|1-\rho_\alpha^2|\}/\{\pi(1+\rho_\alpha)^4\}=\{\rho_\alpha(1-2\rho)+\rho(1-\rho_\alpha)^2)|1-\rho_\alpha^2|\}/\{\pi(1+\rho_\alpha)^4\}\ge 0$, we know that $g(\eta)$ gives a maximum value at point γ_2 and a minimum value at point $\gamma_2\pm\pi$ (mod 2π).



(a)



(b)

Figure 1.2: Discriminant (1.7) for density (1.2) when $\gamma_2 = 0$, as functions of (a) (γ_1, ρ_α) when $\rho = 1/2$, and (b) (γ_1, ρ) when $\rho_\alpha = 1/4$. The region of positive discriminant/bimodality is shown in black, and negative discriminant/unimodality in white.

Conditions for symmetry

Different from Kato and Jones (2010), we start with a measure of skewness (Mardia and Jupp, 1999, Section 3.4) for circular random variable η defined by $s = \mathrm{E}\{\sin 2(\eta - \mu_{\mathrm{MC}})\}/(1-\rho_{\mathrm{MC}})^{3/2}$, where μ_{MC} is the mean direction and ρ_{MC} the mean resultant length of η. The skewness of MC_1 is calculated as

$$\begin{aligned} s(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta) &= -2\rho_{\mathrm{MC}}^{-2}(1-\rho_{\mathrm{MC}})^{-3/2}\rho^2\rho_\alpha(1-\rho_\alpha^2)^2 \\ &\quad \times\{\rho(1-\rho_\alpha^2)+\rho_\alpha\cos\gamma_1\}\sin\gamma_1 \end{aligned} \tag{1.8}$$

for $0 \le \rho_\alpha < 1$ unless $\rho_{\mathrm{MC}} = 0$, and for the case $\rho_\alpha > 1$ we only need to convert the ρ_α to $1/\rho_\alpha$ in equation (1.8).

If the probability density function $g(\eta)$ is symmetric about μ_{MC}, the skewness is 0. This means that the following equation is necessarily satisfied for symmetry:

$$\rho^2\rho_\alpha(1-\rho_\alpha^2)^2\{\rho(1-\rho_\alpha^2)+\rho_\alpha\cos\gamma_1\}\sin\gamma_1 = 0.$$

Thus, we consider the four cases below.

(a) $\rho = 0$.

In this case, MC_1 becomes a symmetric wrapped Cauchy distribution with mean direction γ_2 and mean resultant length ρ_α for $0 \le \rho_\alpha < 1$ and $1/\rho_\alpha$ for $\rho_\alpha > 1$.

(b) $\rho_\alpha = 0$.

In this case, MC_1 becomes a symmetric cardioid distribution $\mathrm{C}(\gamma_1+\gamma_2, \rho)$.

(c) $\sin\gamma_1 = 0$.

In this case, it is obtained that $\gamma_1 = 0$ or $\gamma_1 = -\pi$. And the probability density function becomes

$$g(\eta) = \begin{cases} \frac{|1-\rho_\alpha^2|}{2\pi\{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)\}} \left[1 + 2\rho\frac{(1+\rho_\alpha^2)\cos(\eta-\gamma_2)-2\rho_\alpha}{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)}\right], & \gamma_1 = 0, \\ \frac{|1-\rho_\alpha^2|}{2\pi\{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)\}} \left[1 - 2\rho\frac{(1+\rho_\alpha^2)\cos(\eta-\gamma_2)-2\rho_\alpha}{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)}\right], & \gamma_1 = -\pi. \end{cases}$$

It is not difficult to find that $g(\gamma_2 + \eta) = g(\gamma_2 - \eta)$ for all η. Thus, when $\sin\gamma_1 = 0$, $g(\eta)$ is symmetric about γ_2.

(d) $\rho(1-\rho_\alpha^2) + \rho_\alpha \cos\gamma_1 = 0$.

In this case, we have $\rho = -\{\rho_\alpha/(1-\rho_\alpha^2)\}\cos\gamma_1$, and the probability density function is

$$\begin{aligned} g(\eta) &= \frac{|1-\rho_\alpha^2|}{2\pi(1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2))}\left[1 - 2\frac{\rho_\alpha}{1-\rho_\alpha^2}\cos\gamma_1 \right. \\ &\qquad \left. \times \left\{\frac{\cos(\eta-\gamma_1-\gamma_2) - 2\rho_\alpha\cos\gamma_1 + \rho_\alpha^2\cos(\eta+\gamma_1-\gamma_2)}{1+\rho_\alpha^2-2\rho_\alpha\cos(\eta-\gamma_2)}\right\}\right]. \end{aligned}$$

For each x_0 there exists an η_0 such that $g(x_0-\eta_0) - g(x_0+\eta_0) \neq 0$, and $g(\eta)$ is not symmetric in this case. An explicit example of an asymmetric density which nonetheless has zero circular skewness is seen in Figure 1(c): $\rho_\alpha = 1/2$, $\gamma_1 = 2\pi/3$ and $\rho = 1/3$.

Consequently, the probability density function $g(\eta)$ is symmetric if and only if $\rho = 0$ or $\rho_\alpha = 0$ or $\gamma_1 = 0$ or $\gamma_1 = -\pi$.

Chapter 2

Distributions on the torus

2.1 Introduction

Sometimes it is necessary to consider the joint distribution of two circular random variables, such as the relationship between a pair of wind directions, measured at two locations at same time. Such a pair of realization may be identified with a point on the torus and thus the joint distributions of two circular random variables are called distributions on the torus or bivariate circular distributions.

There are many bivariate circular distributions discussed in the literature. Some methods to get bivariate circular distributions and existing parametric models are introduced in Section 2.2. Bivariate von Mises–Fisher distributions were introduced by Mardia (1975). A theorem for generating families of bivariate circular distributions with specified marginal distributions was proposed by Wehrly and Johnson (1979), and a bivariate circular distribution with von Mises marginals was proposed by Shieh and Johnson (2005).

In Section 2.3, we propose a bivariate cardioid distribution generated from a circular-circular structural model linked with Möbius transformation, which can be viewed as a method of trivariate reduction. An explicit form of the joint probability density function and its trigonometric moments are obtained. An illustration is given in Section 2.4 as an application of the bivariate cardioid distribution to the wind direction data (Johnson and Wehrly, 1977) at 6 a.m. and 12 noon measured each day at a weather station in Milwaukee for 21 consecutive days.

2.2 Some existing distributions on the torus

A family of bivariate von Mises–Fisher distributions was introduced by Mardia (1975) (see also Mardia and Jupp, 1999, Section 3.7.1). Consider the joint distribution of two circular random variables Θ_1 and Θ_2. Then (Θ_1, Θ_2) takes values on the unit torus. One useful set of distribution on the torus is the bivariate von Mises model (Mardia, 1975) with probability density function proportional to

$$\exp\left\{\kappa_1 \cos(\theta_1 - \mu_1) + \kappa_2 \cos(\theta_2 - \mu_2) + (\cos\theta_1, \sin\theta_1)\mathbf{A}(\cos\theta_2, \sin\theta_2)^T\right\},$$

where $\mathbf{A}$ is a 2×2 matrix. The marginal distributions of Θ_1 and Θ_2 are von Mises if and only if either $\mathbf{A} = \mathbf{0}$ or $\kappa_1 = \kappa_2 = 0$ and $\mathbf{A}$ is a multiple of an orthogonal matrix (so that Θ_1 and Θ_2 are uniformly distributed). A submodel of this class was discussed by Jupp and Mardia (1980), and another submodel was considered by Rivest (1988), whose subsets were investigated by Singh et al. (2002) and Mardia et al. (2007). This class of distributions are maximum entropy distributions (see also Jammalamadaka and SenGupta, 2001, Section 2.3.1).

Wehrly and Johnson (1979) have given a theorem for generating families of bivariate circular distributions with specified marginal distributions. The theorem is given below.

Theorem. Let $f_1(\theta)$ and $f_2(\eta)$ be specified densities on the circle and $F_1(\theta)$ and $F_2(\eta)$ be their distribution functions. Also, let $g(\cdot)$ be a density on the circle. Then

$$f(\theta, \eta) = 2\pi g \left[2\pi \left\{F_1(\theta) - F_2(\eta)\right\}\right] f_1(\theta) f_2(\eta)$$

and

$$f(\theta, \eta) = 2\pi g \left[2\pi \left\{F_1(\theta) + F_2(\eta)\right\}\right] f_1(\theta) f_2(\eta)$$

where $-\pi \leq \theta, \eta < \pi$, are densities on the torus having the specified marginal densities $f_1(\theta)$ and $f_2(\eta)$.

Shieh and Johnson (2005) study a bivariate model with von Mises marginals $f(\theta_j)$ for $j = 1, 2$, having probability density function:

$$f_{12}(\theta_1, \theta_2) = 2\pi f_1(\theta_1) f_2(\theta_2) \frac{1}{2\pi I_0(\kappa_{12})} \mathrm{e}^{\kappa_{12} \cos\left[2\pi\{F_1(\theta_1) - F_2(\theta_2)\} - \mu_{12}\right]}, \tag{2.1}$$

where $-\pi \leq \theta_1, \theta_2 < \pi$, $-\pi \leq \mu_{12} < \pi$, $\kappa_{12} \geq 0$, and $F_j(\theta_j) = \int_{-\pi}^{\theta_j} f_j(\theta) d\theta$. Some properties such as maximum likelihood estimation are investigated.

Furthermore, Kato and Shimizu (2008) provide four-variate continuous distributions on certain manifolds with specified bivariate marginal distributions. The theorem is applicable to the construction of distributions on two tori, cylinders or discs.

Bibliographic notes: As other literature, Baba (1981) proposed a wrapped multivariate normal distribution; SenGupta (2004) investigated bivariate circular distributions with the properties of maximum entropy and conditional specifications; Aronld and Strass (1991) characterize the class of bivariate distributions such that the conditional distributions belong to any specified exponential families. Kato (2009) proposed a bivariate circular distribution with circular uniform marginals and wrapped Cauchy conditional distributions; Shieh et al. (2011)

gave a family of bivariate generalized von Mises distributions, whose marginals are generalized von Mises distributions.

2.3 A family of bivariate cardioid distributions

2.3.1 Definition and joint probability density function

Wang and Shimizu (2012) consider a circular-circular structural model, similar to but different from the linear structural model (cf. Cheng and Van Ness, 1999, Section 1.1), or a model using a method of trivariate reduction, i.e. $\Theta = \xi + \delta$ and $\Phi = \eta + \varepsilon$, where ξ is a random variable which follows a cardioid distribution $\mathrm{C}(\mu, \rho)$ and η is linked to ξ with the Möbius transformation (1.1) as $\eta = M(\xi)$. Here δ and ε are independently distributed as cardioid distributions $\mathrm{C}(0, \rho_1)$ and $\mathrm{C}(0, \rho_2)$ respectively and are independent of ξ. The distribution of η and its properties have been given in Section 1.4 of this thesis.

The joint probability density function of Θ and Φ, whose distribution is denoted by $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, \rho_1, \rho_2)$ or MC_2 for short, is calculated as

$$\begin{aligned} f_1(\theta, \phi) \;=\; & \left(\frac{1}{2\pi}\right)^2 [1 + 2\rho_2\rho_\alpha\{1 + 2\rho\rho_1 \cos(\theta - \mu)\} \cos(\mu_\alpha + \mu_\beta - \phi) \\ & + 2\rho\rho_1 \cos(\theta - \mu) + 2\rho\rho_2(1 - \rho_\alpha^2) \cos(\phi - \mu - \mu_\beta) \\ & + 2\rho_1\rho_2(1 - \rho_\alpha^2) \cos(\phi - \theta - \mu_\beta) \\ & - 2\rho\rho_1\rho_2\rho_\alpha(1 - \rho_\alpha^2) \cos(\theta - \phi + \mu - \mu_\alpha + \mu_\beta)], \end{aligned} \tag{2.2}$$

for $0 \le \rho_\alpha < 1$, and for $\rho_\alpha > 1$,

$$
\begin{aligned}
f_2(\theta,\phi) &= \left(\frac{1}{2\pi}\right)^2 \Bigg[1 + 2\rho_2\{1 + 2\rho\rho_1\cos(\theta-\mu)\}\frac{1}{\rho_\alpha}\cos(\mu_\alpha+\mu_\beta-\phi) \\
&\quad +2\rho\rho_1\cos(\theta-\mu) + 2\rho\rho_2\frac{\rho_\alpha^2-1}{\rho_\alpha^2}\cos(\phi+\mu-2\mu_\alpha-\mu_\beta) \\
&\quad +2\rho_1\rho_2\frac{\rho_\alpha^2-1}{\rho_\alpha^2}\cos(\theta+\phi-2\mu_\alpha-\mu_\beta) \\
&\quad -2\rho\rho_1\rho_2\frac{\rho_\alpha^2-1}{\rho_\alpha^3}\cos(\theta+\phi+\mu-3\mu_\alpha-\mu_\beta)\Bigg].
\end{aligned}
$$

Contour plots for $\rho = 1/4,\ \rho_1 = \rho_2 = 1/2,\ \mu_\alpha = \mu_\beta = 0$ and various combinations of the other two parameters μ and ρ_α are displayed in Figure 2.1.

The marginal probability density functions of Θ and Φ are $\mathrm{C}(\mu, \rho\rho_1)$ and $\mathrm{C}(\mu_\phi, \rho_\phi)$ from the reproductive property of cardioid distributions, where $\mu_\phi = \mu_{\mathrm{MC}}$ and $\rho_\phi = \rho_2\rho_{\mathrm{MC}}$. The conditional probability density functions are easily obtainable. If $\rho_1 = 0$ or $\rho_2 = 0$ in MC_2, then Θ is independent of Φ. If $\mu = \mu_\alpha$ or $\rho_\alpha = 0$, the joint probability density function has a mode at $(\mu, \mu+\mu_\beta)$, but we do not know whether the distribution is unimodal or not.

2.3.2 Trigonometric moments

The joint trigonometric moments $\Phi(p,q) = \mathrm{E}\{\mathrm{e}^{\mathrm{i}(p\Theta+q\Phi)}\},\ p,q = 0, \pm1, \pm2, \ldots,$ of MC_2 are listed below for the cases $\Phi(p,q) \neq 0$:

$$
\begin{aligned}
\Phi(p,0) &= \begin{cases} 1, \quad p=0, \\ \rho\rho_1\mathrm{e}^{\mathrm{i}\mu}, \quad p=1, \\ \rho\rho_1\mathrm{e}^{-\mathrm{i}\mu}, \quad p=-1, \end{cases} \\
\Phi(0,1) &= \begin{cases} \rho_2\mathrm{e}^{\mathrm{i}(\mu_\alpha+\mu_\beta)}\left\{\rho_\alpha + \rho(1-\rho_\alpha^2)\mathrm{e}^{\mathrm{i}(\mu-\mu_\alpha)}\right\}, \quad 0 \le \rho_\alpha < 1, \\ \rho_2\mathrm{e}^{\mathrm{i}(\mu_\alpha+\mu_\beta)}\left\{\frac{1}{\rho_\alpha} + \rho\frac{\rho_\alpha^2-1}{\rho_\alpha^2}\mathrm{e}^{\mathrm{i}(\mu_\alpha-\mu)}\right\}, \quad \rho_\alpha > 1, \end{cases}
\end{aligned}
$$

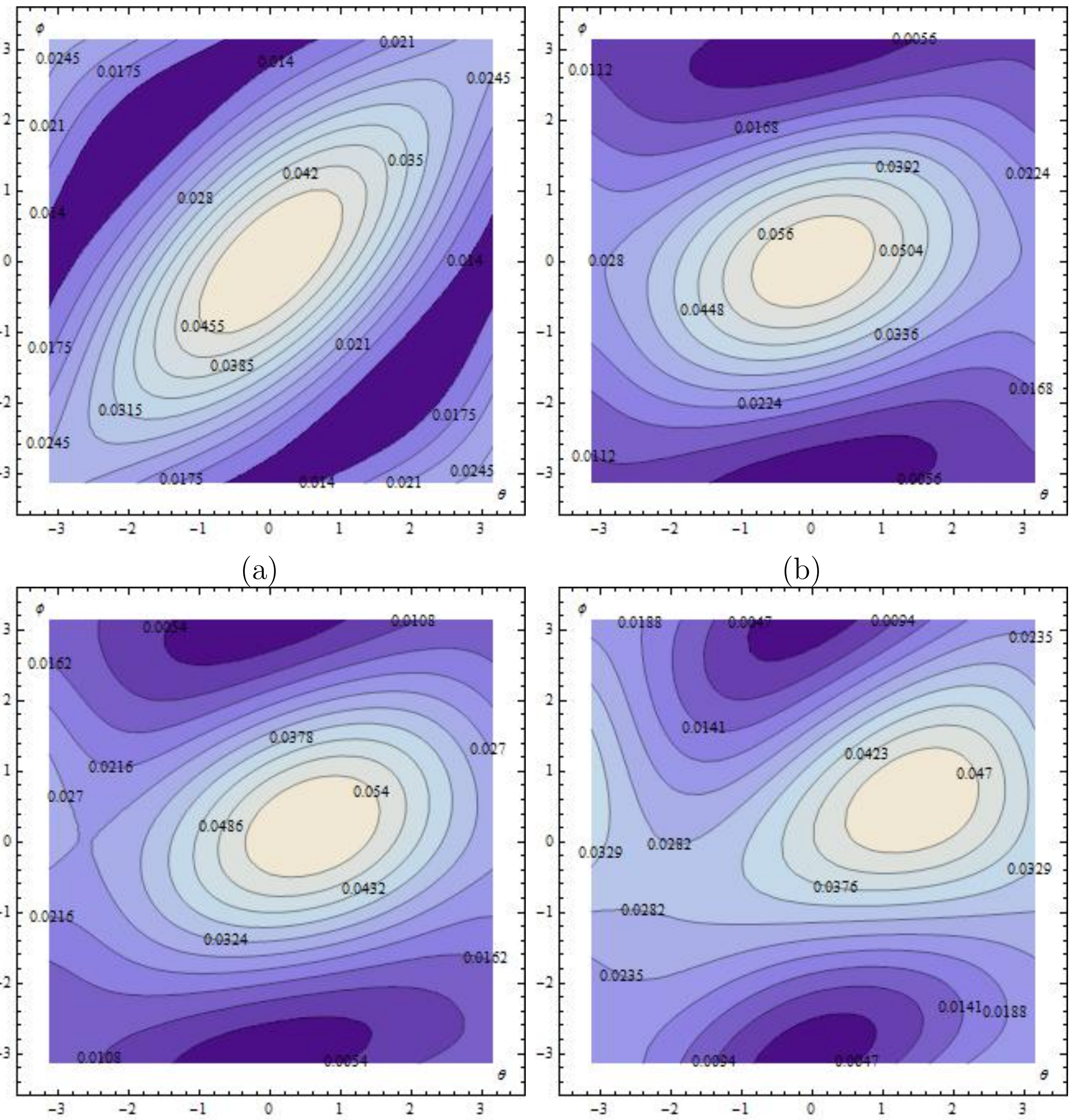


Figure 2.1: Contour plots of (2.2) for $\rho = 1/4,\ \rho_1 = \rho_2 = 1/2,\ \mu_\alpha = \mu_\beta = 0$ and (a) $\mu = 0,\ \rho_\alpha = 0$, (b) $\mu = 0,\ \rho_\alpha = 0.6$, (c) $\rho_\alpha = 0.6,\ \mu = \pi/4$ and (d) $\rho_\alpha = 0.6,\ \mu = 2\pi/3$.

$$\Phi(1,1) = \begin{cases} \rho\rho_1\rho_2\rho_\alpha \mathrm{e}^{\mathrm{i}(\mu_\alpha+\mu_\beta+\mu)}, \quad 0 \le \rho_\alpha < 1, \\ \rho_1\rho_2 \mathrm{e}^{\mathrm{i}(\mu_\alpha+\mu_\beta)} \left\{ \frac{\rho_\alpha^2-1}{\rho_\alpha^2} \mathrm{e}^{\mathrm{i}\mu_\alpha} + \frac{\rho}{\rho_\alpha} \mathrm{e}^{\mathrm{i}\mu} - \rho \frac{\rho_\alpha^2-1}{\rho_\alpha^3} \mathrm{e}^{\mathrm{i}(2\mu_\alpha-\mu)} \right\}, \quad \rho_\alpha > 1, \end{cases}$$

$$\Phi(1,-1) = \begin{cases} \rho_1\rho_2 \mathrm{e}^{-\mathrm{i}(\mu_\alpha+\mu_\beta)} \left\{ (1-\rho_\alpha^2)\mathrm{e}^{\mathrm{i}\mu_\alpha} + \rho\rho_\alpha \mathrm{e}^{\mathrm{i}\mu} - \rho\rho_\alpha(1-\rho_\alpha^2)\mathrm{e}^{\mathrm{i}(2\mu_\alpha-\mu)} \right\}, \\ \qquad 0 \le \rho_\alpha < 1, \\ \frac{\rho\rho_1\rho_2}{\rho_\alpha} \mathrm{e}^{-\mathrm{i}(\mu_\alpha+\mu_\beta-\mu)}, \quad \rho_\alpha > 1. \end{cases}$$

For the case $\Phi(0,-1)$, $\Phi(-1,-1)$, $\Phi(-1,1)$, we only need to convert the i to $-\mathrm{i}$ in the formulas $\Phi(0,1)$, $\Phi(1,1)$ and $\Phi(1,-1)$ respectively.

2.3.3 Correlation coefficient

A circular-circular correlation coefficient for bivariate data (cf. Mardia and Jupp, 1999, Section 11.2) is

$$\begin{aligned} r_{\theta,\phi}^2 = \; & \frac{1}{(1-r_1^2)(1-r_2^2)} \left\{ (r_{cc}^2 + r_{cs}^2 + r_{sc}^2 + r_{ss}^2) + 2(r_{cc}r_{ss} + r_{cs}r_{sc})r_1r_2 \right. \\ & \left. -2(r_{cc}r_{cs} + r_{sc}r_{ss})r_2 - 2(r_{cc}r_{sc} + r_{cs}r_{ss})r_1 \right\}, \end{aligned}$$

where $r_{cc} = \mathrm{corr}(\cos\Theta, \cos\Phi)$, $r_{cs} = \mathrm{corr}(\cos\Theta, \sin\Phi)$, $r_{sc} = \mathrm{corr}(\sin\Theta, \cos\Phi)$, $r_{ss} = \mathrm{corr}(\sin\Theta, \sin\Phi)$, $r_1 = \mathrm{corr}(\cos\Theta, \sin\Theta)$, $r_2 = \mathrm{corr}(\cos\Phi, \sin\Phi)$ are Pearson's correlation coefficients. The correlation coefficient for MC_2 is

$$r_{\theta,\phi}^2 = 2\rho_1^2\rho_2^2(t + 2\rho^2\rho_\alpha^2), \quad 0 \le \rho_\alpha < 1,$$

where $t = (1-\rho_\alpha^2)[(1-\rho_\alpha^2)(1+\rho^2\rho_\alpha^2) + 2\rho\rho_\alpha^3\cos(\mu-\mu_\alpha) - 2\rho^2\rho_\alpha^2\cos 2(\mu-\mu_\alpha)]$. For the case $\rho_\alpha > 1$, we only need to convert the ρ_α to $1/\rho_\alpha$. Figure 2.2 gives a plot of $t + 2\rho^2\rho_\alpha^2$ for $0 < \rho \le 1/2$ and $0 \le \rho_\alpha < 1$ when $\mu - \mu_\alpha = 0$. The values of $t + 2\rho^2\rho_\alpha^2$ approach 1 for any ρ as ρ_α goes to 0.

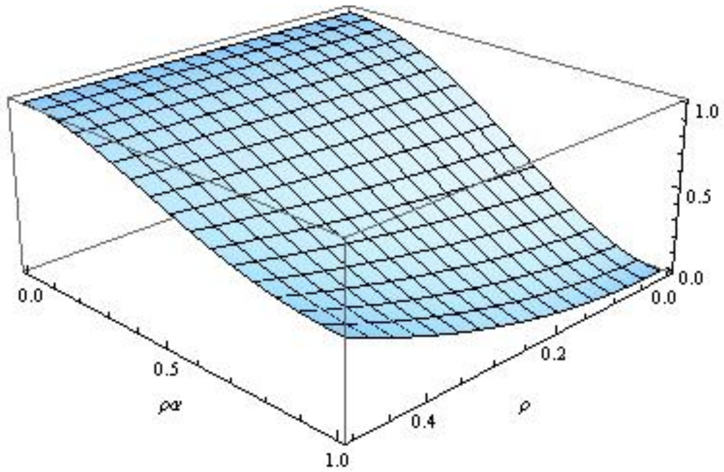


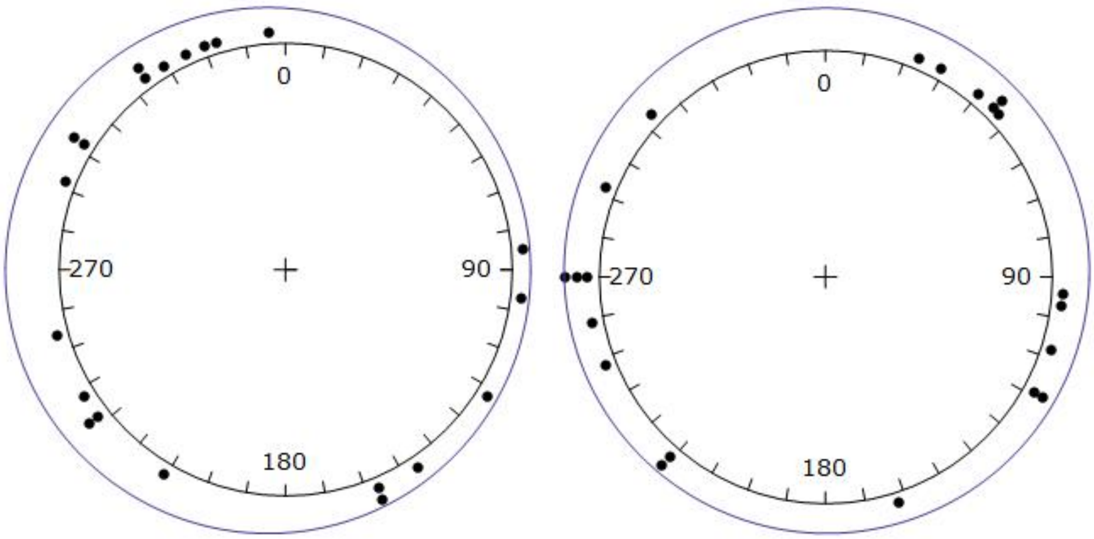
Figure 2.2: Plot of $t + 2\rho^2\rho_\alpha^2$ for $0 < \rho \leq 1/2$ and $0 \leq \rho_\alpha < 1$ when $\mu - \mu_\alpha = 0$.

2.4 Illustrative example

The wind direction at 6 a.m. and 12 noon measured each day at a weather station in Milwaukee for 21 consecutive days (Johnson and Wehrly, 1977) is used for an example.

First, we investigate about testing the hypothesis that the wind direction data $\theta_1, \theta_2, \ldots, \theta_{21}$ at 6 a.m. and $\phi_1, \phi_2, \ldots, \phi_{21}$ at noon come from cardioid distributions. Watson's U^2 goodness-of-fit test gives $U_\theta^2 = 0.071$ and $U_\phi^2 = 0.070$, whose p-values are 0.48 and 0.49 respectively. Thus, we have enough evidence that the data come from cardioid distributions. Figure 2.3 shows circular plots of the data as well as circular plots of densities estimated by the maximum likelihood method under cardioid distributions. The maximum likelihood estimates of parameters are $\hat{\mu}_\theta = 4.670$ (radians) $= 267.608$ (degrees), $\hat{\rho}_\theta = 0.244$, $\hat{\mu}_\phi = 0.214$ (radians) $= 12.273$ (degrees) and $\hat{\rho}_\phi = 0.0853$.

Second, we use a test for serial dependence (Wehrly and Johnson, 1979) for each of the wind direction data at 6 a.m. and noon under a model with cardioid marginal distributions. The test statistic for serial dependence,



(a) (b)

Figure 2.3: (a) Circular plot of the wind data at 6 a.m. and the curve of fitted density of $C(\mu_\theta, \rho_\theta)$. The estimated mean direction and mean resultant length are $\hat{\mu}_\theta = 4.670$ (radians) $= 267.608$ (degrees) and $\hat{\rho}_\theta = 0.244$. (b) Circular plot of the wind data at noon and the curve of fitted density of $C(\mu_\phi, \rho_\phi)$. The estimated mean direction and mean resultant length are $\hat{\mu}_\phi = 0.214$ (radians) $= 12.273$ (degrees) and $\hat{\rho}_\phi = 0.0853$.

$$\sqrt{2/(n-1)}\sum_{j=1}^{n-1}\cos[\theta_{j+1}-\theta_j+2\rho\{\sin(\theta_{j+1}-\mu)-\sin(\theta_j-\mu)\}]$$

for $\theta_1,\ldots,\theta_{n-1}$, has standard normal as an asymptotic distribution under independence. We replace ρ and μ with their maximum likelihood estimates $\hat{\rho}$ and $\hat{\mu}$ under the model of cardioid distribution. The values of the test statistic for the wind direction data at 6.a.m. and noon are 0.181 (p-value 0.428) and 0.864 (p-value 0.194) respectively and there is no strong indication of serial correlation in the sequence of wind direction. As far as we know, testing serial dependence of bivariate circular datapoints is unknown, and this should be a future study.

Third, independence of Θ and Φ is investigated. It is known (cf. Mardia and Jupp (1999, p. 249)) that independence of Θ and Φ is rejected for large values of $r^2_{\theta,\phi}$, since $nr^2_{\theta,\phi}\approx\chi^2_4$ as $n\to\infty$ under independence. Numerical estimation of $r^2_{\theta,\phi}$ gives $r^2_{\theta,\phi}=0.439$, and its p-value is 0.0558, which is marginal to 5% and there is no clear evidence for rejection or acceptance of independence.

Finally, we fit MC_2 to the data set. An interpretation of the MC_2 model is as follows. Let ξ and η denote unobservable angular random variables which represent wind direction at 6 a.m. and noon, and it is assumed that there exists a structural relationship $\eta=M(\xi)$, where ξ follows a cardioid distribution and M is the Möbius transformation. Angular variables δ and ε are independent measurement errors which are assumed to follow cardioid distributions with zero mean direction, independent of ξ, and Θ $(=\xi+\delta)$ and Φ $(=\eta+\varepsilon)$ are observed. The maximum likelihood (ML) estimates of parameters for the model MC_2, the values of maximum log-likelihood (MLL) and the values of Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) are summarized in Table 2.1 for (a) the identity transformation $M(\xi)=\xi$, i.e. $\mathrm{MC}_2(\mu,\rho,-,0,0,\rho_1,\rho_2)$, with common mean resultant length $\rho_1=\rho_2$, (b) angular rotation $M(\xi)=\xi+\mu_\beta$, i.e. $\mathrm{MC}_2(\mu,\rho,-,0,\mu_\beta,\rho_1,\rho_2)$, with $\rho_1=\rho_2$, independent models (c) $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, 0, \rho_2) \equiv \mathrm{MC}_2(\mu \pm \pi, \rho, \mu_\alpha \pm \pi, \rho_\alpha, \mu_\beta \pm \pi, 0, \rho_2)$ and (d) $\mathrm{MC}_2(\mu, \rho, -, -, -, \rho_1, 0)$, (e) a model $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, \rho_1, \rho_2)$ with $\rho_1 = \rho_2$, and (f) the full model $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, \rho_1, \rho_2)$. The potential structure is estimated as $\eta = \hat{M}(\xi)$ using the ML estimates $\hat{\mu}_\alpha$, $\hat{\rho}_\alpha$ and $\hat{\mu}_\beta$. For this data set, a simple distribution $\mathrm{MC}_2(\mu, \rho, -, 0, \mu_\beta, \rho_1, \rho_2)$ with $\rho_1 = \rho_2$ linked to angular rotation $M(\xi) = \xi + \mu_\beta$ is selected as an optimal model between these six in the sense of minimum AIC. Approximate 90% confidence intervals for μ, ρ, μ_β and ρ_1 of the model based on the Fisher information are $(4.921, 6.105)$, $(0.119, 0.425)$, $(0.417, 1.164)$ and $(0.411, 0.5)$. Figure 2.4(a) shows a contour plot of the optimal model with a plot of the data and the Möbius curve, line for this data set, as a function $\phi = \hat{M}(\theta) = \theta + \hat{\mu}_\beta$ with the ML estimate $\hat{\mu}_\beta = 0.790$ (radians) $= 45.264$ (degrees) of μ_β under the model. Figure 2.4(b) shows a 3-D plot of the fitted joint density.

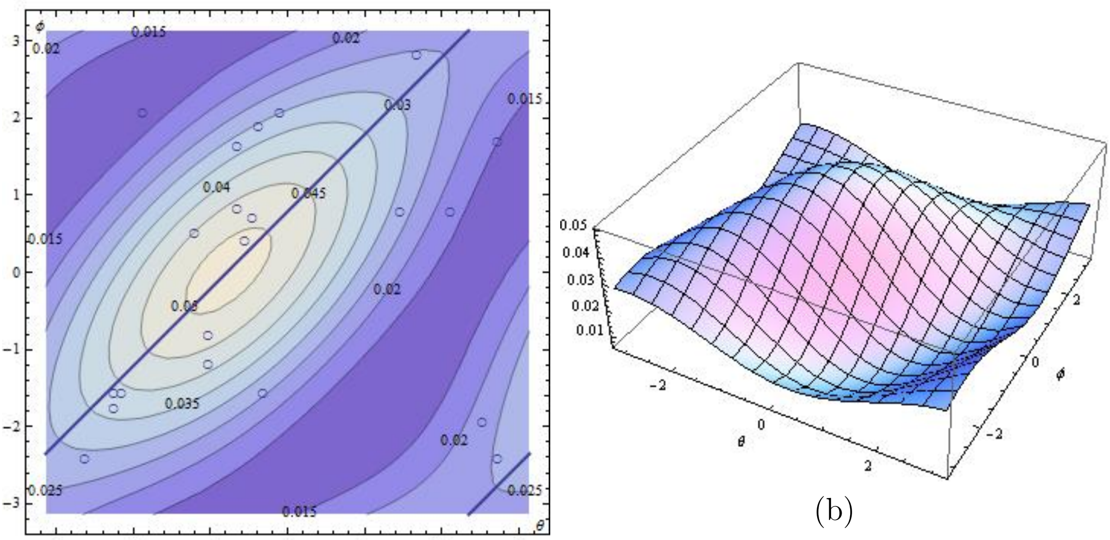
Table 2.1: Maximum likelihood estimates of the parameters, the maximum log-likelihood (MLL), AIC and BIC values for (a) the identity transformation $M(\xi) = \xi$ with $\rho_1 = \rho_2$, (b) angular rotation $M(\xi) = \xi + \mu_\beta$ with $\rho_1 = \rho_2$, independent models (c) $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, 0, \rho_2)$ and (d) $\mathrm{MC}_2(\mu, \rho, -, -, -, \rho_1, 0)$, and (e) a model $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, \rho_1, \rho_2)$ with $\rho_1 = \rho_2$, and (f) the full model $\mathrm{MC}_2(\mu, \rho, \mu_\alpha, \rho_\alpha, \mu_\beta, \rho_1, \rho_2)$.

Model	$\hat{\mu}$	$\hat{\rho}$	$\hat{\mu}_\alpha$	$\hat{\rho}_\alpha$	$\hat{\mu}_\beta$	$\hat{\rho}_1$	$\hat{\rho}_2$	MLL	AIC	BIC
(a)	5.708	0.196	–	–	–	0.5	0.5	-74.588	155.176	158.310
(b)	5.513	0.272	–	–	0.790	0.5	0.5	-73.183	154.367	158.545
(c)	0.010	0.230	0.000	0.803	0.213	–	0.096	-77.076	166.151	172.418
	(3.152)		(3.142)		(3.355)					
(d)	4.671	0.496	–	–	–	0.491	–	-75.884	157.767	160.901
(e)	5.065	0.357	1.266	0.314	0.532	0.5	0.5	-72.705	157.410	163.677
(f)	5.065	0.357	1.266	0.314	0.532	0.5	0.5	-72.705	159.410	166.722

As a comment, the fact that the estimate of ρ_1 $(= \rho_2)$ is 0.5 which is a boundary of the parameter space, suggests better fit of more sharply peaked distributions such as wrapped Cauchy and von Mises distributions. Shieh and Johnson (2005) study a bivariate distribution whose joint probability density function is given by function (2.1). For the same data set, maximum likelihood estimates of the parameters are given in Table 2.2. The bivariate distribution with von Mises marginals is selected as a better model than MC_2(b) in the sense of smaller AIC and BIC values, but MC_2 Model has benefit of inferring the structural relationship between two unobservable circular variables. Circular-circular structural relationship models based on wrapped Cauchy and von Mises distributions are beyond the scope of the current thesis.

Table 2.2: Maximum likelihood estimates of the parameters, the maximum log-likelihood (MLL), AIC and BIC values for the Shieh and Johnson (2005) model.

$\hat{\mu}_1$	$\hat{\kappa}_1$	$\hat{\mu}_2$	$\hat{\kappa}_2$	$\hat{\mu}_{12}$	$\hat{\kappa}_{12}$	MLL	AIC	BIC
4.769	0.645	4.952	0.246	5.137	1.239	-70.115	152.229	158.496



(a) (b)

Figure 2.4: (a) Contour plot of the fitted distribution with plots of the wind direction data at 6 a.m. and noon. The Möbius curve, line for this data set, is drawn as a function $\phi = \hat{M}(\theta) = \theta + \hat{\mu}_\beta$ with the estimated value of μ_β. (b) 3-D plot of the fitted joint density of $\mathrm{MC}_2(\mu, \rho, -, 0, \mu_\beta, \rho_1, \rho_2)$ with $\rho_1 = \rho_2$.

Chapter 3

Distributions on the cylinder

3.1 Introduction

Sometimes it is necessary to consider models for bivariate data, such as wind direction and speed, wind direction and the concentration of a pollutant. The feature is: one variable is angular and the other one is linear. Such a pair of realization may be identified with a point on the cylinder. Thus sometimes angular-linear random variables are called cylindrical random variables and the joint distributions of the cylindrical variables are called cylindrical distributions or distributions on the cylinder.

Some existing parametric models are introduced for cylindrical distributions in Section 3.2. Mardia and Sutton (1978) proposed a suitable model for cylindrical data, and obtained the maximum likelihood estimators for the parameters of the model. Johnson and Wehrly (1978) proposed some angular-linear distributions based on the principle of maximum entropy, and proposed a model with specified marginal distributions. Recently, Kato and Shimizu (2008) investigated the Johnson and Wehrly model in detail, proposed an extension of the Mardia and Sutton model, and provided a theorem which constructs four-dimensional distributions with specified bivariate marginal distributions on certain manifolds such as two tori, cylinders or discs.

In Section 3.3, we propose an angular-linear distribution whose density is a combination of von Mises and transformed Kumaraswamy distributions. Some properties such as marginal distribution, conditional distribution and mode are obtained. Illustrative examples are given in Section 3.4.

Distributional studies and regression models have played important roles in statistical analysis of circular data. Symmetric and possibly asymmetric circular-linear multivariate regression models (SenGupta and Ugwuowo, 2006) are motivated by and applied to predict some environmental characteristics based on both circular and linear predictors. Noting that the circular-linear regression models are derived from the conditional distributions of cylindrical distributions, we give, in Appendix, a likelihood approach (Cook, 1986) to study influence diagnostic analysis for cosine, cosine-sine and cosine-cosine models (see Liu et al., 2013, manuscript).

3.2 Existing distributions on the cylinder

There are various situations which involve both circular random variable Θ and linear random variable X, such as wind direction and speed, direction and movement distance for animal, wind direction and the concentration of air pullutant. Then the random vector (Θ, X) takes values on the cylinder. Mardia and Sutton (1978) proposed a model with density

$$f(\theta, x) = \frac{1}{2\pi I_0(\kappa)} \exp\left\{\kappa \cos(\theta - \mu_0)\right\} \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left\{-\left[\frac{(x-\mu_c)^2}{2\sigma_c^2}\right]\right\}$$

where $-\pi \leq \theta < \pi$, $-\infty < x < \infty$, $-\pi \leq \mu_0 < \pi$, $\kappa \geq 0$, $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero and

$$\begin{aligned}\mu_c &= \mu + \sqrt{\kappa}\sigma\{\rho_1(\cos\theta - \cos\mu_0) + \rho_2(\sin\theta - \sin\mu_0)\},\\ \sigma_c &= \sigma^2(1-\rho^2) \quad \text{and} \quad \rho = (\rho_1^2+\rho_2^2)^{1/2}, \quad 0 \le \rho < 1.\end{aligned}$$

for $-\infty < \mu < \infty$, $\sigma > 0$, $0 \le \rho_1 < 1$, $0 \le \rho_2 < 1$ The parameters of the model are μ, μ_0, κ, ρ_1, ρ_2 and σ. Maximum likelihood estimates of parameters and a practical example are given by Mardia and Sutton (1978).

Johnson and Wehrly (1978) proposed some distributions on the cylinder which maximize the entropy subject to constraints on certain moments. One of the distributions has density

$$f(\theta, x) = \frac{1}{2\pi}\left(\lambda^2 - \kappa^2\right)^{1/2} \exp\{-\lambda x + \kappa x \cos(\theta - \mu)\}, \tag{3.1}$$

where $-\pi \le \theta < \pi$, $x > 0$, $0 < \kappa < \lambda$, $-\pi \le \mu < \pi$. This distribution is the maximum entropy distribution subject to constraints on $\mathrm{E}(X)$, $\mathrm{E}(X\sin\Theta)$, and $\mathrm{E}(X\cos\Theta)$. The marginal distribution of Θ is a wrapped Cauchy distribution $\mathrm{WC}(\mu, \kappa\{\lambda + (\lambda^2-\kappa^2)^{1/2}\}^{-1})$, the conditional distribution of Θ given $X = x$ is a von Mises distribution $\mathrm{VM}(\mu, \kappa x)$ and the conditional distribution of X given $\Theta = \theta$ is an exponential distribution with mean $\{\lambda - \kappa\cos(\theta-\mu)\}^{-1}$.

Another distribution proposed by Johnson and Wehrly (1978) has density

$$f(\theta, x) = c\exp\left\{-\frac{x^2}{2\sigma} + \frac{\lambda x}{\sigma^2} + \frac{\kappa x}{\sigma^2}\cos(\theta-\mu)\right\}, \tag{3.2}$$

where c (> 0) is a normalizing constant, $-\pi \le \theta < \pi$, $-\infty < x < \infty$, $-\infty < \lambda < \infty$, $\kappa > 0$, and $-\pi < \mu < \pi$, gives the maximum entropy cylindrical distribution subject to $\mathrm{E}(X)$, $\mathrm{E}(X^2)$, $\mathrm{E}(X\cos\Theta)$ and $\mathrm{E}(X\sin\Theta)$ taking specified values (see also SenGupta, 2004). The marginal distributions are not of familiar form, but the conditional distributions are

$$f_1(\theta|x) = \left\{2\pi I_0(\kappa x/\sigma^2)\right\}^{-1} \exp\left\{(\kappa x/\sigma^2)\cos(\theta-\mu)\right\}$$

and

$$f_2(x|\theta) = (2\pi\sigma^2)^{-1/2} \exp\left\{-(1/2\sigma^2)[x - \{\lambda + \kappa\cos(\theta-\mu)\}]^2\right\},$$

which are the densities of a von Mises distribution $\mathrm{VM}(\mu, \kappa x/\sigma^2)$ and a normal distribution $\mathrm{N}(\lambda + \kappa\{\cos(\theta-\mu)\}, \sigma^2)$ respectively.

Furthermore, Johnson and Wehrly (1978) proposed a distribution, whose marginal densities $f_1(\theta)$, $f_2(x)$ are specified, with density

$$f(\theta, x) = 2\pi f_1(\theta) f_2(x) g\left[2\pi\left\{F_1(\theta) - F_2(x)\right\}\right],$$

where $F_1(\theta)$ and $F_2(x)$ denote the distribution functions of $f_1(\theta)$ and $f_2(x)$ respectively, and $g(\cdot)$ is a probability density function on the circle.

Kato and Shimizu (2008) gave a theorem which constructs four-dimensional distributions with specified bivariate marginals on certain manifolds such as two tori, cylinders or discs, and also proposed an extention of the Mardia and Sutton model, which has the joint probability density function:

$$f(\theta, x) = C^{-1} \exp\left[-\frac{\{x - \mu(\theta)\}^2}{2\sigma^2} + \kappa_1\cos(\theta-\mu_1) + \kappa_2\cos 2(\theta-\mu_2)\right], \quad (3.3)$$

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

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

The distribution has the following properties.

(1) $f(\theta, x)$ is the maximum entropy probability density function on the cylinder subject to $\mathrm{E}(X^2)$, $\mathrm{E}(X)$, $\mathrm{E}(X\cos\Theta)$, $\mathrm{E}(X\sin\Theta)$, $\mathrm{E}(\cos\Theta)$, $\mathrm{E}(\sin\Theta)$, $\mathrm{E}(\cos 2\Theta)$ and $\mathrm{E}(\sin 2\Theta)$ taking specified values consistent with expectation.

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

Note: Kato (2009) proposed a cylindrical distribution with circular uniform and standard Cauchy marginals.

3.3 New distributions on the cylinder

The angular-linear distribution we propose has density

$$f(\theta, x) = A^{-1}\, \mathrm{e}^{\{-\lambda x + \kappa x \cos(\theta - \mu)\}} \left(1 - \mathrm{e}^{-\lambda x}\right)^{\beta - 1}, \tag{3.4}$$

where $-\pi \leq \theta < \pi,\ x \geq 0;\ -\pi \leq \mu < \pi,\ \beta > 0,\ \lambda > \kappa > 0$, and the normalizing constant is

$$A = 2\pi \sum_{j=0}^{\infty} \binom{\beta - 1}{j} (-1)^j \{\lambda^2 (j+1)^2 - \kappa^2\}^{-1/2}.$$

This distribution is a combination of von Mises and transformed Kumaraswamy distributions, and it is also an extended model of the Johnson–Wehrly model (3.1). When $\beta = 1$, the normalizing constant is $A = 2\pi(\lambda^2 - \kappa^2)^{-1/2}$, and our model is reduced to Johnson–Wehrly model (3.1). When parameters satisfy the condition $1 < \beta < 1 + (\lambda - \kappa)/\lambda$, the distribution with (3.4) has a mode at (μ, $\lambda^{-1} \ln\{(\lambda\beta - \kappa)/(\lambda - \kappa)\}$) and an antimode at ($\mu + \pi, \lambda^{-1} \ln\{(\lambda\beta + \kappa)/(\lambda + \kappa)\}$).

The marginal density of Θ can be expressed as

$$f_\Theta(\theta) \;=\; A^{-1}\sum_{j=0}^{\infty}\binom{\beta-1}{j}(-1)^j\{\lambda(j+1)-\kappa\cos(\theta-\mu)\}^{-1}, \qquad (3.5)$$

where $-\pi \le \theta < \pi$. The distribution with density (3.5) is symmetric about μ and unimodal. It has mode at μ and antimode at $\mu+\pi$. As a special case, when $\beta = 1$, the density function (3.5) is reduced to

$$f_\Theta(\theta) \;=\; \frac{(\lambda^2-\kappa^2)^{1/2}}{2\pi}\{\lambda-\kappa\cos(\theta-\mu)\}^{-1}, \quad -\pi \le \theta < \pi.$$

It is a wrapped Cauchy distribution with mean direction μ and mean resultant length $\kappa/\{\lambda+(\lambda^2-\kappa^2)^{1/2}\}$.

The marginal distribution of X is given by

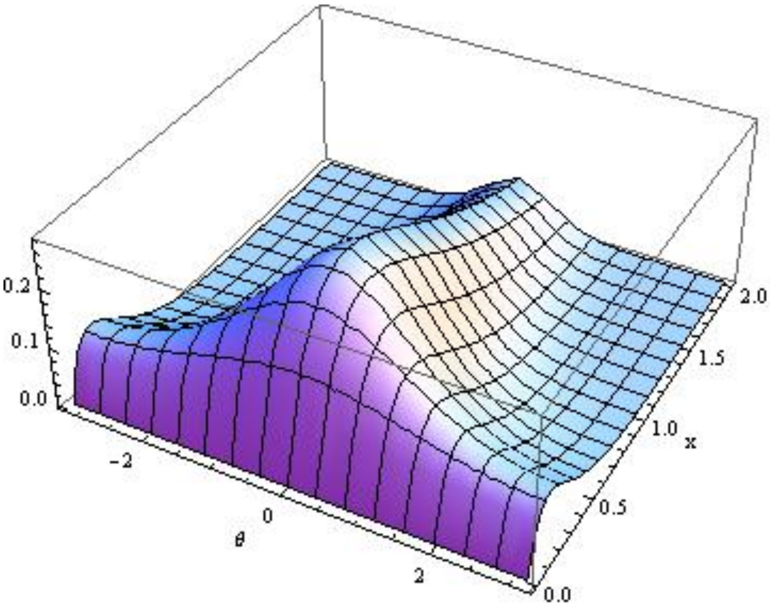
$$f_X(x) \;=\; 2\pi A^{-1}\mathrm{e}^{-\lambda x}\left(1-\mathrm{e}^{-\lambda x}\right)^{\beta-1} I_0(\kappa x), \quad x > 0. \qquad (3.6)$$

Thus, the conditional distribution of Θ given $X = x$ is a von Mises distribution $\mathrm{VM}(\mu,\, \kappa x)$. As a special case, when $\beta = 1$, the density function (3.6) is reduced to

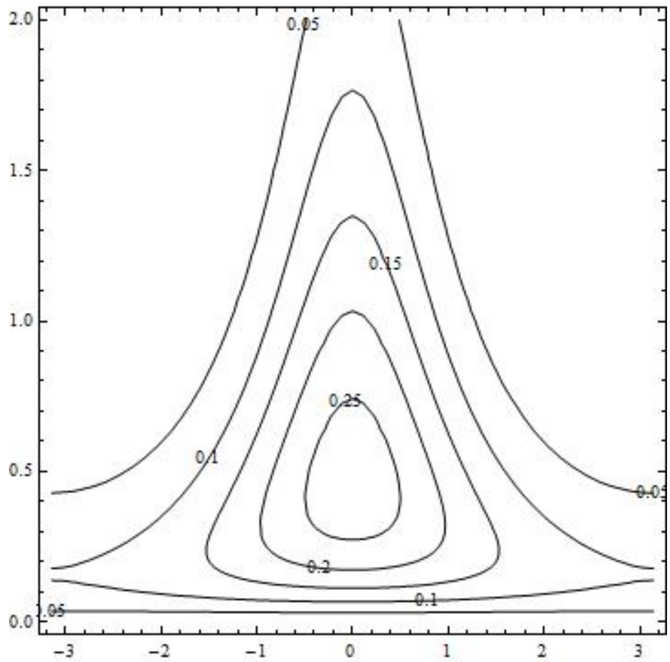
$$f_X(x) \;=\; (\lambda^2-\kappa^2)^{1/2}\mathrm{e}^{-\lambda x}I_0(\kappa x), \quad x > 0.$$

Since $f'_X(x) = \kappa^2\mathrm{e}^{-\lambda x}[1-(\lambda/\kappa)^2-(\lambda/\kappa)\{A(\kappa x)\}^{-1}] < 0$, $f_X(x)$ is a monotonically decreasing function, and takes maximum value $(\lambda^2-\kappa^2)^{1/2}$ at point $x = 0$. Figures 3.1 and 3.2 show the plots of model (3.4) with parameters $\beta = 2$ and $\beta = 0.8$ for $\lambda = 3$, $\kappa = 2$, $\mu = 0$ respectively.

We also consider some other cylindrical distributions. McClintock et al. (2012) proposed a model which has the form of a combination of Weibull and wrapped Cauchy distributions. Different from the model used in McClintock et

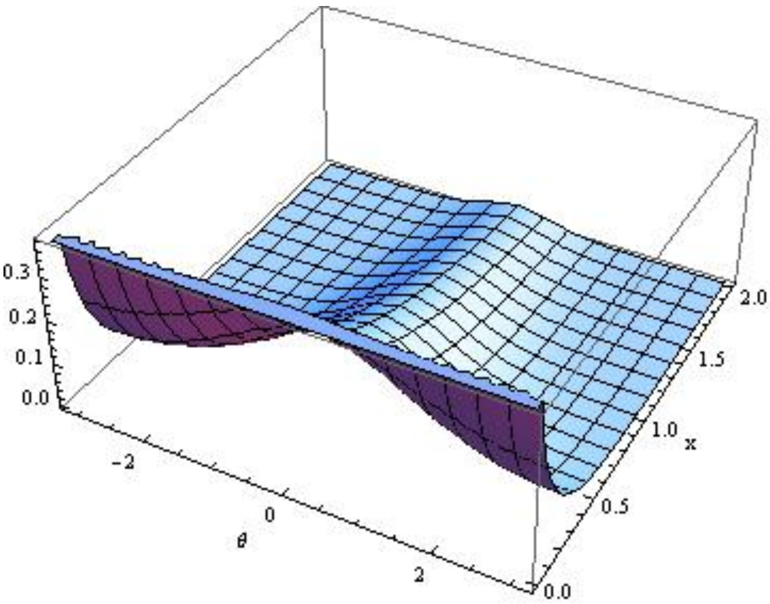


(a)

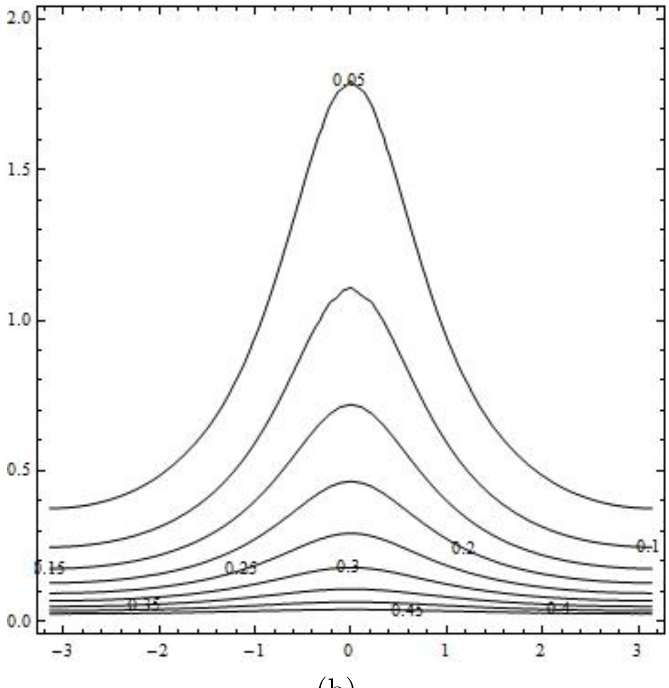


(b)

Figure 3.1: (a) 3-D plot and (b) contour plot of model (3.4) with parameters: $\lambda = 3$, $\beta = 2$, $\kappa = 2$, $\mu = 0$.



(a)



(b)

Figure 3.2: (a) 3-D plot and (b) contour plot of model (3.4) with parameters: $\lambda = 3$, $\beta = 0.8$, $\kappa = 2$, $\mu = 0$.

al. (2012), we let θ and x be dependent, and think of a cylindrical distribution whose density is given by

$$
\begin{aligned}
f(\theta, x) &= \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} \mathrm{e}^{-\left(\frac{x}{a}\right)^{b}} \\
&\quad \times \left(\frac{1}{2\pi}\right) \frac{1-\{\tanh(\rho x)\}^2}{1+\{\tanh(\rho x)\}^2 - 2\tanh(\rho x)\cos(\theta-\mu)}, \qquad (3.7)
\end{aligned}
$$

where $-\pi \leq \theta < \pi$, $x \geq 0$; $-\pi \leq \mu < \pi$, $a > 0$, $b > 0$, $\rho > 0$. Figure 3.3 shows 3-D and contour plots of the density (3.7).

We also propose a joint distribution by using the method of generating a cylindrical distribution with specified marginal distributions given by Johnson and Wehrly (1978). The joint density function is given by

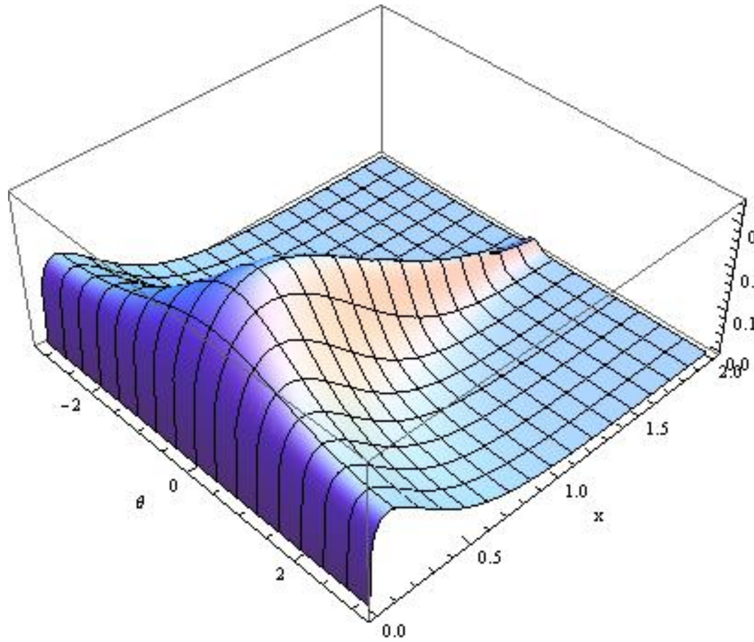
$$
\begin{aligned}
f(\theta, x) &= \frac{1}{2\pi}\,\alpha\beta\, e^{-\alpha x}\left(1-e^{-\alpha x}\right)^{\beta-1}\{1+2\rho\cos(\theta-\mu_c)\} \\
&\quad \times \frac{1}{I_0(\kappa)} \mathrm{e}^{\kappa\cos\left[2\pi\left(1-e^{-\alpha x}\right)^{\beta} - \{\theta+2\rho\sin(\theta-\mu_c)+2\rho\sin\mu_c\}-\mu\right]}, \qquad (3.8)
\end{aligned}
$$

where $-\pi \leq \theta < \pi$, $x \geq 0$; $-\pi \leq \mu < \pi$, $-\pi \leq \mu_c < \pi$ $\kappa > 0$, $\alpha \geq 0$, $\beta > 0$, $0 \leq \rho < 1/2$. This distribution has a cardioid marginal distribution $\mathrm{C}(\mu_c, \rho)$ for circular random variable Θ, and the linear random variable X is distributed as Kumaraswamy's distribution after the exponential transformation $\exp(-X)$. The circular link function of (3.8) is the von Mises distribution $\mathrm{VM}(\mu, \kappa)$.

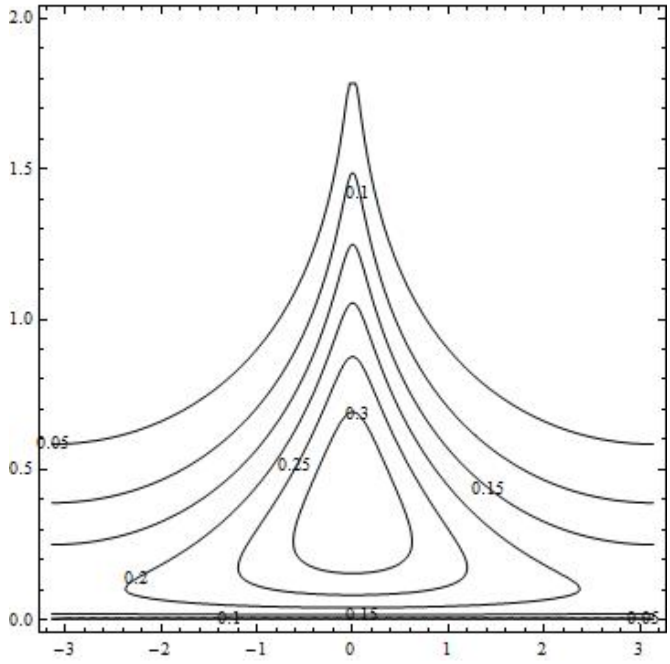
3.4 Illustrative examples

Example 1

In this example, we use the latitude, longitude and magnitude data for foreshocks during 72 hours before the earthquake off the Pacific coast of Tohoku, Japan, a magnitude 9.0 $\mathrm{M_W}$, occurred at 14 : 46 JST (05 : 46 UTC) on 11 March, 2011 with the epicenter 38.30 degrees for latitude and 142.37 degrees for



(a)



(b)

Figure 3.3: (a) 3-D plot and (b) contour plot of model (3.7) with parameters $a = 0.5$, $b = 1.3$, $\rho = 0.8$, $\mu = 0$.

longitude approximately 70 km east of the coast. The data were taken from the U.S. Geological Survey website at

http://earthquake.usgs.gov/earthquakes/eqarchives/epic/epic_global.php

on 8 July 2011. We pick up the first 28 earthquakes off the Pacific coast of Tohoku whose magnitudes, m, are greater than 4. Figures 3.4 and 3.5 show 3-D and contour plots for the set of the magnitudes $m_j - 4$ ($j = 1, \ldots, 28$) and corresponding consecutive angles θ_j ($0 \leq \theta_j < 2\pi$; $j = 1, \ldots, 28$) calculated from epicenters. From the AIC values 146.742 for model (3.1) and 117.886 for model (3.8), we select model (3.8) as an optimal model between the two.

Example 2

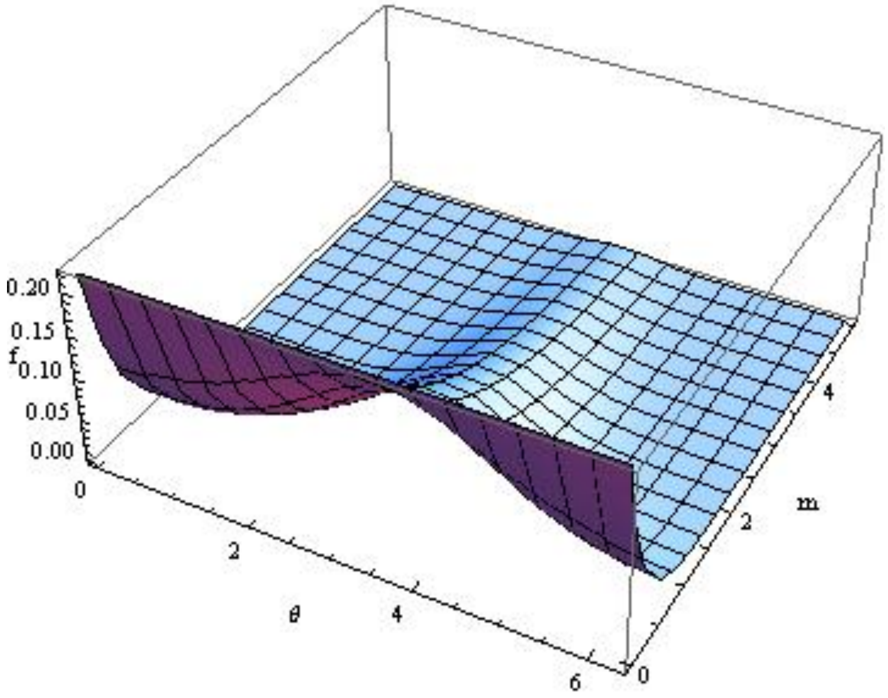
The cylindrical dataset on movements of blue periwinkles is considered next. The observations are directions (θ) and distances (x) moved by small blue periwinkles after they had been transplanted downshore from the height at which they normally live. The data are taken from Table B.20 of Fisher (1993). The planar plot of the cylindrical data (θ, x) is shown on Figure 3.6.

For the directions, the sample mean direction and mean resultant length are 1.620 radians (92.819 degrees) and 0.775 respectively. Figure 3.7 shows the circular plot of the directions. We test symmetry (Pewsey, 2005) for the directions. The p-value of the test is 0.006, and thus the directions can be seen as coming from an asymmetric circular distribution.

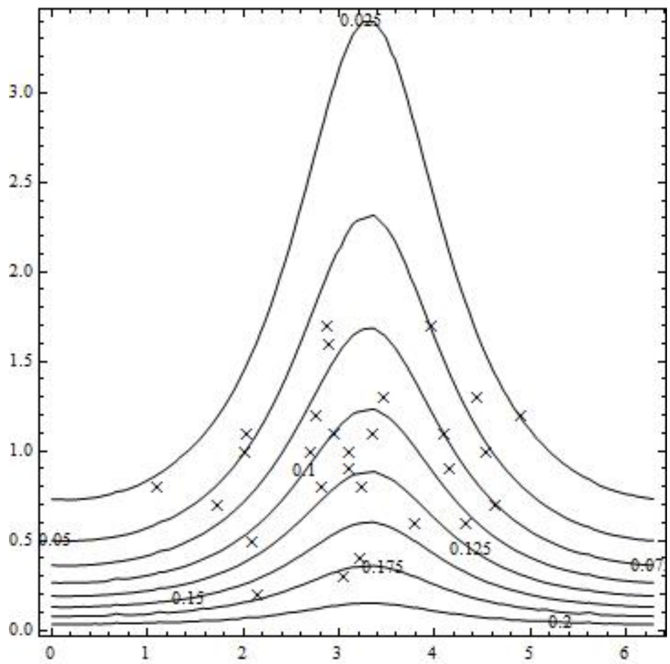
We use the sine skewed model of (3.4), whose density is given by

$$\begin{aligned} f_{ssk1}(\theta, x) &= \{1 + \gamma \sin(\theta - \mu)\} \\ &\quad \times A\, \mathrm{e}^{\{-\lambda x + \kappa x \cos(\theta - \mu)\}} \left(1 - \mathrm{e}^{-\lambda x}\right)^{\beta - 1}, \end{aligned} \tag{3.9}$$

where $-1 \leq \gamma \leq 1$, and the sine skewed model of (3.7) with density

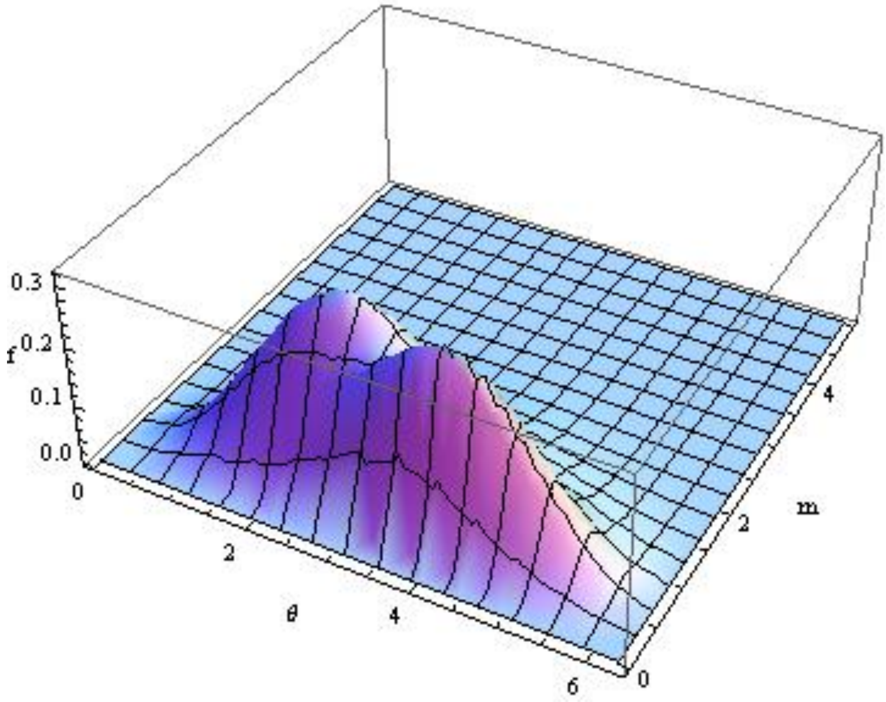


(a)

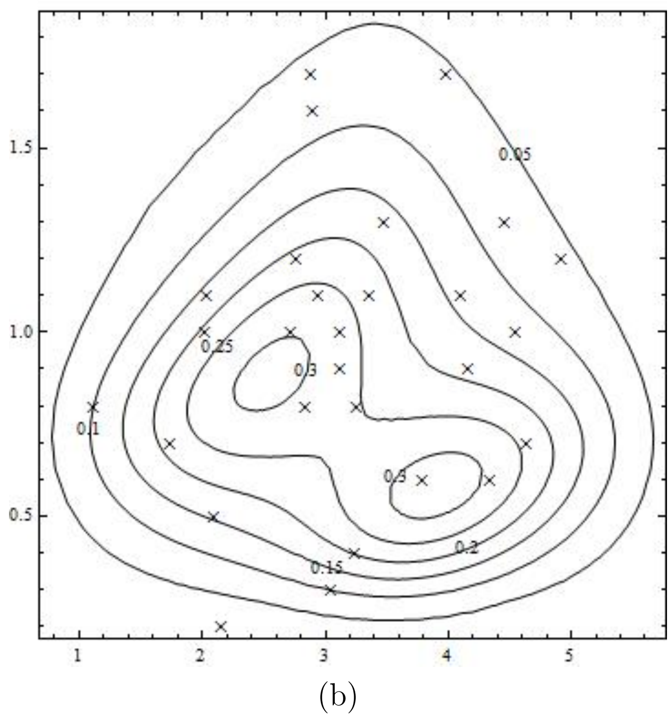


(b)

Figure 3.4: (a) 3-D plot and (b) contour plot of fitted model (3.1) with scatter plot: $\hat{\lambda} = 1.818$, $\hat{\kappa} = 1.180$, $\hat{\mu} = 3.308$, MLL= -70.371, AIC= 146.742, BIC= 150.738.



(a)



(b)

Figure 3.5: (a) 3-D plot and (b) contour plot of fitted model (3.8) with scatter plot: $\hat{\alpha} = 2.648$, $\hat{\beta} = 6.345$, $\hat{\rho} = 0.500$, $\hat{\mu}_c = 3.221$, $\hat{\kappa} = 0.259$, $\hat{\mu} = 2.272$, MLL$= -52.943$, AIC$= 117.886$, BIC$= 125.879$.

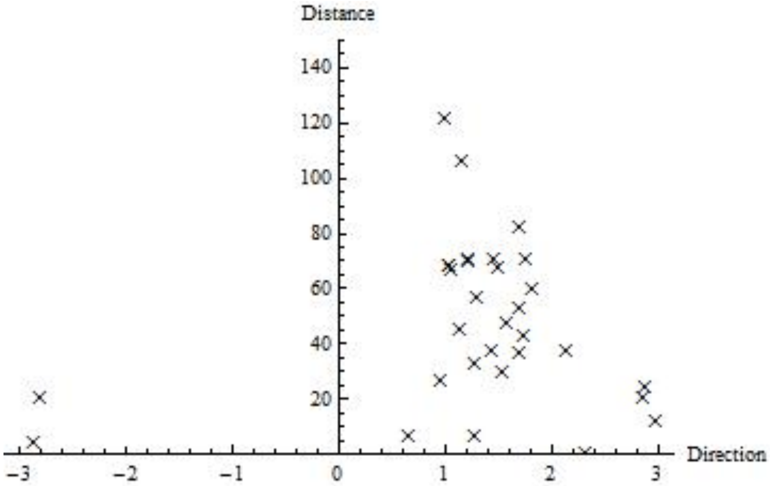


Figure 3.6: Plot of observations of directions (θ) and distances (x) moved by small blue periwinkles.

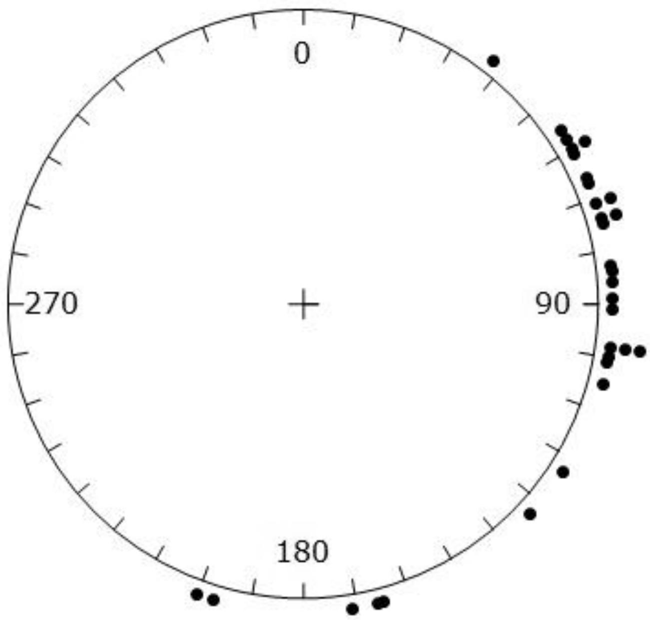


Figure 3.7: Circular plot of the directions.

$$
\begin{aligned}
f_{ssk2}(\theta, x) &= \{1 + \gamma \sin(\theta - \mu)\} \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} \mathrm{e}^{-\left(\frac{x}{a}\right)^b} \\
&\quad \times \left(\frac{1}{2\pi}\right) \frac{1 - \{\tanh(\rho x)\}^2}{1 + \{\tanh(\rho x)\}^2 - 2\tanh(\rho x)\cos(\theta - \mu)}, \quad (3.10)
\end{aligned}
$$

where $-1 \leq \gamma \leq 1$. This idea of skewing comes from Abe and Pewsey (2011), who proposed skew circular distributions that are generated by perturbation of symmetric circular distributions. We fit models (3.9) and (3.10) to the cylindrical data set of movements of blue periwinkles, and compare the results with those using the Kato–Shimizu and Mardia–Sutton models.

Table 3.1: Maximum likelihood estimates of the parameters, the maximum log-likelihood, and AIC of model (3.9).

$\hat{\lambda}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\mu}$	$\hat{\gamma}$	MLL	AIC
0.095	2.334	0.078	1.227	1.	-174.694	359.387

Table 3.2: Maximum likelihood estimates of the parameters, the maximum log-likelihood, and AIC of model (3.10).

$\hat{a}$	$\hat{b}$	$\hat{\rho}$	$\hat{\mu}$	$\hat{\gamma}$	MLL	AIC
52.368	1.536	0.017	1.165	1.	-171.119	352.237

The maximum likelihood estimates of the parameters, the maximum log-likelihood, and AIC of models (3.9) and (3.10) are given in Table 3.1 and Table 3.2 respectively. Figures 3.8 and 3.9 show the contour plots of fitted models as well as the scatter plots. The AIC value of model (3.10) is 352.2, which is less than the AIC values 359.387 of model (3.9), 357.5 of Kato–Shimizu model, and 365.8 of Mardia–Sutton model. Thus, model (3.10) is selected as an optimal model among these four in the sense of minimum AIC.

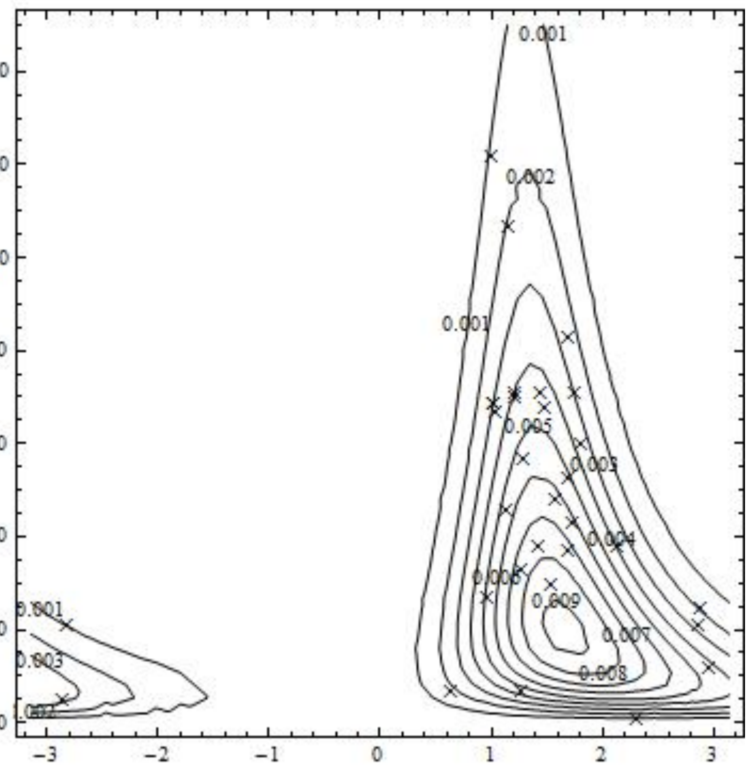


Figure 3.8: Contour plot of fitted model (3.9) with scatter plot for the data set of movements of blue periwinkles.

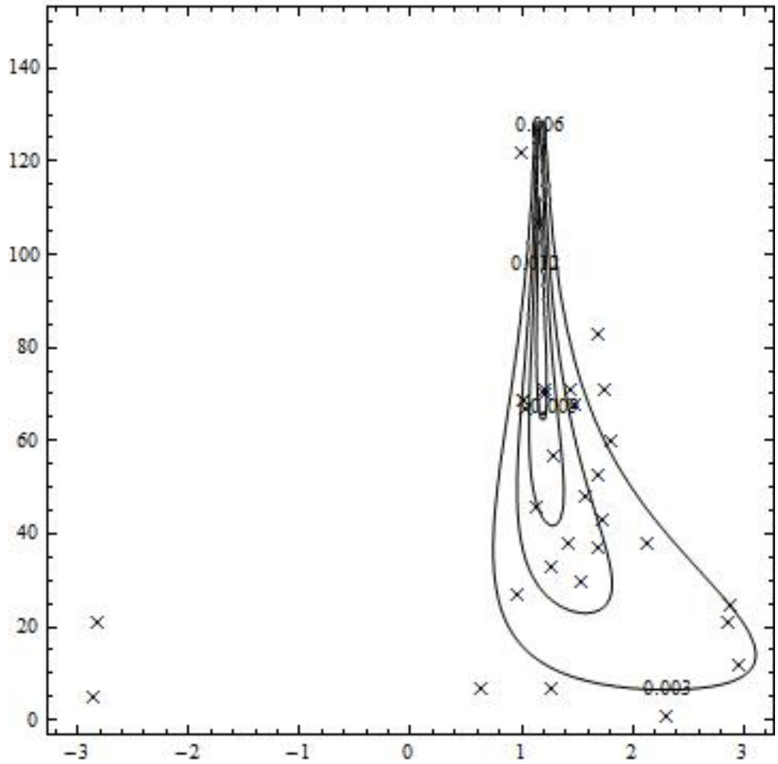


Figure 3.9: Contour plot of fitted model (3.10) with scatter plot for the data set of movements of blue periwinkles.

3.5 Circular-linear regression models

In this section, we remark that circular-linear regression models are derived from distributions on the cylinder.

If a random vector (Θ, Y) is distributed according to the Johnson–Wehrly distribution (3.2), the conditional distribution of Y given $\Theta = \theta$ is normal with mean $\lambda + \kappa \cos(\theta - \mu)$ and variance σ^2. This provides the usual model for trigonometric regression. Similarly if we start with a modified joint density

$$f(\theta, y) \propto \exp\left[-\frac{y^2}{2\sigma^2} + \frac{by}{\sigma^2} + \frac{\kappa y}{\sigma^2} \cos\{\theta - \phi + \gamma \sin(\theta - \phi)\}\right], \quad |\gamma| \leq 1,$$

then the conditional distribution of Y given $\Theta = \theta$ is normal with mean $b + \kappa \cos\{\theta - \phi + \gamma \sin(\theta - \phi)\}$ and variance σ^2. If we adopt an alternative joint density

$$f(\theta, y) \propto \exp\left[-\frac{y^2}{2\sigma^2} + \frac{by}{\sigma^2} + \frac{\kappa y}{\sigma^2} \cos\{\theta - \phi + \gamma \cos(\theta - \phi)\}\right], \quad |\gamma| \leq 1,$$

then the conditional distribution of Y given $\Theta = \theta$ is normal with mean $b + \kappa \cos\{\theta - \phi + \gamma \cos(\theta - \phi)\}$ and variance σ^2.

Thus, SenGupta and Ugwuowo (2006) propose the following Cosine, Cosine-sine and Cosine-cosine (SU) regression models:

$$\begin{aligned} y_i &= x_i'\beta + \alpha \cos(\theta_i - \phi) + \varepsilon_i, && (3.11) \\ y_i &= x_i'\beta + \alpha \cos\{\theta_i - \phi + \gamma \sin(\theta_i - \phi)\} + \varepsilon_i, && (3.12) \\ y_i &= x_i'\beta + \alpha \cos\{\theta_i - \phi + \gamma \cos(\theta_i - \phi)\} + \varepsilon_i, && (3.13) \end{aligned}$$

where y_i $(-\infty < y_i < \infty)$ is the observed linear response value, $x_i = (1, x_{i1}, x_{i2})'$ $(\in \mathbb{R}^3)$ is the observed linear covariate value, θ_i $(-\pi \leq \theta_i < \pi)$ is the observed angular or circular value, ε_i is independently distributed and $\varepsilon_i \sim N(0, \sigma^2)$ for

$i = 1, \ldots, n$. The parameter vectors in (3.11), (3.12) and (3.13) are respectively $\eta = (\beta', \alpha, \phi, \sigma^2)'$, $\eta = (\beta', \alpha, \phi, \gamma, \sigma^2)'$ and $\eta = (\beta', \alpha, \phi, \gamma, \sigma^2)'$, where σ^2 (> 0) is the variance assumed for the normal distribution, $\beta = (\beta_0, \beta_1, \beta_2)'$ $(\in \mathbb{R}^3)$ is the regression coefficient vector, $\alpha \in \mathbb{R}$, $-\pi \leq \phi < \pi$, and particularly γ $(|\gamma| \leq 1)$ in (3.12) is the parameter of kurtosis (indicating to what extent the shape differs from a sinusoidal oscillation) and γ in (3.13) is the parameter of skewness (for the oscillation pattern when the peaks and troughs do not follow each other).

An application of the local influence method proposed by Cook (1986) to the SU regression models will be treated in Appendix.

Chapter 4

Distributions on the disc

4.1 Introduction

Some distributions with support on the unit disc in two dimensions are introduced in Section 4.2. As far as we know, only few articles investigate distributions on the disc. Jones (2002) proposed a beta distribution on the disc. An alternative distribution on the disc named the Möbius distribution was proposed by Jones (2004), whose density function is asymmetric or skew for the length from the center of the disc, but symmetric for a line through the origin. Therefore, in Section 4.3.2 we propose a family of modified Möbius distributions skew for both of the length and angle. The modified Möbius distributions have six parameters. They can be reduced to the Möbius distribution and uniform distribution as special cases, but in general the family is skew for both of the length and angle. Some properties such as marginal distribution are obtained. In Section 4.4, an illustrative example of fitting the models to data which consist of values of ozone concentration and wind direction (Johnson and Wehrly, 1977) is given.

4.2 Möbius distribution

The Möbius distribution was proposed by Jones (2004). The density function is asymmetric for the length from the center of the disc, and it is obtained by applying the Möbius transformation to the bivariate spherically symmetric beta (or Pearson type II) distribution with density

$$g_\beta(x,y) = \frac{\beta}{\pi}\left(1-x^2-y^2\right)^{\beta-1}, \quad \beta>0, \quad 0\le x^2+y^2<1.$$

The distribution includes the uniform distribution on the disc when $\beta = 1$ and is symmetric with mode at zero when $\beta > 1$ and antimode at zero when $\beta < 1$, whose density function could be written as

$$g(\rho,\varphi) = \frac{\beta}{\pi}\rho(1-\rho^2)^{\beta-1}, \quad 0<\rho<1,\ -\pi\le\varphi<\pi,\ \beta>0$$

in polar coordinates. Write $z = (x, y)$ and $w = (u, v)$ as complex numbers with, in particular, $w = r\mathrm{e}^{\mathrm{i}\theta}, 0 \le r \le 1, -\pi \le \theta \le \pi$. Applying the inverse Möbius transformation $w \equiv r\,\mathrm{e}^{\mathrm{i}\theta} = M_{\mathrm{D}}(z) = (z + c)/(1 + \bar{c}z)$ $(c = a\,\mathrm{e}^{\mathrm{i}\mu},\ 0 \le a < 1,\ 0 \le \mu < 2\pi)$ to z, then we obtain the Möbius distribution. The density function for (r, θ) can be written as

$$f(r,\theta) = \frac{\beta(1-a^2)^{\beta+1}r(1-r^2)^{\beta-1}}{\pi\left\{1+a^2r^2-2ar\cos(\theta-\mu)\right\}^{\beta+1}}, \quad 0<r<1,\ -\pi\le\theta<\pi.$$

The parameter a controls the off-centeredness of the distribution, and the Möbius distribution reduces to the bivariate spherically symmetric beta distribution when $a = 0$. The parameter β plays the role of concentration parameter. The Möbius distribution is asymmetric for the length r, but still symmetric for the angle θ. Figure 4.1 shows (a) a 3-D plot and (b) a contour plot for the Möbius distribution with parameters $a = 0.25, \beta = 3$ and $\mu = 0$. The marginal distribution of R is

$$f(r) = \frac{2\beta(1-a^2)^{\beta+1} r(1-r^2)^{\beta-1}}{(1-a^2r^2)^{\beta+1}} P_\beta\left(\frac{1+a^2r^2}{1-a^2r^2}\right) \quad 0 \le r < 1,$$

where P denotes the associated Legendre function and is defined as

$$P_\beta(z) = \frac{1}{\pi}\int_0^\pi \frac{d\psi}{\left(z+\sqrt{z^2-1}\cos\psi\right)^{\beta+1}}.$$

The conditional distribution of $\Theta|(R=r)$ is

$$f(\theta|r) = \left[2\pi P_\beta(z)\left\{z - \sqrt{z^2-1}\cos(\theta-\mu)\right\}^{\beta+1}\right]^{-1},$$

where $z = (1+a^2r^2)/(1-a^2r^2)$. It belongs to the family of Jones and Pewsey (2005), actually a circular t distribution given in Section 1.3.5.

4.3 A modified Möbius distribution

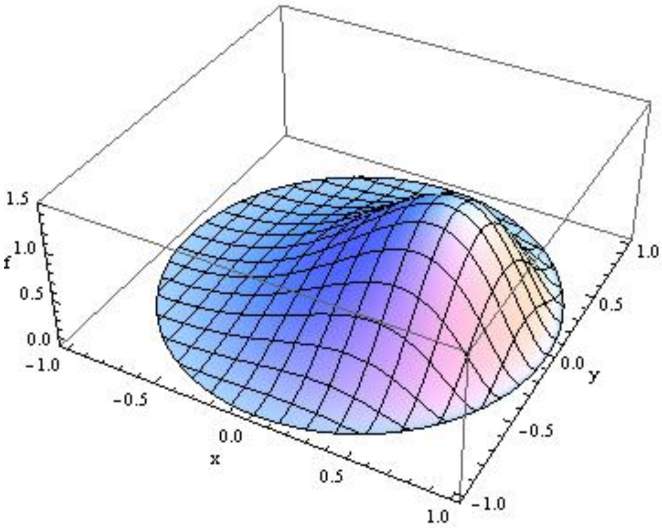
4.3.1 Kumaraswamy's distribution

Jones (2009) systematically investigated Kumaraswamy's (1980) distribution. Its density is

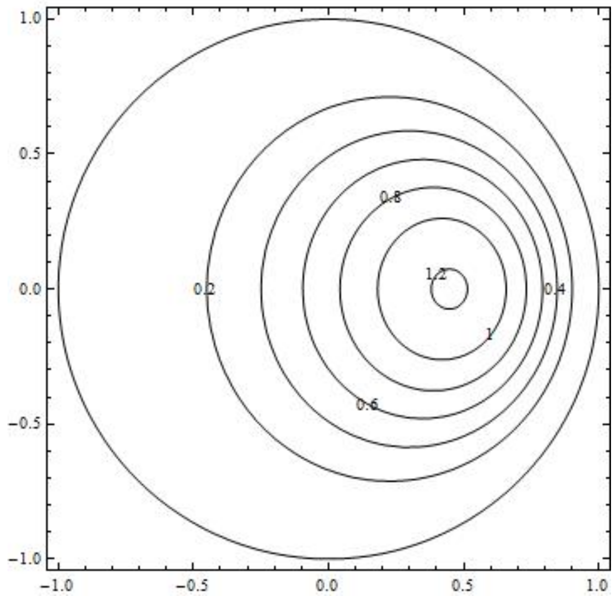
$$g(\rho) = \alpha\beta\rho^{\alpha-1}\left(1-\rho^\alpha\right)^{\beta-1}, \quad 0 < \rho < 1, \tag{4.1}$$

where α and β are positive shape parameters.

Kumaraswamy's densities are unimodal, uniantimodal, increasing, decreasing or constant depending in the same way as the beta distribution on the values of its parameters. Some density plots of Kumaraswamy's distribution are given in Figure 4.2. The distribution function of Kumaraswamy's distribution is



(a)



(b)

Figure 4.1: (a) 3-D plot and (b) contour plot of the Möbius distribution with parameters $a = 0.25$, $\beta = 3$ and $\mu = 0$.

$$G(\rho) = 1 - (1 - \rho^{\alpha})^{\beta}, \quad 0 < \rho < 1,$$

and it can be shown that Kumaraswamy's distribution has the shape:

(1) unimodal, if $\alpha > 1$, $\beta > 1$,

(2) uniantimodal, if $0 < \alpha < 1$, $0 < \beta < 1$,

(3) increasing, if $\alpha > 1$, $0 < \beta \leq 1$,

(4) decreasing, if $0 < \alpha \leq 1$, $\beta > 1$,

(5) constant, if $\alpha = \beta = 1$.

In the first two cases, the mode/antimode is at

$$\rho = \left(\frac{\alpha - 1}{\alpha\beta - 1} \right)^{1/\alpha}.$$

At the boundaries $\rho \to 0$ and $\rho \to 1$, Kumaraswamy's distribution has density:

(1) $g(\rho) \sim \rho^{\alpha-1}$, as $\rho \to 0$,

(2) $g(\rho) \sim (1 - \rho)^{\beta-1}$, as $\rho \to 1$.

4.3.2 A modified Möbius distribution

We propose a distribution on the unit disc, which is asymmetric not only for the length r but also for the direction θ. Similar to Jones (2004), we start with a bivariate independent circular distribution with joint density function

$$g(\rho, \varphi) = \frac{\alpha\beta}{2\pi} \rho^{\alpha-1} (1 - \rho^{\alpha})^{\beta-1}, \tag{4.2}$$

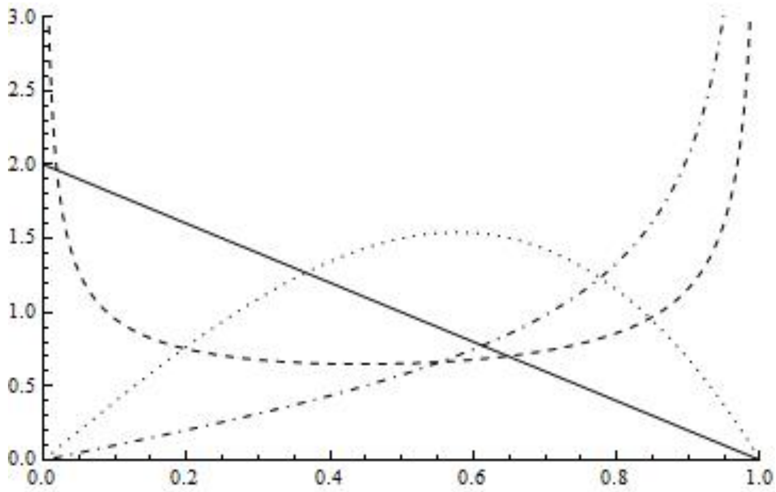


Figure 4.2: Density plots of Kumaraswamy's distribution with parameters $\alpha = 2$, $\beta = 2$ (dots), $\alpha = 0.5$, $\beta = 0.5$ (dashed), $\alpha = 2$, $\beta = 0.5$ (dotdashed), $\alpha = 1$, $\beta = 2$ (solid).

where $0 < \rho < 1$, $-\pi \leq \varphi < \pi$, $\alpha \geq 2$ and $\beta > 0$. This distribution is an independent model for the length ρ and angle φ which are made from Kumaraswamy's distribution and a circular uniform distribution. It reduces to the bivariate spherically symmetric beta distribution when $\alpha = 2$. By applying the inverse Möbius transformation to the distribution (4.2), we have equations

$$
\begin{aligned}
z &= \frac{\tau - c}{1 - \bar{c}\tau}, \\
\rho &= \sqrt{\frac{a^2 + r_\tau^2 - 2ar_\tau \cos(\theta_\tau - \mu)}{1 + a^2 r_\tau^2 - 2ar_\tau \cos(\theta_\tau - \mu)}}, \\
\varphi &= \arctan\left\{\frac{r_\tau \sin\theta_\tau - a^2 r_\tau \sin(\theta_\tau - 2\mu) - a\,(1 + r_\tau^2)\sin\mu}{r_\tau \cos\theta_\tau + a^2 r_\tau \cos(\theta_\tau - 2\mu) - a\,(1 + r_\tau^2)\cos\mu}\right\},
\end{aligned}
$$

for $z = \rho e^{i\varphi}$ and $\tau = r_\tau e^{i\theta_\tau}$. The Jacobian for this transformation is

$$
J_1 = \left|\left|\frac{\partial(\rho, \varphi)}{\partial(r_\tau, \theta_\tau)}\right|\right| = \frac{r_\tau\,(1 - a^2)^2}{\sqrt{\frac{a^2 + r_\tau^2 - 2ar_\tau \cos(\theta_\tau - \mu)}{1 + a^2 r_\tau^2 - 2ar_\tau \cos(\theta_\tau - \mu)}}\,\left(1 + a^2 r_\tau^2 - 2ar_\tau \cos(\theta_\tau - \mu)\right)^2}.
$$

After this transformation, we obtain the joint density function of R_τ and Θ_τ

$$f(r_\tau,\theta_\tau)=\frac{\alpha\beta(1-a^2)^2 r_\tau}{2\pi\{B(r_\tau,\theta_\tau)\}^2}\left\{\frac{A(r_\tau,\theta_\tau)}{B(r_\tau,\theta_\tau)}\right\}^{\alpha/2-1}\left[1-\left\{\frac{A(r_\tau,\theta_\tau)}{B(r_\tau,\theta_\tau)}\right\}^{\alpha/2}\right]^{\beta-1}, \quad (4.3)$$

where

$$\begin{aligned}A(r_\tau,\theta_\tau) &= a^2+r_\tau^2-2ar_\tau\cos(\theta_\tau-\mu),\\ B(r_\tau,\theta_\tau) &= 1+a^2r_\tau^2-2ar_\tau\cos(\theta_\tau-\mu).\end{aligned}$$

The joint distribution with (4.3) is reduced to the Möbius distribution when $\alpha=2$, and it is asymmetric for the length r_τ, but still symmetric for the angle θ_τ.

Alternatively we propose a Möbius type transformation given by

$$M_{\mathrm{C}}(\tau)=|\tau|\,\frac{\tau/|\tau|+c_0}{1+\overline{c_0}\tau/|\tau|}, \qquad (4.4)$$

where c_0 is a complex number in the unit disc. This transformation maps the circle with radius r onto itself. We apply this transformation to the joint distribution (4.3) to obtain an asymmetric distribution for both of the length and direction.

Model I

Let $c_0=a_0\mathrm{e}^{\mathrm{i}\mu_0}$, where $0\le a_0<1$, $-\pi\le\mu_0<\pi$, and apply the transformation (4.4) to the joint distribution (4.3), then

$$\begin{aligned}\tau &= |w|\frac{w/|w|-c_0}{1-\bar{c}_0w/|w|},\\ r_\tau &= r,\\ \theta_\tau &= \arctan\left\{\frac{\sin\theta-2a_0\sin\mu_0-a_0^2\sin(\theta-2\mu_0)}{\cos\theta-2a_0\cos\mu_0+a_0^2\cos(\theta-2\mu_0)}\right\}\end{aligned}$$

for $\tau=r_\tau\mathrm{e}^{\mathrm{i}\theta_\tau}$ and $w=r\mathrm{e}^{\mathrm{i}\theta}$. The Jacobian for this transformation is

$$J=\left|\left|\frac{\partial(r_\tau,\theta_\tau)}{\partial(r,\theta)}\right|\right|=\frac{1-a_0^2}{1+a_0^2-2a_0\cos(\theta-\mu_0)},$$

so the joint density function of R and Θ is

$$\begin{aligned}f_1(r,\theta) &= \frac{\alpha\beta\,(1-a^2)^2\,(1-a_0^2)\,r}{2\pi\,\{1+a_0^2-2a_0\cos(\theta-\mu_0)\}\,\{B_1(r,\theta)\}^2} \\ &\quad\times\left\{\frac{A_1(r,\theta)}{B_1(r,\theta)}\right\}^{\frac{\alpha}{2}-1}\left[1-\left\{\frac{A_1(r,\theta)}{B_1(r,\theta)}\right\}^{\frac{\alpha}{2}}\right]^{\beta-1},\end{aligned}$$

where

$$\begin{aligned}A_1(r,\theta) &= a^2+r^2-2ar\cos(\theta_\tau-\mu) \\ &= a^2+r^2-2ar\frac{\cos(\theta-\mu)-2a_0\cos(\mu-\mu_0)+a_0^2\cos(\theta+\mu-2\mu_0)}{1+a_0^2-2a_0\cos(\theta-\mu_0)}, \\ B_1(r,\theta) &= 1+a^2r^2-2ar\cos(\theta_\tau-\mu) \\ &= 1+a^2r^2-2ar\frac{\cos(\theta-\mu)-2a_0\cos(\mu-\mu_0)+a_0^2\cos(\theta+\mu-2\mu_0)}{1+a_0^2-2a_0\cos(\theta-\mu_0)}.\end{aligned}$$

As a special case, if $\alpha=2$, the joint density function of R and Θ reduces to

$$f_1(r,\theta) = \frac{\alpha\beta\,(1-a^2)^{\beta+1}\,(1-a_0^2)\,r(1-r^2)^{\beta-1}}{2\pi\,\{1+a_0^2-2a_0\cos(\theta-\mu_0)\}\,\{B_1(r,\theta)\}^{\beta+1}}.$$

The marginal density function of R is, independently of α,

$$f_r(r)=\int_0^{2\pi}f(r,\theta)\,\mathrm{d}\theta=\frac{2\beta\,(1-a^2)^{\beta+1}\,r(1-r^2)^{\beta-1}}{(1-a^2r^2)^{\beta+1}}P_\beta\left(\frac{1+a^2r^2}{1-a^2r^2}\right),$$

which is the same marginal density function as that of the Möbius distribution, where $P_s(x)$ denotes the associated Legendre function. If $a_0=0$, the joint density function of R and Θ reduces to (4.3) because (4.4) reduces to the identity mapping. Finally, when $a=0$ the inverse Möbius transformation is just the identity transformation, and the joint density function of R and Θ reduces to

$$f(r,\theta) \;=\; \frac{\alpha\beta\,(1-a_0^2)\,r^{\alpha-1}(1-r^{\alpha})^{\beta-1}}{2\pi\left\{1+a_0^2-2a_0\cos\left(\theta-\mu_0\right)\right\}}.$$

This is an independent model for R and Θ, and in this case, R is distributied as Kumaraswamy's distribution, and Θ is distributed as a wrapped Cauchy distribution.

4.3.3 Some other models

In the transformation (4.4), let $c_0 = a_0 r \mathrm{e}^{\mathrm{i}\mu_0}$ instead of $c_0 = a_0 \mathrm{e}^{\mathrm{i}\mu_0}$. Then,

$$\begin{aligned}
\tau &= |w|\frac{w/|w|-c_0}{1-\bar{c}_0 w/|w|},\\
r_\tau &= r,\\
\theta_\tau &= \arctan\left\{\frac{\sin\theta-2a_0\sin\mu_0-a_0^2\sin\left(\theta-2\mu_0\right)}{\cos\theta-2a_0\cos\mu_0+a_0^2\cos\left(\theta-2\mu_0\right)}\right\}
\end{aligned}$$

for $\tau = r_\tau \mathrm{e}^{\mathrm{i}\theta_\tau}$ and $w = r\mathrm{e}^{\mathrm{i}\theta}$. The distribution (Model II) obtained has joint density

$$\begin{aligned}
f_2(r,\theta) \;=\;& \frac{\alpha\beta\,(1-a^2)^2\,(1-a_0^2 r^2)\,r}{2\pi\left\{1+a_0^2 r^2-2a_0 r\cos\left(\theta-\mu_0\right)\right\}\left\{B_2(r,\theta)\right\}^2}\\
&\times\left\{\frac{A_2(r,\theta)}{B_2(r,\theta)}\right\}^{\frac{\alpha}{2}-1}\left[1-\left\{\frac{A_2(r,\theta)}{B_2(r,\theta)}\right\}^{\frac{\alpha}{2}}\right]^{\beta-1}.
\end{aligned}$$

Here $A_2(r,\theta)$ and $B_2(r,\theta)$ are given by replacing a_0 in $A_1(r,\theta)$ and $B_1(r,\theta)$ with $a_0 r$. As a special case, when $\alpha = 2$, the marginal density function of R is the same as the marginal density of Model I. If $a_0 = 0$, the Model II reduces to the Möbius distribution. If $a = 0$, the distribution does not yield independence of R and Θ. Thus, Model II is completely different from Model I.

In the transformation (4.4), let $c_0 = \tanh\left(a_0 r_\tau\right)\mathrm{e}^{\mathrm{i}\mu_0}$, where $a_0 \geq 0,\ 0 \leq \mu_0 <$

2π, to obtain a distribution (Model III). Then,

$$
\begin{aligned}
\tau &= |w| \frac{w/|w| - c_0}{1 - \bar{c}_0 w/|w|}, \\
r_\tau &= r, \\
\theta_\tau &= \arctan \left\{ \frac{\sin\theta - 2\tanh(a_0 r)\sin\mu - \{\tanh(a_0 r)\}^2 \sin(\theta - 2\mu_0)}{\cos\theta - 2\tanh(a_0 r)\cos\mu + \{\tanh(a_0 r)\}^2 \cos(\theta - 2\mu_0)} \right\}.
\end{aligned}
$$

The Jacobian of the transformation is

$$
J = \left|\left| \frac{\partial(r_\tau, \theta_\tau)}{\partial(r, \theta)} \right|\right| = \frac{1}{\cosh(2a_0 r) - \cos(\theta - \mu_0)\sinh(2a_0 r)}
$$

and thus the joint density function of R and Θ is

$$
\begin{aligned}
f_3(r,\theta) &= \frac{\alpha\beta\left(1-a^2\right)^2 r}{2\pi\left(\cosh(2a_0 r) - \cos(\theta-\mu_0)\sinh(2a_0 r)\right)\{B_3(r,\theta)\}^2} \left\{ \frac{A_3(r,\theta)}{B_3(r,\theta)} \right\}^{\frac{\alpha}{2}-1} \\
&\quad \times \left[1 - \left\{ \frac{A_3(r,\theta)}{B_3(r,\theta)} \right\}^{\frac{\alpha}{2}} \right]^{\beta-1},
\end{aligned}
$$

where $A_3(r,\theta)$ and $B_3(r,\theta)$ are given by replacing a_0 in $A_1(r,\theta)$ and $B_1(r,\theta)$ with $\tanh(a_0 r)$. As a special case, R has the same marginal distribution as the Möbius distribution when $\alpha = 2$.

4.4 Illustrative example

The ozone concentration and wind direction data, collected at a weather station in Milwaukee, U.S.A., in 1975, are used for an example. The size of the dataset is 19, given in Table 1 of Johnson and Wehrly (1977). For the purposes of this example, we divide the ozone concentration data by a maximal value taken to be 120, and wind direction converted from degrees on [0 360) to radians on $[-\pi\ \pi)$, and plotted in Figure 4.5.

For the wind directions, the sample mean direction and mean resultant length

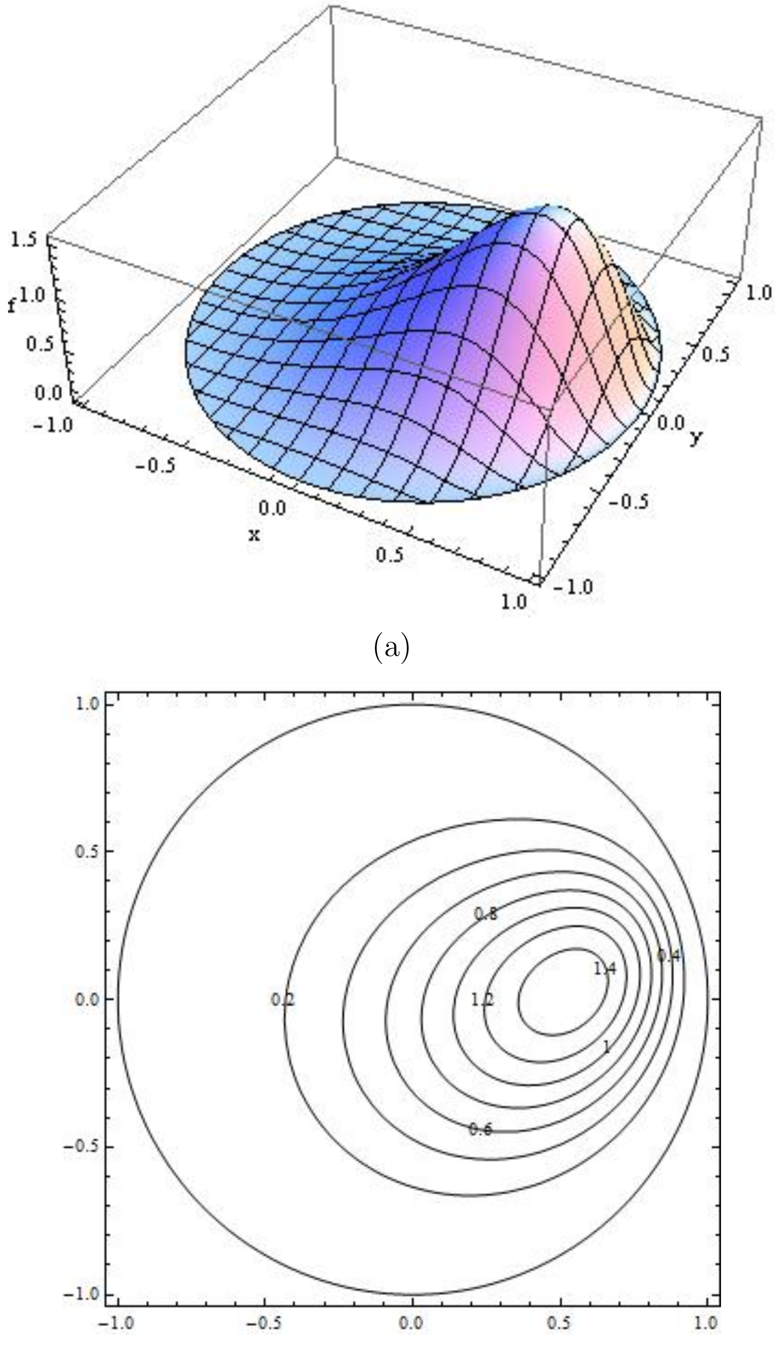
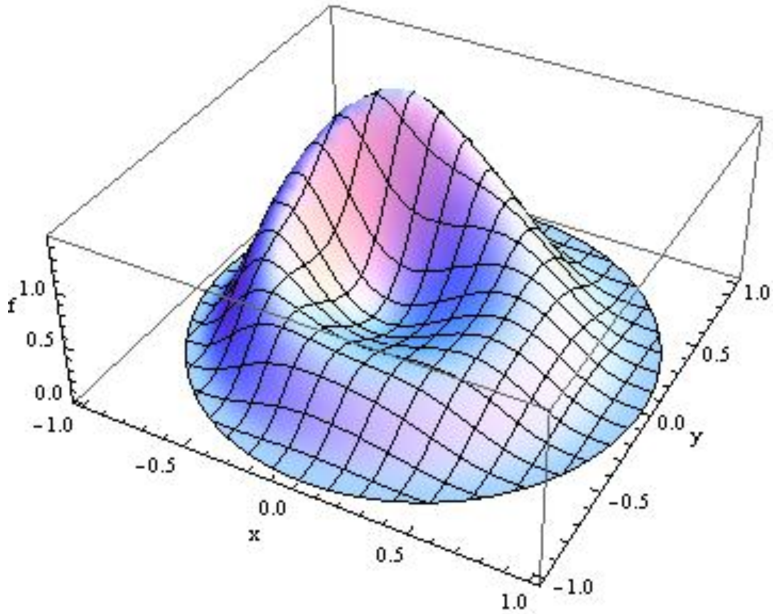
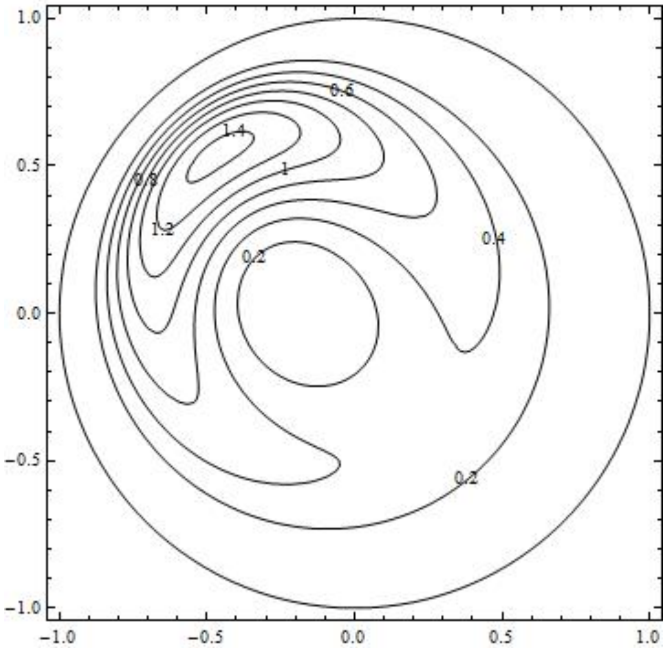


Figure 4.3: (a) 3-D plot and (b) contour plot of Model III with parameters $a = 0.25$, $a_0 = 0.5$, $\alpha = 2$, $\beta = 3$, $\mu = -\pi/5$ and $\mu_0 = \pi/4$.



(a)



(b)

Figure 4.4: (a) 3-D plot and (b) contour plot of Model III with parameters $a = 0.2$, $a_0 = 0.4$, $\alpha = 4$, $\beta = 5$, $\mu = \pi$ and $\mu_0 = \pi/2$.

are 0.292 radians (16.707 degrees) and 0.517 respectively. We test symmetry (Pewsey, 2002) for the directions. The p-value of the test is 0.460, and thus we do not reject the hypothesis that the wind directions are coming from a symmetric circular distribution. Then we fit the Möbius distribution and model Ⅲ to this paired dataset. The maximum likelihood estimates of the parameters, the maximum log-likelihood (MLL), AIC and BIC values for the Möbius distribution and model Ⅲ are given in Table 4.1. Figures 4.6 and 4.7 show the 3-D and contour plots of the fitted models as well as the scatter plots. For this data set, the Möbius distribution is selected as an optimal model in the sense of minimum AIC, although model Ⅲ has greater MLL value.

Table 4.1: Maximum likelihood estimates of the parameters, the maximum log-likelihood (MLL), AIC and BIC values for the Möbius distribution and model Ⅲ.

Model	$\hat{a}$	$\hat{a_0}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\mu_0}$	MLL	AIC	BIC
Möbius	0.284	–	–	3.783	0.648	–	-27.655	61.309	64.142
model Ⅲ	0.263	0.233	2.000	3.610	0.875	-0.327	-27.639	67.279	72.945

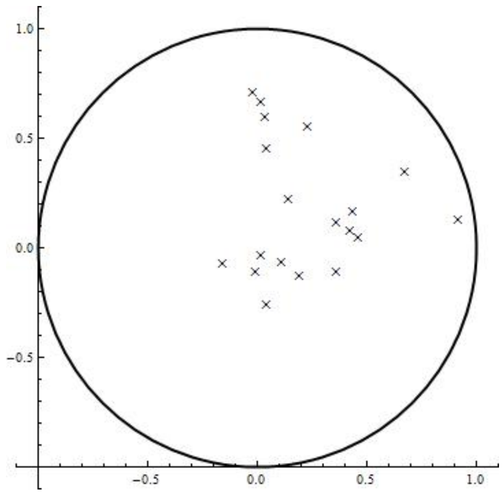
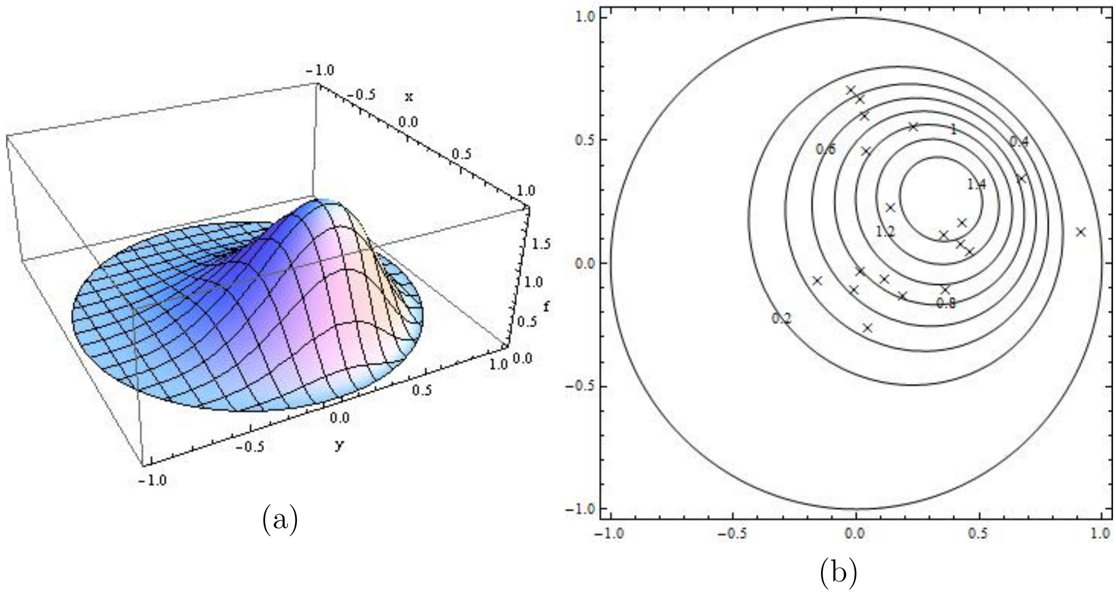
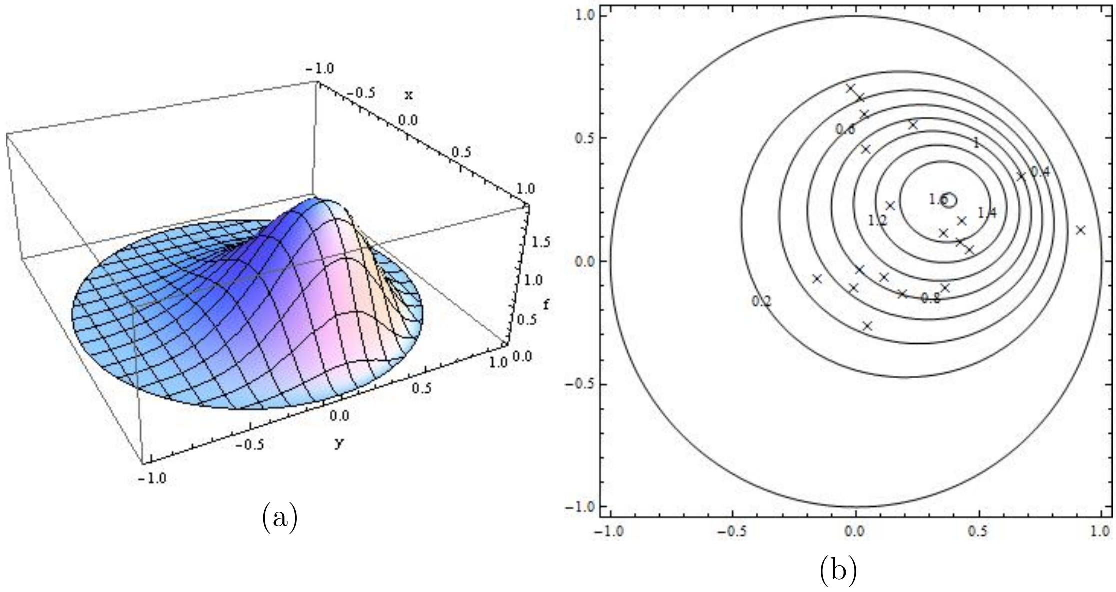


Figure 4.5: A scatterplot of the ozone concentration/wind direction data from Johnson and Wehrly (1977).



(a) (b)

Figure 4.6: (a) 3-D plot and (b) contour plot of fitted Möbius distribution with scatter plot for the data set of Johnson and Wehrly (1977): $\hat{a} = 0.284$, $\hat{\beta} = 3.783$, $\hat{\mu} = 0.648$.



(a) (b)

Figure 4.7: (a) 3-D plot and (b) contour plot of fitted model III with scatter plot for the data set of Johnson and Wehrly (1977): $\hat{a} = 0.263$, $\hat{a_0} = 0.233$, $\hat{\alpha} = 2.000$, $\hat{\beta} = 3.610$, $\hat{\mu} = 0.875$, $\hat{\mu_0} = -0.327$.

Appendix

Local influence method

The local influence method was first proposed by Cook (1986) for assessing the influence of small perturbations in a general statistical model. Let $L(\eta)$ represent the log-likelihood for the postulated (i.e. unperturbed) model and observed data, where η is a vector of unknown parameters with its maximum likelihood estimator $\hat{\eta}$. Let w denote a vector of the (small) perturbations in the model, Ω represent the open set of relevant perturbations such that $w \in \Omega$, and then $L(\eta|w)$ be the log-likelihood of the perturbed model and $\hat{\eta}_w$ denote the corresponding maximum likelihood estimator of η. Let $w_0 \in \Omega$ denote a no-perturbation vector such that $L(\eta) = L(\eta|w_0)$. Suppose that $L(\eta|w)$ is twice continuously differentiable in a neighborhood of $(\hat{\eta}, w_0)$. We are interested in comparing the parameter estimates $\hat{\eta}$ and $\hat{\eta}_w$ by using the idea of local influence. To implement the idea is to investigate the extent to which the inference is affected by the corresponding perturbation.

As in Cook (1986), the likelihood displacement is chosen to be

$$LD(w) = 2\{L(\hat{\eta}) - L(\hat{\eta}_w)\}$$

which can be used to assess the influence of the perturbation w. It is not difficult to see that large values of $LD(w)$ indicate that $\hat{\eta}$ and $\hat{\eta}_w$ differ considerably relative to the contours of the unperturbed log-likelihood $L(\eta)$. This method is based on studying the local behaviour of an influence graph $a(w) = (w', LD(w))'$ around w_0. Cook (1986) suggests investigating the direction in which this influence measure changes most rapidly locally, i.e. the maximum curvature of the surface $a(w)$. Upon $LD(w)$ the maximum curvature $C_{\max}$ is given by $C_{\max} = \max_{||l||=1} C_l$, where $C_l = 2|l'Fl|$. To find $C_{\max}$ and the corresponding direction $l_{\max}$, we need to calculate the matrix F, which is defined by

$$F = -\Delta' H^{-1} \Delta,$$

where Δ is a matrix for the perturbed model

$$\Delta \equiv \Delta(\hat{\eta}, w_0) = \left. \frac{\partial^2 L(\eta|w)}{\partial \eta \partial w'} \right|_{\eta=\hat{\eta}, w=w_0}$$

evaluated at $\hat{\eta}$ and w_0, and $-H$ is the observed information matrix for the postulated model

$$H \equiv H(\hat{\eta}) = \left. \frac{\partial^2 L(\eta)}{\partial \eta \partial \eta'} \right|_{\eta=\hat{\eta}}$$

evaluated at $\hat{\eta}$; then $l_{\max}$ is a unit-length eigenvector that is associated with the largest absolute eigenvalue of F, and large values of those elements of $l_{\max}$ indicate the corresponding observations are likely to be influential.

For each model of (3.11), (3.12) and (3.13), H and Δ for the five perturbation schemes are similarly obtainable. In the cosine model (3.11), we have the following H and Δ:

$$\begin{aligned} L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}(\sigma^2)^{-1}\varepsilon_i^2, \\ g(\eta, w) &= \sum_{i=1}^{n} g_i, \quad g_i = \frac{\partial l_i}{\partial \eta}, \end{aligned}$$

where

$$\begin{aligned}
\frac{\partial l_i}{\partial \beta} &= (\sigma^2)^{-1}\varepsilon_i x_i, \\
\frac{\partial l_i}{\partial \alpha} &= (\sigma^2)^{-1}\varepsilon_i \cos(\theta_i - \phi), \\
\frac{\partial l_i}{\partial \phi} &= (\sigma^2)^{-1}\alpha\varepsilon_i \sin(\theta_i - \phi), \\
\frac{\partial l_i}{\partial \sigma^2} &= -(1/2)(\sigma^2)^{-1} + (1/2)(\sigma^2)^{-2}\varepsilon_i^2
\end{aligned}$$

and

$$H(\eta) = \sum_{i=1}^{n} H_i, \quad H_i = \begin{pmatrix} h_{\beta\beta'} & h_{\beta\alpha} & h_{\beta\phi} & h_{\beta\sigma^2} \\ h_{\alpha\beta'} & h_{\alpha\alpha} & h_{\alpha\phi} & h_{\alpha\sigma^2} \\ h_{\phi\beta'} & h_{\phi\alpha} & h_{\phi\phi} & h_{\phi\sigma^2} \\ h_{\sigma^2\beta'} & h_{\sigma^2\alpha} & h_{\sigma^2\phi} & h_{\sigma^2\sigma^2} \end{pmatrix}.$$

Here

$$\begin{aligned}
h_{\beta\beta'} &= \frac{\partial^2 l_i}{\partial\beta\partial\beta'} = -(\sigma^2)^{-1}x_i x_i', \\
h_{\beta\alpha} &= h'_{\alpha\beta'} = \frac{\partial^2 l_i}{\partial\beta\partial\alpha} = -(\sigma^2)^{-1}x_i\cos(\theta_i - \phi), \\
h_{\beta\phi} &= h'_{\phi\beta'} = \frac{\partial^2 l_i}{\partial\beta\partial\phi} = -(\sigma^2)^{-1}x_i\alpha\sin(\theta_i - \phi), \\
h_{\beta\sigma^2} &= h'_{\sigma^2\beta'} = \frac{\partial^2 l_i}{\partial\beta\partial\sigma^2} = -(\sigma^2)^{-2}x_i\varepsilon_i, \\
h_{\alpha\alpha} &= \frac{\partial^2 l_i}{\partial\alpha^2} = -(\sigma^2)^{-1}\{\cos(\theta_i - \phi)\}^2, \\
h_{\alpha\phi} &= h_{\phi\alpha} = \frac{\partial^2 l_i}{\partial\alpha\partial\phi} = (\sigma^2)^{-1}\{\varepsilon_i - \alpha\cos(\theta_i - \phi)\}\sin(\theta_i - \phi), \\
h_{\alpha\sigma^2} &= h_{\sigma^2\alpha} = \frac{\partial^2 l_i}{\partial\alpha\partial\sigma^2} = -(\sigma^2)^{-2}\varepsilon_i\cos(\theta_i - \phi), \\
h_{\phi\phi} &= \frac{\partial^2 l_i}{\partial\phi^2} = -(\sigma^2)^{-1}\alpha\left[\varepsilon_i\cos(\theta_i - \phi) + \alpha\{\sin(\theta_i - \phi)\}^2\right],
\end{aligned}$$

$$
\begin{aligned}
h_{\phi\sigma^2} &= h_{\sigma^2\phi} = \frac{\partial^2 l_i}{\partial\phi\partial\sigma^2} = -(\sigma^2)^{-2}\alpha\varepsilon_i \sin(\theta_i - \phi), \\
h_{\sigma^2\sigma^2} &= \frac{\partial^2 l_i}{\partial{\sigma^2}^2} = \frac{1}{2}(\sigma^2)^{-2} - (\sigma^2)^{-3}\varepsilon_i^2.
\end{aligned}
$$

y-**scheme** with no-perturbation $n \times 1$ vector $w_0 = (0, \ldots, 0)'$:

$$
\begin{aligned}
y_i + w_i &= x_i'\beta + \alpha\cos(\theta_i - \phi) + \varepsilon_i, \quad i = 1, \ldots, n, \\
L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}(\sigma^2)^{-1}\varepsilon_i^2, \\
\Delta(\eta, w) &= (\Delta_1(\eta, w_1), \ldots, \Delta_n(\eta, w_n)), \\
\Delta_i(\eta, w_i) &= \begin{pmatrix} \frac{\partial^2 l_i}{\partial\beta\partial w_i} \\ \frac{\partial^2 l_i}{\partial\alpha\partial w_i} \\ \frac{\partial^2 l_i}{\partial\phi\partial w_i} \\ \frac{\partial^2 l_i}{\partial\sigma^2\partial w_i} \end{pmatrix} = (\sigma^2)^{-1} \begin{pmatrix} x_i \\ \cos(\theta_i - \phi) \\ \alpha\sin(\theta_i - \phi) \\ (\sigma^2)^{-1}\varepsilon_i \end{pmatrix}.
\end{aligned}
$$

x-**scheme** with no-perturbation $(2n) \times 1$ vector $w_0 = (0, \ldots, 0)'$:

$$
\begin{aligned}
y_i &= \beta_0 + (x_i^* + w_i^*)'\beta^* + \alpha\cos(\theta_i - \phi) + \varepsilon_i, \\
x_i^* &= (x_{1i}, x_{2i})', \ \beta^* = (\beta_1, \beta_2)', \\
w &= (w_1^{*\prime}, \ldots, w_n^{*\prime})', \ w_i^* = (w_{1i}, w_{2i})', \quad i = 1, \ldots, n, \\
L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}(\sigma^2)^{-1}\varepsilon_i^2, \\
\Delta(\eta, w) &= (\Delta_1(\eta, w_1^*), \ldots, \Delta_n(\eta, w_n^*)), \\
\Delta_i(\eta, w_i^*) &= \begin{pmatrix} \frac{\partial^2 l_i}{\partial\beta\partial w_i^{*\prime}} \\ \frac{\partial^2 l_i}{\partial\alpha\partial w_i^{*\prime}} \\ \frac{\partial^2 l_i}{\partial\phi\partial w_i^{*\prime}} \\ \frac{\partial^2 l_i}{\partial\sigma^2\partial w_i^{*\prime}} \end{pmatrix} = -(\sigma^2)^{-1} \begin{pmatrix} \beta^{*\prime} \\ (x_i^* + w_i^*)\beta^{*\prime} - \varepsilon_i I_2 \\ \beta^{*\prime}\cos(\theta_i - \phi) \\ \beta^{*\prime}\alpha\sin(\theta_i - \phi) \\ \beta^{*\prime}(\sigma^2)^{-1}\varepsilon_i \end{pmatrix},
\end{aligned}
$$

where I_2 denotes the 2×2 identity matrix.

θ-scheme with no-perturbation $n \times 1$ vector $w_0 = (0, \ldots, 0)'$:

$$\begin{aligned}
y_i &= x_i'\beta + \alpha \cos(\theta_i + w_i - \phi) + \varepsilon_i, \quad i = 1, \ldots, n, \\
L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}(\sigma^2)^{-1}\varepsilon_i^2, \\
\Delta(\eta, w) &= (\Delta_1(\eta, w_1), \ldots, \Delta_n(\eta, w_n)), \\
\Delta_i(\eta, w_i) &= \left(\frac{\partial^2 l_i}{\partial\beta\partial w_i}, \frac{\partial^2 l_i}{\partial\alpha\partial w_i}, \frac{\partial^2 l_i}{\partial\phi\partial w_i}, \frac{\partial^2 l_i}{\partial\sigma^2\partial w_i}\right)' \\
&= (\sigma^2)^{-1} \begin{pmatrix} x_i\alpha\sin(\theta_i + w_i - \phi) \\ -\{\varepsilon_i - \alpha\cos(\theta_i + w_i - \phi)\}\sin(\theta_i + w_i - \phi) \\ \alpha\varepsilon_i\cos(\theta_i + w_i - \phi) + \alpha^2\{\sin(\theta_i + w_i - \phi)\}^2 \\ (\sigma^2)^{-1}\alpha\varepsilon_i\sin(\theta_i + w_i - \phi) \end{pmatrix}.
\end{aligned}$$

σ^2-scheme with no-perturbation $n \times 1$ vector $w_0 = (1, \ldots, 1)'$:

$$\begin{aligned}
y_i &= x_i'\beta + \alpha\cos(\theta_i - \phi) + \varepsilon_i, \quad i = 1, \ldots, n, \\
L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 + \frac{1}{2}\ln w_i - \frac{1}{2}(\sigma^2/w_i)^{-1}\varepsilon_i^2, \\
\Delta(\eta, w) &= (\Delta_1(\eta, w_1), \ldots, \Delta_n(\eta, w_n)), \\
\Delta_i(\eta, w_i) &= \begin{pmatrix} \frac{\partial^2 l_i}{\partial\beta\partial w_i} \\ \frac{\partial^2 l_i}{\partial\alpha\partial w_i} \\ \frac{\partial^2 l_i}{\partial\phi\partial w_i} \\ \frac{\partial^2 l_i}{\partial\sigma^2\partial w_i} \end{pmatrix} = (\sigma^2)^{-1}\varepsilon_i \begin{pmatrix} x_i \\ \cos(\theta_i - \phi) \\ \alpha\sin(\theta_i - \phi) \\ (1/2)(\sigma^2)^{-1}\varepsilon_i \end{pmatrix}.
\end{aligned}$$

l_i-scheme with no-perturbation $n \times 1$ vector $w_0 = (1, \ldots, 1)'$:

$$\begin{aligned}
y_i &= x_i'\beta + \alpha\cos(\theta_i - \phi) + \varepsilon_i, \quad i = 1, \ldots, n, \\
L(\eta, w) &= \sum_{i=1}^{n} l_i, \quad l_i = w_i\left\{-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}(\sigma^2)^{-1}\varepsilon_i^2\right\},
\end{aligned}$$

$$
\begin{aligned}
\Delta(\eta, w) &= (\Delta_1(\eta, w_1), \ldots, \Delta_n(\eta, w_n)) \,, \\
\Delta_i(\eta, w) &= \begin{pmatrix} \frac{\partial^2 l_i}{\partial\beta\partial w_i} \\ \frac{\partial^2 l_i}{\partial\alpha\partial w_i} \\ \frac{\partial^2 l_i}{\partial\phi\partial w_i} \\ \frac{\partial^2 l_i}{\partial\sigma^2\partial w_i} \end{pmatrix} = (\sigma^2)^{-1} \begin{pmatrix} \varepsilon_i x_i \\ \varepsilon_i \cos(\theta_i - \phi) \\ \alpha\varepsilon_i \sin(\theta_i - \phi) \\ -(1/2) + (1/2)(\sigma^2)^{-1}\varepsilon_i^2 \end{pmatrix}.
\end{aligned}
$$

As an illustration, we use the original data studied with the SU models. The data consist of the measurements of the solar radiation for a period of six days and half-hourly record starting from 9:00 am to 5:30 pm. This period was chosen due to the uninterrupted weather condition which provided sunshine throughout. Same as SenGupta and Ugwuowo (2006) we associate these hours with the angles $0°, 20°, 40°, \ldots, 340°$, respectively, which correspond to θ_i. Thus, the group 340° - 360° corresponds to 5:30 pm - 6:00 pm, and 6:00 pm and 9:00 am are interpreted as the same beginning/end point on the circle. The interval [6:00 pm, 9:00 am) is irrelevant for us and is treated as a vacuous one. The other predictor variables x_{1i} and x_{2i} are ambient temperature, i.e. the atmospheric temperature observed at the experimental site, and control temperature to predict the absorber temperature y_i, i.e. the temperature of the water, in a well-constructed Thermosyphon solar water heater. The list plot of the solar energy data is given in Figure 1, followed by our MLE's as well as the values of Maximized Log-Likelihood (MLL), Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) in Table 1. The cosine-cosine model is selected as an optimal model among the three in the sense of minimum AIC, while the simple cosine model is selected in the sense of minimum BIC. However, it can be observed that there are not so much differences about the AIC and BIC values between the cosine-cosine and cosine models.

We only give the results of influential diagnostics from the Cos-cos model. We do not deal with the θ-scheme in this example because θ_i indicate angles trans-

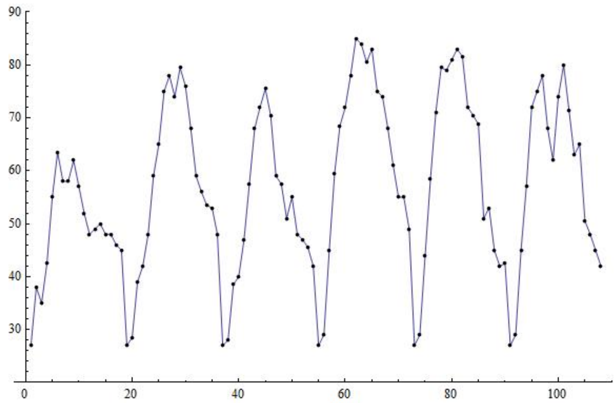


Figure 1: Plot of the solar energy data.

Table 1: MLE's, the values of Maximized Log-Likelihood (MLL), Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Model	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}$	$\hat{\phi}$	$\hat{\gamma}$	$\hat{\sigma}^2$	MLL	AIC	BIC
Cos	-156.53	5.20	1.44	-6.18	0.038	–	26.70	-330.6	673.2	689.3
Cos-sin	-156.53	5.20	1.44	-6.18	0.038	0.00	26.70	-330.6	675.2	694.0
Cos-cos	-153.21	5.23	1.32	-7.03	0.018	-0.37	26.00	-329.2	672.4	691.3

formed from hours. Moreover, note that the perturbation in the x-scheme for the cosine-cosine model is $y_i = \beta_0 + (x_{1i} + w_{1i})\beta_1 + x_{2i}\beta_2 + \alpha \cos\{\theta_i - \phi + \gamma \cos(\theta_i - \phi)\} + \varepsilon_i$ because x_{2i} denotes control temperature. From the plots we see that the elements corresponding to observations 7 and 8 appear to be standing out from the rest, and so these observations are possibly most influential.

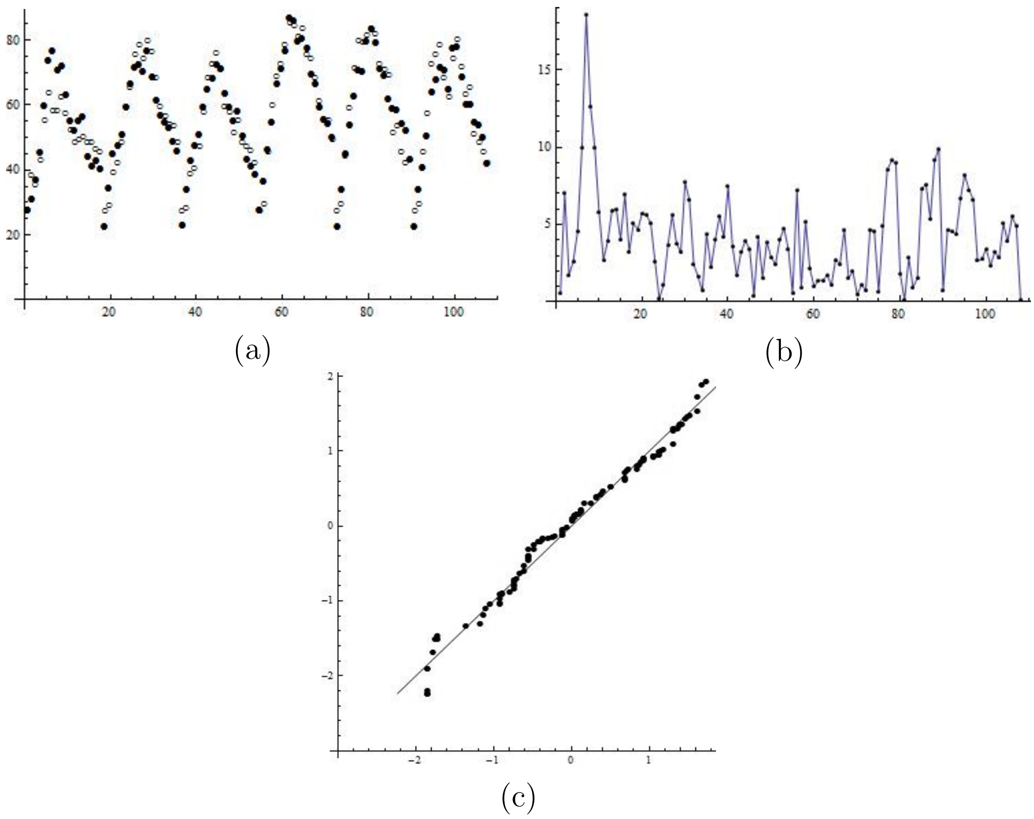


Figure 2: Cosine-cosine model. (a) Index plot of absorber temperatures (y_i $\circ$) and their prediction ($\hat{y}_i$ $\bullet$), (b) Index plot of $|y_i - \hat{y}_i|$, (c) Q-Q plot of the standardized residuals to the empirical distribution on the horizontal axis.

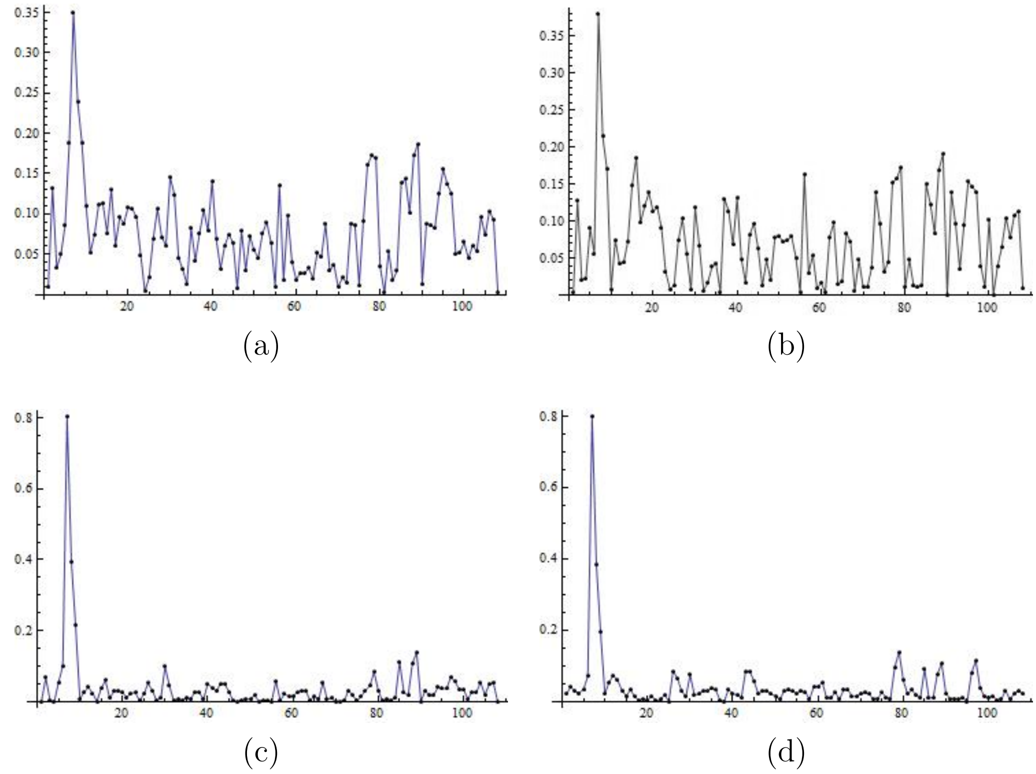

(a) (b) (c) (d)

Figure 3: Cosine-cosine model. Index plots of the elements of $l_{\max}$ for (a) y-scheme, (b) x-scheme, (c) σ^2-scheme, and (d) l_i-scheme.

Acknowledgement

First of all, I would like to express my deepest gratitude to Professor Kunio Shimizu for his supervision over the entire work. My sincere gratitude also goes to Professor Ritei Shibata, Professor Mihoko Minami and Professor Hideo Suzuki for various insightful comments and suggestions on the paper. I also most greatful to Professor Ashis SenGupta, Dr. Shuangzhe Liu, Dr. Kagumi Uesu, Dr. Shogo Kato and Dr. Toshihiro Abe for their valuable advice during my research.

Bibliography

[1] Abe, T. and Pewsey, A. (2011). Sine-skewed circular distributions. *Statistical Papers*, **52**, 683–707.

[2] Aronld, B. C. and Strass, D. J. (1991). Bivariate distributions with conditionals in prescribed exponential families. *Journal of the Royal Statistical Society, Series B* , **53**, 365–375.

[3] Baba, Y. (1981). Statistics of angular data: Wrapped normal distribution model. *Proceedings of the Institute of Statistical Mathematics*, **28**, 41–54. (in Japanese)

[4] Batschelet, E. (1981). *Circular Statistics in Biology*. Academic Press, London.

[5] Cartwright, D. E. (1963). The Use of Directional Spectra in Studying the Output of a Wave Recorder on a Moving Ship. in *Ocean Wave Spectra*, Englewood Cliffs, NJ: Prentice-Hall, pp. 203–218.

[6] Cheng, C. L. and Van Ness, J. W. (1999). *Statistical Regression with Measurement Error*. Arnold, London.

[7] Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society: Series B*, **48**, 133–169.

[8] Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press, Cambridge.

[9] Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. World Scientific, Singapore.

[10] Jeffreys, H. (1961). *Theory of Probability*. Third Edition. Clarendon Press, Oxford.

[11] Johnson, R. A. and Wehrly, T. E. (1977). Measures and models for angular correlation and angular-linear correlation. *Journal of the Royal Statistical Society: Series B*, **39**, 222–229.

[12] Johnson, R. A. and Wehrly, T. E. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, **73**, 602–606.

[13] Jones, M. C. (2002). Marginal replacement in multivariate densities, with application to skewing spherically symmetric distributions. *Journal of Multivariate Analysis*, **81**, 85–99.

[14] Jones, M. C. (2004). The Möbius distribution on the disc. *Annals of the Institute of Statistical Mathematics*, **56**, 733–742.

[15] Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, **6**, 70–81.

[16] Jones, M. C. and Pewsey, A. (2005). A family of symmetric distributions on the circle. *Journal of the American Statistical Association*, **100**, 1422–1428.

[17] Jupp, P. E. and Mardia, K. V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika*, **67**, 163–173.

[18] Kato, S. (2009). A distribution for a pair of unit vectors generated by Brownian motion. *Bernoulli*, **15**, 898–921.

[19] Kato, S. and Jones, M. C. (2010). A family of distributions on the circle with links to, and applications arising from, Möbius transformation. *Journal of the American Statistical Association*, **105**, 249–262.

[20] Kato, S. and Shimizu, K. (2008). Dependent models for observations which include angular ones. *Journal of Statistical Planning and Inference*, **138**, 3538–3549.

[21] Kato, S., Shimizu, K. and Shieh, G. S-R. (2008). A circular-circular regression model. *Statistica Sinica*, **18**, 633–645.

[22] Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, **46**, 79–88.

[23] Liu, S., Wang, M.-Z., Shimizu, K. and SenGupta, A. (2013). Influence diagnostics in possibly asymmetric circular-linear multivariate regression models. Manuscript.

[24] Mardia, K. V. (1975). Statistics of directional data (with discussion). *Journal of the Royal Statistical Society: Series B*, **37**, 349–393.

[25] Mardia, K. V. and Jupp, P. E. (1999). *Directional Statistics*. Wiley, Chichester.

[26] Mardia, K. V. and Sutton, T. W. (1978). A model for cylindrical variables with applications. *Journal of the Royal Statistical Society: Series B*, **40**, 229–233.

[27] Mardia, K. V., Taylor, C. C. and Subramaniam, G. K. (2007). Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data. *Biometrics*, **63**, 505–512.

[28] McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J. and Morales, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Society of America*, **82**, 335–349.

[29] Pewsey, A. (2002). Testing circular symmetry. *The Canadian Journal of Statistics*, **30**, 591–600.

[30] Rivest, L. P. (1988). A distribution for dependent unit vectors. *Communications in Statistics*, A, **17**, 461–483.

[31] SenGupta, A. (2004). On the constructions of probability distributions for directional data. *Bulletin of Calcutta Mathematical Society*, **96**, 139–154.

[32] SenGupta, A. and Ugwuowo, F. I. (2006). Asymmetric circular-linear multivariate regression models with applications to environmental data. *Environmental and Ecological Statistics*, **13**, 299–309.

[33] Shieh, G. S. and Johnson, R. A. (2005). Inferences based on a bivariate distribution with von Mises marginals. *Annals of the Institute of Statistical Mathematics*, **57**, 789–802.

[34] Shieh, G. S., Zheng, S., Johnson, R. A., Chang, Y.-F., Shimizu, K., Wang, C.-C. and Tang, S.-L. (2011). Modeling and comparing the organization of circular genomes. *Bioinformatics*, **27**, 912–918.

[35] Shimizu, K. and Iida, K. (2002). Pearson type VII distributions on spheres. *Communications in Statistics–Theory and Methods*, **31**, 513–526.

[36] Singh, H., Hnizdo, V., and Demchuk, E. (2002). Probabilistic model for two dependent circular variables. *Biometrika*, **89**, 719–723.

[37] Wang, M.-Z. and Shimizu, K. (2012). On applying Möbius transformation to cardioid random variables. *Statistical Methodology*, **9**, 604–614.

[38] Wang, M.-Z., Shimizu, K. and Uesu, K. An analysis of earthquakes latitude, longitude and magnitude data by use of directional statistics. *Japanese Journal of Applied Statistics*, to appear. (in Japanese)

[39] Wehrly, T. E. and Johnson, R. A. (1979). Bivariate models for dependence of angular observations and a related Markov process. *Biometrika*, **66**, 255–256.