An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion

SHOGO KATO[1] and M.C. JONES[2]

[1]*Institute of Statistical Mathematics, 10-3 Midori-Cho, Tachikawa, Tokyo 190-8562, Japan.*
E-mail: skato@ism.ac.jp
[2]*Department of Mathematics & Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA,*
UK. E-mail: m.c.jones@open.ac.uk

We introduce a four-parameter extended family of distributions related to the wrapped Cauchy distribution on the circle. The proposed family can be derived by altering the settings of a problem in Brownian motion which generates the wrapped Cauchy. The densities of this family have a closed form and can be symmetric or asymmetric depending on the choice of the parameters. Trigonometric moments are available, and they are shown to have a simple form. Further tractable properties of the model are obtained, many by utilizing the trigonometric moments. Other topics related to the model, including alternative derivations and Möbius transformation, are considered. Discussion of the symmetric submodels is given. Finally, generalization to a family of distributions on the sphere is briefly made.

Keywords: asymmetry; circular Cauchy distribution; directional statistics; four-parameter distribution; trigonometric moments

1. Introduction

As a unimodal and symmetric model on the circle, the wrapped Cauchy or circular Cauchy distribution has played an important role in directional statistics. It has density

$$f(\theta; \mu, \rho) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)}, \qquad -\pi \leq \theta < \pi, \tag{1}$$

where μ ($\in [-\pi, \pi)$) is a location parameter, and ρ ($\in [0, 1)$) controls the concentration of the model. This distribution has some desirable mathematical features, as discussed in Kent and Tyler [11] and McCullagh [15]. We write $\Theta \sim \mathrm{WC}(\mu, \rho)$ if a random variable Θ has density (1).

In modeling symmetric circular data, the wrapped Cauchy distribution could be one choice as well as other familiar symmetric models such as the von Mises and wrapped normal distributions; see also Jones and Pewsey [7]. In reality, however, it is not so common to find such data in fields of application. In many cases the data of interest are asymmetrically distributed, and therefore probability distributions having asymmetric densities are desired.

Construction of a tractable circular model with an asymmetric shape has been a problem in statistics of circular data. To tackle this problem, some asymmetric extensions of well-known circular models have been proposed in the literature. Maksimov [12], Yfantis and Borgman [20] and

Gatto and Jammalamadaka [4] discussed an extension of the von Mises distribution, generated through maximization of Shannon's entropy with restrictions on certain trigonometric moments. Batschelet [2] proposed a mathematical method of skewing circular distributions that has seen renewed interest very recently. Pewsey [16] presented a four-parameter family of distributions on the circle by wrapping the stable distribution, which is asymmetric, onto the circle. Recent work by Kato and Jones [9] proposed a family of distributions arising from the Möbius transformation which includes the von Mises and wrapped Cauchy distributions. Unlike familiar symmetric distributions, it is often difficult to deal with skew models in statistical analysis. This difficulty is partly due to the lack of some mathematical properties that many of the well-known symmetric models have. For example, existing asymmetric models often have complex normalizing constants and trigonometric moments, which could cause trouble in analysis.

In this paper, we provide a four-parameter extended family of circular distributions, based on the wrapped Cauchy distribution, by applying Brownian motion which, to our knowledge, has not been used to propose a skew distribution on the circle. The four parameters enable the model to describe not only symmetric shapes but also asymmetric ones. An advantage of the proposed model is its mathematical tractability. For instance, it has a simple normalizing constant and trigonometric moments. The current proposal is more tractable than the family discussed by Kato and Jones [9], but is complementary to it in the sense that the latter has other advantages (particularly some associated with Möbius transformation).

The subsequent sections are organized as follows. In Section 2 we make the main proposal of this paper. The derivation of the proposed model is given, and the probability density function and probabilities of intervals under the density are discussed. Also, special cases of the model are briefly considered. Section 3 concerns the shapes of the density. Conditions for symmetry and for unimodality are explored, and the interpretation of the parameters is discussed through pictures of the density. In Section 4 we discuss the trigonometric moments and problems related to them. It is shown that the trigonometric moments can be expressed in a simple form. The mean direction, mean resultant length and skewness of the model are also considered. Some other topics concerning our model are provided in Section 5. Apart from the derivation given in Section 2, there are other methods to generate the family, which are discussed in that section. In addition, we study the conformal transformation properties of the distribution. In Section 6 we investigate some properties of the symmetric cases of the proposed model. These submodels have some properties which the general family does not have. Finally, in Section 7, generalization to a family of distributions on the sphere is made, and its properties are briefly discussed.

Further material on the topic of this paper can be found in its technical report version, Kato and Jones [10]. This includes the path of the expected exit points of a Brownian particle, random variate generation using the Markov chain Monte Carlo method, and proofs of Theorems 1, 3 and 10.

2. A family of distributions on the circle

2.1. Definition

It is a well-known fact that the wrapped Cauchy distribution can be generated as the distribution of the position of a Brownian particle which exits the unit disc in the two-dimensional plane; see,

for example, Section 1.10 in Durrett [3]. In addition to the wrapped Cauchy circular model, some distributions on certain manifolds are derived from, or have relationships with, Brownian motion. The Cauchy distribution on the real line is the distribution of the position where a Brownian particle exits the upper half plane. Considering the hitting time of the particle, one can obtain the inverse Gaussian distribution on the positive half-line. Kato [8] proposed a distribution for a pair of unit vectors by recording points where a Brownian particle hits circles with different radii.

By applying Brownian motion, we provide a family of asymmetric distributions on the circle, which includes the wrapped Cauchy distribution as a special case. The proposed model is defined as follows.

Definition 1. *Let* $\{B_t; t \geq 0\}$ *be* $\mathbb{R}^2$-*valued Brownian motion without drift starting at* $B_0 = \rho_1(\cos\mu_1, \sin\mu_1)'$, *where* $0 \leq \rho_1 < 1$ *and* $-\pi \leq \mu_1 < \pi$. *This Brownian particle will eventually hit the unit circle. Let* τ_1 *be the first time at which the particle exits the circle, that is,* $\tau_1 = \inf\{t; \|B_t\| = 1\}$. *After leaving the unit circle, the particle will hit a circle with radius* ρ_2^{-1} $(0 < \rho_2 < 1)$ *first at the time* τ_2, *meaning* $\tau_2 = \inf\{t; \|B_t\| = \rho_2^{-1}\}$. *Then the proposed model is defined by the conditional distribution of* B_{τ_1} *given* $B_{\tau_2} = \rho_2^{-1}(\cos\mu_2, \sin\mu_2)'$, *where* $-\pi \leq \mu_2 < \pi$.

To put it another way, the proposed random vector, B_{τ_1} given B_{τ_2}, represents the position where a Brownian particle first hits the unit circle, given the future point at which the particle exits a circle with a larger radius. From the next subsection, we investigate some properties of the proposed model.

2.2. Probability density function

One feature of the proposed model is that it has a closed form of density with simple normalizing constant. It is given in the following theorem. See Appendix A of Kato and Jones [10] for the proof.

Theorem 1. *Let* Θ *be a* $[-\pi, \pi)$-*valued random variable defined as* $B_{\tau_1} = (\cos\Theta, \sin\Theta)'$. *Then the conditional density of* Θ *given* $B_{\tau_2} = \rho_2^{-1}(\cos\mu_2, \sin\mu_2)'$ *is given by*

$$f(\theta) = C[\{1 + \rho_1^2 - 2\rho_1\cos(\theta - \mu_1)\}\{1 + \rho_2^2 - 2\rho_2\cos(\theta - \mu_2)\}]^{-1}, \qquad -\pi \leq \theta < \pi, \quad (2)$$

where $-\pi \leq \mu_1, \mu_2 < \pi$, $0 \leq \rho_1, \rho_2 < 1$, *and the normalizing constant* C *is*

$$C = \frac{(1 - \rho_1^2)(1 - \rho_2^2)\{1 + \rho_1^2\rho_2^2 - 2\rho_1\rho_2\cos(\mu_1 - \mu_2)\}}{2\pi(1 - \rho_1^2\rho_2^2)}.$$

For convenience, write $\Theta \sim \text{EWC}(\mu_1, \mu_2, \rho_1, \rho_2)$ if a random variable Θ has density (2) ("E" for "Extended"). Note that the density does not involve any infinite sums or special functions.

It is easy to see that distribution (2) reduces to the wrapped Cauchy (1) when either ρ_1 or ρ_2 is equal to zero.

Quite often, it is advantageous to express the random variable and parameters of the proposed model in terms of complex numbers rather than real numbers. Define a random vector by $Z = e^{i\Theta}$ where $\Theta \sim \text{EWC}(\mu_1, \mu_2, \rho_1, \rho_2)$. Then Z has density

$$f(z) = \frac{1}{2\pi} \frac{|1 - \phi_1 \overline{\phi_2}|^2}{1 - |\phi_1 \overline{\phi_2}|^2} \frac{1 - |\phi_1|^2}{|z - \phi_1|^2} \frac{1 - |\phi_2|^2}{|z - \phi_2|^2}, \qquad z \in \partial D, \tag{3}$$

with respect to arc length on the circle, where $\phi_1 = \rho_1 e^{i\mu_1}$, $\phi_2 = \rho_2 e^{i\mu_2}$ and $\partial D = \{z \in \mathbb{C}; |z| = 1\}$. It is clear that the parameters in this formulation, ϕ_1 and ϕ_2, take values on the unit disc in the complex plane denoted by $D = \{z \in \mathbb{C}; |z| < 1\}$. For brevity, we denote the distribution (3) by $\text{EC}^*(\phi_1, \phi_2)$. Also, as in McCullagh [15], write $Z \sim C^*(\phi_1)$ if Z follows distribution (3) with $\phi_2 = 0$.

We remark that the above change-of-variable and reparametrization do not actually change the distribution. In later sections, we utilize either of the representations (2) or (3), whichever is the more convenient.

2.3. Probabilities

As well as the density of the proposed model, the probabilities of intervals under the density can also be expressed without using infinite sums or special functions.

Theorem 2. *Let a random variable* Θ *follow the* $\text{EWC}(\mu_1, \mu_2, \rho_1, \rho_2)$ *distribution. If* $\rho_1 \neq \rho_2$ *or* $\mu_1 \neq \mu_2$, *then the probability of intervals under the density of* Θ *is given by*

$$P(a < \Theta \leq b)$$

$$= \int_a^b C[\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}\{1 + \rho_2^2 - 2\rho_2 \cos(\theta - \mu_2)\}]^{-1} \, d\theta$$

$$= \frac{C}{D} \left\{ \rho_1 \rho_2 \sin(\mu_1 - \mu_2) \left[\log\left\{ \frac{1 + \rho_2^2 - 2\rho_2 \cos(\theta - \mu_2)}{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)} \right\} \right]_a^b \right.$$

$$+ \frac{2\rho_1\{\rho_1(1 + \rho_2^2) - \rho_2(1 + \rho_1^2)\cos(\mu_1 - \mu_2)\}}{1 - \rho_1^2}$$

$$\times \left(\left[\arctan\left\{ \frac{1 + \rho_1}{1 - \rho_1} \tan\left(\frac{\theta - \mu_1}{2} \right) \right\} \right]_a^b + A_{\mu_1} \right)$$

$$+ \frac{2\rho_2\{\rho_2(1 + \rho_1^2) - \rho_1(1 + \rho_2^2)\cos(\mu_1 - \mu_2)\}}{1 - \rho_2^2}$$

$$\left. \times \left(\left[\arctan\left\{ \frac{1 + \rho_2}{1 - \rho_2} \tan\left(\frac{\theta - \mu_2}{2} \right) \right\} \right]_a^b + A_{\mu_2} \right) \right\},$$

where C is defined as in Theorem 1, $-\pi \le a < b < \pi$,

$$D = (\rho_1^2 + \rho_2^2)(1 + \rho_1^2 \rho_2^2) - 2\rho_1 \rho_2(1 + \rho_1^2)(1 + \rho_2^2)\cos(\mu_1 - \mu_2) + 4\rho_1^2 \rho_2^2 \cos^2(\mu_1 - \mu_2)$$

and

$$A_\mu = \begin{cases} \pi, & \tan\{\frac{1}{2}(a - \mu)\} > \tan\{\frac{1}{2}(b - \mu)\}, \\ 0, & \textit{otherwise.} \end{cases}$$

If $\rho_1 = \rho_2$ and $\mu_1 = \mu_2$, then

$$P(a < \Theta \le b) = \int_a^b C\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}^{-2} \, d\theta$$

$$= \frac{2C}{(1 - \rho_1^2)^2}\left\{\left[\frac{\rho_1 \sin(\theta - \mu_1)}{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)}\right]_a^b\right.$$

$$\left. + \frac{1 + \rho_1^2}{1 - \rho_1^2}\left(\left[\arctan\left\{\frac{1 + \rho_1}{1 - \rho_1}\tan\left(\frac{\theta - \mu_1}{2}\right)\right\}\right]_a^b + A_{\mu_1}\right)\right\}.$$

Proof. The result is straightforward from equations (2.553.3), (2.554.3) and (2.559.2) of Gradshteyn and Ryzhik [5]. $\square$

2.4. Special cases

The proposed family (2) contains some known distributions as special cases.

Case 1: The wrapped Cauchy distribution. As mentioned in Section 2.2, model (2) becomes the wrapped Cauchy WC(μ_1, ρ_1) if $\rho_2 = 0$. Similarly, the model is WC(μ_2, ρ_2) when $\rho_1 = 0$.

Case 2: A special case of the Jones and Pewsey [7] family. If $\rho_1 = \rho_2 (\equiv \rho)$ and $\mu_1 = \mu_2 (\equiv \mu)$, then density (2) reduces to

$$f(\theta) = \frac{1}{2\pi}\frac{1 - \rho^2}{1 + \rho^2}\left(\frac{1 - \rho^2}{1 + \rho^2 - 2\rho\cos(\theta - \mu)}\right)^2, \qquad -\pi \le \theta < \pi. \tag{4}$$

This submodel corresponds to a special case of the family presented by Jones and Pewsey [7]. Their model has density

$$f_{\text{JP}}(\theta) = \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi)\cos(\theta - \mu)\}^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}, \qquad -\pi \le \theta < \pi,$$

where $\kappa \ge 0$, $\psi \in \mathbb{R}$, $-\pi \le \mu < \pi$ and $P_{1/\psi}(z)$ is the associated Legendre function of the first kind of degree $1/\psi$ and order 0 (Gradshteyn and Ryzhik [5], Sections 8.7, 8.8). Our submodel is equivalent to their family with $\psi = -1/2$ and $\kappa = 2\log((1 + \rho)/(1 - \rho))$. In this case the associated Legendre function simplifies to $P_{-2}(z) = P_1(z) = z$. (See equations (8.2.1) and (8.4.3) of Abramowitz and Stegun [1].) This fairly heavy-tailed circular distribution is shown by Jones

and Pewsey to tend, suitably normalized, to the t distribution on $\sqrt{3}$ degrees of freedom as ($\rho \to 1$ and hence) $\kappa \to \infty$!

Case 3: One-point distribution. As ρ_1 (ρ_2) tends to one, the model converges to a point distribution with singularity at $\theta = \mu_1$ ($\theta = \mu_2$). Normalizing by a scale factor of $(1 - \rho)/\sqrt{\rho}$, the limiting version of the wrapped Cauchy distribution as $\rho \to 1$ is the ordinary Cauchy distribution. That argument can be extended to show that the Cauchy limiting distribution also arises in this case.

Case 4: Two-point distribution. Assume that $\rho_1 = \rho_2$ ($\equiv \rho$) and $\mu_1 \neq \mu_2$. When ρ goes to one, model (2) converges to the distribution of a random variable which takes values on μ_1 or μ_2 with probability 0.5. If $\mu_1 < \mu_2$ and ρ tends to one, $\{\sqrt{\rho}/(1 - \rho)\}(\Theta - \mu_1)$ converges to $\frac{1}{2}C(0, 1) + \frac{1}{2}I$, where $C(0, 1)$ is the standard Cauchy, and I has the distribution function $F(x) = \lim_{\lambda \to \infty} \int_{-\infty}^{x} 1/[\pi\{1 + (t - \lambda)^2\}] \, dt$. Roughly speaking, this limiting distribution is a 50:50 mixture of the standard Cauchy and a point distribution with singularity at $X = \infty$. Similarly, one can show that if $\mu_1 > \mu_2$ and ρ tends to one, $\{\sqrt{\rho}/(1 - \rho)\}(\Theta - \mu_1)$ converges to a 50:50 mixture of the standard Cauchy and a point distribution with singularity at $X = -\infty$. The limiting distribution of $\{\sqrt{\rho}/(1 - \rho)\}(\Theta - \mu_2)$ can be discussed in a similar manner.

Case 5: The circular uniform distribution. When $\rho_1 = \rho_2 = 0$, the distribution reduces to the circular uniform distribution.

3. Shapes of probability density function

3.1. Conditions for symmetry

As briefly stated in Section 1, the proposed model can be symmetric or asymmetric, depending on the values of the parameters. The condition for symmetry is clearly written out in the following theorem.

Theorem 3. *Density (2) is symmetric if and only if $\rho_1 = \rho_2$, $\rho_j = 0$ or $\mu_1 = \mu_2 + (j - 1)\pi$ ($j = 1, 2$).*

See Appendix B of Kato and Jones [10] for the proof. The three-parameter symmetric submodels arising under the above conditions will be discussed in Section 6.

3.2. Conditions for unimodality

Turning to conditions for unimodality, it is possible to express these in terms of an inequality. The process to obtain the inequality is similar to that in Kato and Jones [9], Section 2.5. In the following discussion, take $\mu_2 = 0$ without loss of generality. First we calculate the first derivative of density (2) with respect to θ, which is given by

$$\frac{d}{d\theta} f(\theta) \propto \frac{[\rho_1 \sin(\theta - \mu_1)(1 + \rho_2^2 - 2\rho_2 \cos\theta) + \rho_2 \sin\theta\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}]}{\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}^2 (1 + \rho_2^2 - 2\rho_2 \cos\theta)^2}.$$

Then it follows that the extrema of the density are obtainable as solutions of the following equation:

$$a_0 + a_1 \cos\theta + a_2 \sin\theta + a_3 \cos\theta \sin\theta + a_4 \cos^2\theta = 0, \tag{5}$$

where

$$a_0 = 2\rho_1\rho_2 \sin\mu_1, \qquad a_1 = \rho_1(1+\rho_2^2)\sin\mu_1, \qquad a_2 = -\rho_1(1+\rho_2^2)\cos\mu_1 - \rho_2(1+\rho_1^2),$$

$$a_3 = 4\rho_1\rho_2 \cos\mu_1, \qquad a_4 = -4\rho_1\rho_2 \sin\mu_1.$$

Following Yfantis and Borgman [20] and Kato and Jones [9], Section 2.5, put $x = \tan(\theta/2)$ so that $\cos\theta = (1-x^2)/(1+x^2)$, $\sin\theta = 2x/(1+x^2)$. It follows that equation (5) can be expressed as a quartic equation in x whose discriminant D (Uspensky [19]) can be written down in terms of $a_0, \ldots, a_4$ as in (6) of Kato and Jones [9]. The quartic equation has four real roots or four complex ones if $D > 0$, and two real roots and two complex ones if $D < 0$. Therefore the distribution is bimodal when $D > 0$ and unimodal when $D < 0$. Since a_j $(0 \le j \le 4)$ are functions of μ_1, ρ_1 and ρ_2, we can write the conditions for unimodality as a function of these three parameters. For general μ_2, it is easy to see that the discriminant is expressed in terms of $\mu_1 - \mu_2$, ρ_1 and ρ_2.

Figure 1 exhibits areas of positivity (bimodality of density) and negativity (unimodality of density) of the discriminant D for two pairs of parameters. Figure 1(a) suggests that density (2) with $\rho_1 = 0.5$ becomes unimodal for most cases when $-\pi/2 \le \mu_1 \le \pi/2$. It seems that bimodality is most likely to occur if $\pi/2 \le \mu_1 \le 3\pi/2$. The other frame of Figure 1 seems to show that distribution (2) with $\mu_2 = 2\pi/3$ is bimodal if ρ_1 or ρ_2 is sufficiently large, and ρ_2/ρ_1 takes a value

(a)   (b)
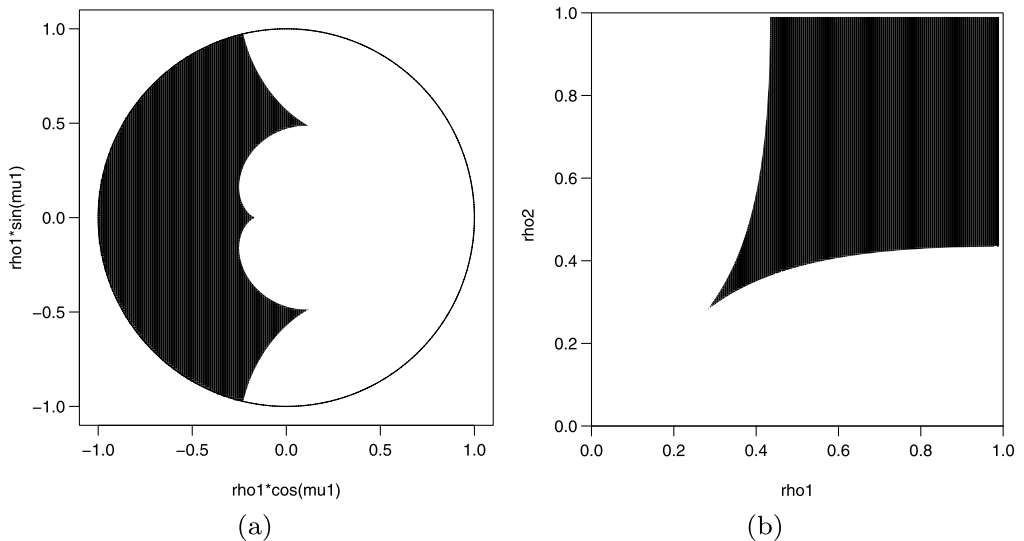
Figure 1. Discriminant D as functions of: (a) $(\rho_1\cos\mu_1, \rho_1\sin\mu_1)$ with $\rho_2 = 0.5$, and (b) (ρ_1, ρ_2) with $\mu_1 = 2\pi/3$. Positive D, corresponding to bimodality, is displayed in black, and negative D, corresponding to unimodality, is exhibited in white.

close to 1. In particular, when $\rho_1 = \rho_2$, the range of parameters corresponding to unimodality seems less wide than that when $\rho_1 = a\rho_2$ with $a \neq 1$. As will be shown in Section 6.2, for $\mu_1 = 2\pi/3$ and $\rho_1 = \rho_2$, the density becomes unimodal if $\rho_1 \leq 2 - \sqrt{3} \simeq 0.268$ and bimodal otherwise.

3.3. Pictures of density

We plot density (2) for selected values of the parameters in Figure 2. Confirming the results in the previous two subsections, this figure shows that the density can be symmetric or asymmetric, unimodal or bimodal, depending on the choice of the parameters. Figure 2(a), displaying the density for fixed values of μ_2, ρ_1 and ρ_2, suggests that the density is symmetric if $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$ and asymmetric otherwise. Figure 2(b) exhibits how the shape of density (2) changes as ρ_1 increases. It is clear from this frame that the greater the value of ρ_1, the greater the concentration of the density. The frame also implies that the absolute value of the circular skewness of the density with fixed μ_1 and μ_2 takes a large value if ρ_1/ρ_2 is close to 1. Some symmetric cases of the density are shown in Figure 2(c). It appears to be that the density is unimodal when $\mu_1 = \mu_2$, whereas bimodality can occur when $\mu_2 = \mu_1 + \pi$. Figure 2(d) is a case in which two parameters, ρ_1 and ρ_2, are fixed to be equal (to 0.5), also implying that density (2) is symmetric. We will focus on this submodel in Section 6.2.

4. Trigonometric moments and related problems

4.1. Trigonometric moments

It is often the case that asymmetric distributions on the circle have complicated trigonometric moments or Fourier coefficients. One feature of our asymmetric model, however, is the relative simplicity of its trigonometric moments. The expression for the moments is greatly simplified if the variables and parameters are represented in terms of complex numbers.
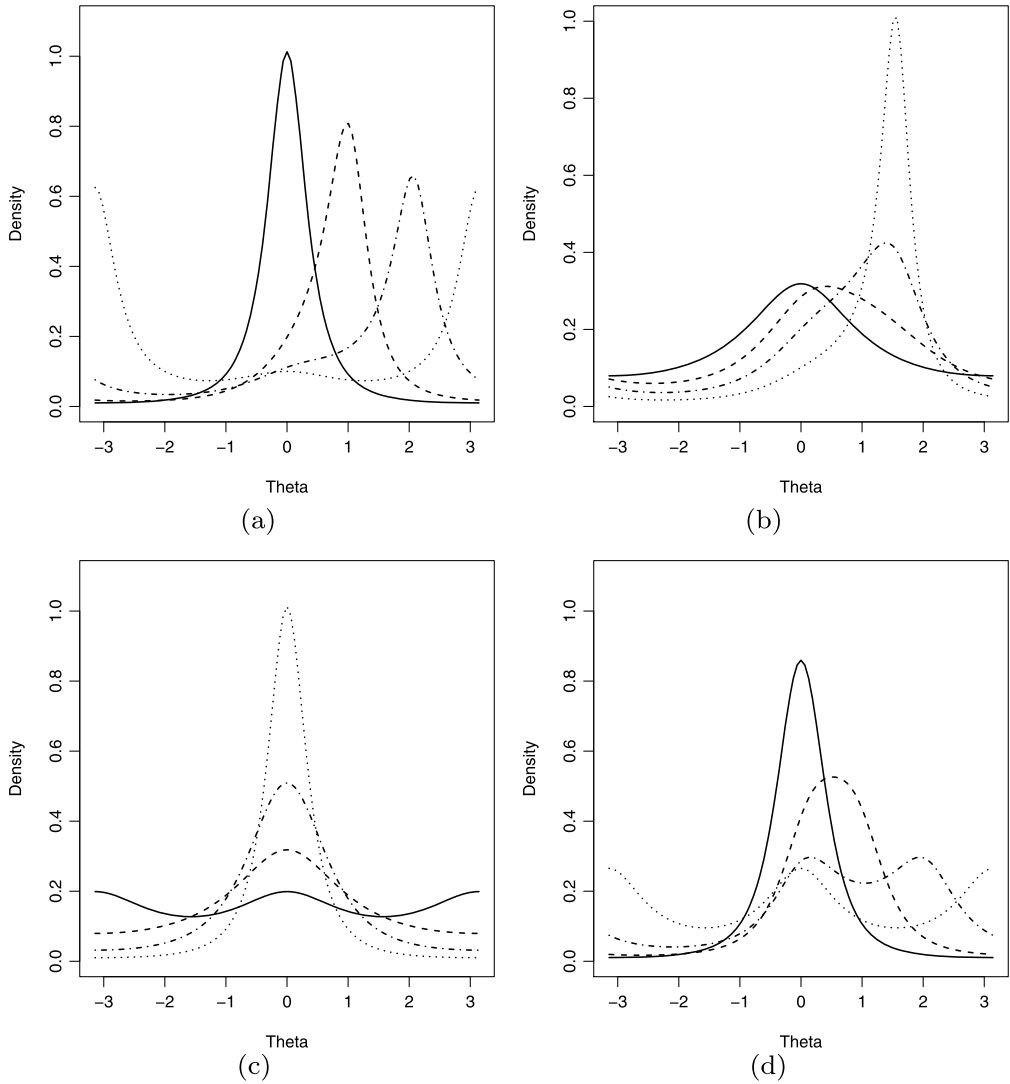
Theorem 4. *Let a random variable Z have the $\text{EC}^*(\phi_1, \phi_2)$ distribution. Then the nth trigonometric moment of Z is given by*

$$
E(Z^n) = \begin{cases} \dfrac{(1 - |\phi_2|^2)(1 - \overline{\phi_1}\phi_2)\phi_1^{n+1} - (1 - |\phi_1|^2)(1 - \phi_1\overline{\phi_2})\phi_2^{n+1}}{(\phi_1 - \phi_2)(1 - |\phi_1\overline{\phi_2}|^2)}, & \phi_1 \neq \phi_2, \\[4mm] \dfrac{1 + n + (1 - n)|\phi_1|^2}{1 + |\phi_1|^2}\phi_1^n, & \phi_1 = \phi_2. \end{cases}
$$

Proof. The nth trigonometric moment can be expressed as

$$
\begin{aligned}
E(Z^n) &= \frac{1}{2\pi} \frac{|1 - \overline{\phi_1}\phi_2|^2}{1 - |\phi_1\overline{\phi_2}|^2} \int_{\partial D} z^n \left(\prod_{j=1}^{2} \frac{1 - |\phi_j|^2}{|z - \phi_j|^2} \right) |\mathrm{d}z| \\
&= \frac{1}{2\pi \mathrm{i}} \frac{|1 - \overline{\phi_1}\phi_2|^2}{1 - |\phi_1\overline{\phi_2}|^2} \int_{\partial D} z^{n+1} \left\{ \prod_{j=1}^{2} \frac{1 - |\phi_j|^2}{(z - \phi_j)(1 - \overline{\phi_j}z)} \right\} \mathrm{d}z.
\end{aligned}
\tag{6}
$$

Figure 2. Density (2) for $\mu_2 = 0$ and (a) $\rho_1 = 2/3$, $\rho_2 = 1/3$ and: $\mu_1 = 0$ (solid), $\pi/3$ (dashed), $2\pi/3$ (dot-dashed) and π (dotted), (b) $\mu_1 = \pi/2$, $\rho_2 = 1/3$ and: $\rho_1 = 0$ (solid), $1/4$ (dashed), $1/2$ (dot-dashed) and $3/4$ (dotted), (c) $\rho_2 = 1/3$ and: $(\mu_1, \rho_1) = (\pi, 1/3)$ (solid), $(0, 0)$ (dashed), $(0, 1/3)$ (dot-dashed) and $(0, 2/3)$ (dotted), and (d) $\rho_1 = \rho_2 = 0.5$ and: $\mu_1 = 0$ (solid), $\pi/3$ (dashed), $2\pi/3$ (dot-dashed) and π (dotted).

If $\phi_1 \neq \phi_2$, the integrand in (6) is holomorphic on the unit disc, except at two poles of order 1, that is, $z = \phi_j$, $j = 1, 2$. When $\phi_1 = \phi_2$, the integrand in (6) has a single pole of order 2 at $z = \phi_1$. Hence, from the residue theorem (e.g., Rudin [18], Theorem 10.42), we obtain Theorem 4. □

4.2. Mean direction and mean resultant length

Let Z be distributed as $EC^*(\phi_1, \phi_2)$. From Theorem 4 it is easy to see that the first trigonometric moment for Z is simply expressed as

$$E(Z) = \frac{(1 - |\phi_2|^2)\phi_1 + (1 - |\phi_1|^2)\phi_2}{1 - |\phi_1\phi_2|^2} \qquad \text{for any } \phi_1 \text{ and } \phi_2. \qquad (7)$$

The following result provides the condition under which the first trigonometric moment is equal to zero.

Corollary 1. *The necessary and sufficient condition for $E(Z) = 0$ is given by $\phi_1 = -\phi_2$.*

Proof. It is straightforward to show the sufficient condition for $E(Z) = 0$. The necessary condition is proved as follows. Let $E(Z) = 0$ and $\phi_1 \neq -\phi_2$. Then, from Theorem 4, the following equation holds between the two parameters: $\phi_1\phi_2 = (\phi_1 + \phi_2)/\overline{(\phi_1 + \phi_2)}$. Taking the absolute values of both sides of the above equation, we have $|\phi_1\phi_2| = 1$. Since $|\phi_j| < 1$ ($j = 1, 2$), there does not exist ϕ_1 and ϕ_2 such that $|\phi_1\phi_2| = 1$. Therefore, if $\phi_1 \neq -\phi_2$, then $E(Z) \neq 0$. $\square$

The mean direction is a measure of location which is defined as $\arg\{E(Z)\}$ if $E(Z) \neq 0$ and is undefined if $E(Z) = 0$. The mean resultant length, which is a measure of concentration, is defined by $|E(Z)|$. Figure 3 shows the mean direction and mean resultant length of a random variable having density (3) with fixed $\phi_2 = 0.5$. Figure 3(a) suggests that the mean direction
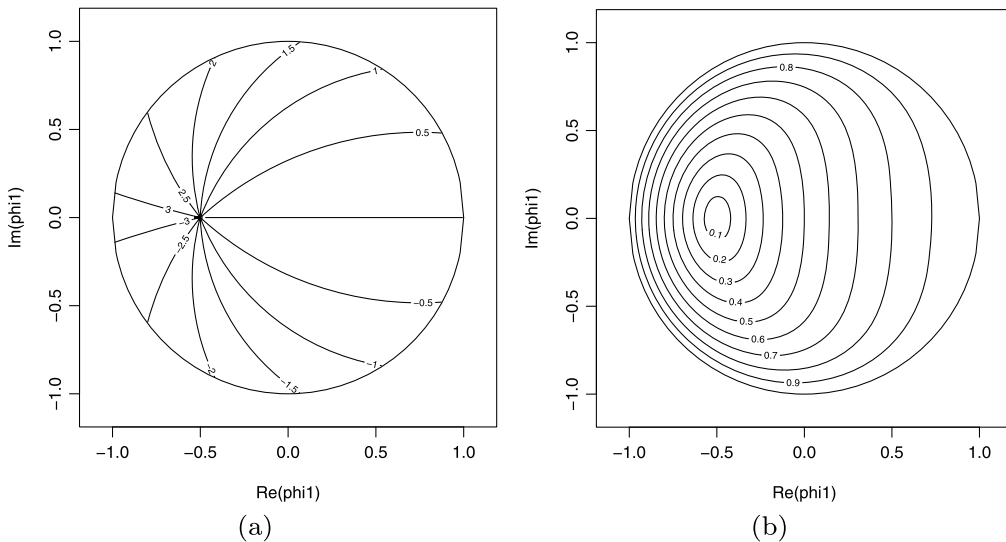
Figure 3. Plot of (a) mean direction and (b) mean resultant length of a random variable having the distribution $EC^*(\phi_1, 0.5)$, as a function of ϕ_1.

monotonically increases as $\arg(\phi_1)$ increases from $-\pi$ to π. As seen in Corollary 1, Figure 3(b) confirms that the mean resultant length is zero if $\phi_1 = -0.5$. It seems from this frame that unimodality holds for the mean resultant length as a function of ϕ_1.

4.3. Skewness

A measure of skewness for circular distributions (Mardia [13]), s, is defined by

$$s = E[\text{Im}\{(Ze^{-i\zeta})^2\}]/(1 - \delta)^{3/2}, \tag{8}$$

where ζ and δ are the mean direction and mean resultant length for a ∂D-valued random variable Z, respectively. For our model, it is possible to express the skewness for the proposed model in a fairly simple form as follows.

Corollary 2. *Let Z be distributed as $\text{EC}^*(\phi_1, \phi_2)$. Then the skewness for the distribution of Z is given by*

$$s = \frac{|1 - \phi_1\overline{\phi_2}|^2}{(1 - |\phi_1\overline{\phi_2}|^2)^3} \rho^{-2}(1 - \rho)^{-3/2} \text{Im}(\phi_1\overline{\phi_2})(|\phi_1|^2 - |\phi_2|^2)(1 - |\phi_1|^2)(1 - |\phi_2|^2), \tag{9}$$

where $\rho = |(1 - |\phi_2|^2)\phi_1 + (1 - |\phi_1|^2)\phi_2|/(1 - |\phi_1\overline{\phi_2}|^2)$.

Proof. Note that the numerator of (8) can be expressed as

$$E[\text{Im}\{(Ze^{-i\mu})^2\}] = \text{Im}[E(Z^2)\{\overline{E(Z)}/|E(Z)|\}^2].$$

Then, by using Theorem 4 and equation (7), the skewness (9) can be obtained after lengthy, but straightforward, calculations. $\qquad\square$

Now we write $s = s(\phi_1, \phi_2)$ in order to view the skewness (9) as a function of ϕ_1 and ϕ_2. It follows from Corollary 2 and Theorem 3 that the following properties hold for s.

Theorem 5. *The following properties hold for the skewness (9):*

1. $s(\phi_1, \phi_2) = 0 \Longleftrightarrow$ *density (3) is symmetric.*
2. $s(\phi_1, \phi_2) > 0 \ (< 0) \Longleftrightarrow \text{Im}(\phi_1\overline{\phi_2})(|\phi_1| - |\phi_2|) > 0 \ (< 0)$.
3. $s(\phi_1, \phi_2) = -s(\phi_1, \overline{\phi_2}\phi_1^2/|\phi_1|^2) = -s(\overline{\phi_1}\phi_2^2/|\phi_2|^2, \phi_2), \phi_1, \phi_2 \neq 0$.
4. $s(\phi_2, \phi_1) = s(\phi_1, \phi_2)$.
5. $s(\overline{\phi_1}, \overline{\phi_2}) = -s(\phi_1, \phi_2)$.
6. $s(\alpha\phi_1, \alpha\phi_2) = s(\phi_1, \phi_2), \alpha \in \partial D$.
7. *Assume $\rho_1, \rho_2 > 0$ and $\sin(\mu_1 - \mu_2) > 0 \ (< 0)$. Then $\lim_{\rho_1 \to 1} s(\rho_1 e^{i\mu_1}, \rho_2 e^{i\mu_2}) = \infty \ (-\infty)$.*

Figure 4 plots this skewness when $\phi_2 = 0.5$. As the first property of Theorem 5 shows, the skewness is equal to zero if and only if $\text{Im}(\phi_1\overline{\phi_2}) = 0$, $|\phi_1| = |\phi_2|$ or $|\phi_j| = 0 \ (j = 1, 2)$.
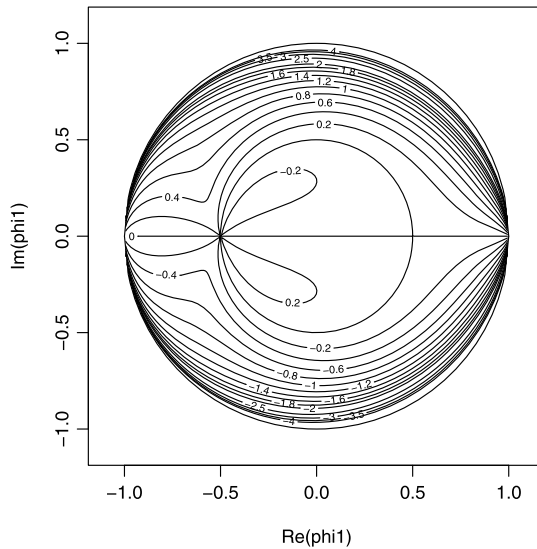
Figure 4. Skewness (9) of a random variable having the distribution EC*$(\phi_1, 0.5)$, as a function of ϕ_1.

This figure also confirms the second property of Theorem 5 that the model is positively (negatively) skewed if $|\phi_1| > |\phi_2|$ and $\text{Im}(\phi_1\overline{\phi_2}) > 0$ ($\text{Im}(\phi_1\overline{\phi_2}) < 0$), or $|\phi_1| < |\phi_2|$ and $\text{Im}(\phi_1\overline{\phi_2}) < 0$ ($\text{Im}(\phi_1\overline{\phi_2}) > 0$). It can also be seen that the fifth and seventh properties of Theorem 5 hold in this figure.

5. Some other topics

5.1. Alternative derivations

As discussed in Section 2.1, the proposed model can be derived by considering a problem in Brownian motion. Apart from the derivation given there, here is another method to generate our family via this stochastic process, kindly suggested by the referee.

Remember that, given a $\mathbb{C}$-valued Brownian path without drift starting from ϕ ($\in D$), the point of first exit from the unit disc D has a wrapped Cauchy distribution with density proportional to $1/|z - \phi|^2$. (Although we assume $\phi \in D$ by convention, the Brownian motion starting at $1/\overline{\phi}$ has the same hitting distribution from the exterior because of the conformal invariance of Brownian motion.) From this it follows that another derivation of the proposed model (3) is given in the following theorem.

Theorem 6. *Let Z_1 and Z_2 be the points of first exit from the unit disc for two independent $\mathbb{C}$-valued Brownian motions without drift starting from ϕ_1 and ϕ_2, respectively, where $\phi_1, \phi_2 \in D$. Then the conditional distribution of Z_1 ($\equiv Z$) given $Z_1 = Z_2$ is given by density (3).*

Remark 1. The referee embedded the above in a generalized formulation leading to densities of the form

$$f(z) \propto \prod_{j=1}^{d} \frac{1 - |\phi_j|^2}{|z - \phi_j|^2}, \qquad z \in \partial D; \phi_1, \ldots, \phi_d \in D, \tag{10}$$

many of whose properties are also readily obtainable using the residue theorem.

In addition to the derivations using Brownian motion, there are other methods to generate our family. The following result shows that the proposed model appears in a Bayesian context.

Theorem 7. *Let $\Theta|\nu$ have the wrapped Cauchy distribution $\mathrm{WC}(\nu, \rho_1)$ with known ρ_1. Assume that the prior distribution of ν is distributed as the wrapped Cauchy $\mathrm{WC}(\mu_2, \rho_2)$. Then the posterior distribution of ν given $\Theta = \mu_1$ has density* (2).

The above derivation enables us to generate random samples from our model using the Markov chain Monte Carlo method; see Kato and Jones [10], Section 5.3, for details.

Remark 2. In Theorem 7, the marginal distribution of Θ is given by the wrapped Cauchy distribution, $\mathrm{WC}(\mu_1 + \mu_2, \rho_1\rho_2)$.

Before we discuss a third method to derive our model, here we briefly recall a known property of the wrapped Cauchy distribution or the Poisson kernel as follows.

Theorem 8 (Rudin [18], Theorem 11.9). *Assume $\overline{D}$ is the closed unit disc in the complex plane. Suppose a function u is continuous on $\overline{D}$ and harmonic in D, and suppose that $|u(z)| < \infty$ for any $z \in \overline{D}$. Then*

$$\int_{\partial D} u(z) \frac{1}{2\pi} \frac{1 - |\phi_1|^2}{|z - \phi_1|^2} \, \mathrm{d}z = u(\phi_1), \qquad \phi_1 \in D.$$

Note that the integrand in the left-hand side of the above equation is actually the product of $u(z)$ and the density of the wrapped Cauchy $C^*(\phi_1)$. From the fact that

$$g(z) = \frac{|\phi_2^{-1}|^2 - |z|^2}{|z - \overline{\phi_2}^{-1}|^2}, \qquad \phi_2 \in D,$$

is a continuous real function on $\overline{D}$ and harmonic in D, it follows that

$$f(z) = \frac{g(z)}{2\pi g(\phi_1)} \frac{1 - |\phi_1|^2}{|z - \phi_1|^2}, \qquad z \in \partial D,$$

satisfies the definition of a density. Clearly, $f(z)$ is equal to density (3).

The following theorem implies that our model appears as a special case of a further general family of distributions. The proof is straightforward and omitted.

Theorem 9. *Let f and g be probability density functions on the circle. Assume that h is the convolution of f and g, namely, $h(\tau) = (f * g)(\tau) = \int_{-\pi}^{\pi} f(\theta)g(\tau - \theta)\,d\theta$. Then a function $k(\theta) = f(\theta)g(\tau - \theta)/h(\tau)$ is a probability density function on the circle.*

Density (2) can be derived on setting $\tau = \mu_2$ and f and g as the densities of WC(μ_1, ρ_1) and WC$(0, \rho_2)$, respectively. In this case $h(\tau)$ also has the form of the wrapped Cauchy density WC$(\mu_1, \rho_1\rho_2)$ since the wrapped Cauchy distribution has the additive property.

5.2. Möbius transformation

The Möbius transformation is well known in complex analysis as a conformal mapping which projects the unit disc onto itself. It is defined as

$$\mathcal{M}(w) = \alpha \frac{w + \beta}{\overline{\beta}w + 1}, \qquad w \in \overline{D}; \alpha \in \partial D, \beta \in D.$$

Although this projection is usually defined on the interior of the unit disc, here we extend the domain of the mapping so that the boundary of the unit disc, that is, the unit circle, is included. It is easy to see that the unit circle is mapped onto itself via the transformation, namely, $\mathcal{M}(\partial D) = \partial D$. In directional statistics the transformation appears in some papers such as McCullagh [15], Jones [6] and Kato and Jones [9].

Here we study the conformal invariance properties of our family. The distribution (3) has a density that is a relative invariant of weight 2 in the sense that

$$f(z; \phi_1, \phi_2) = f\{\mathcal{M}(z); \mathcal{M}(\phi_1), \mathcal{M}(\phi_2)\}|\mathcal{M}'(z)|^2.$$

Similarly, the extended family (10) has a density that is relatively invariant with weight k, in the sense that $f(z; \phi_1, \dots, \phi_k) = f\{\mathcal{M}(z); \mathcal{M}(\phi_1), \dots, \mathcal{M}(\phi_k)\}|\mathcal{M}'(z)|^k$. The relative invariance with weight 1 of the wrapped Cauchy distribution can be obtained by putting $k = 1$.

The proposed family (3) is not closed under the Möbius transformation except for the wrapped Cauchy special cases. To say the same thing in a different way, if Z follows the distribution (3), then the density for $\mathcal{M}(Z)$ is not of the form (3) except for $\phi_1 = 0$ or $\phi_2 = 0$. This fact can be understood clearly in the following context; as seen in Theorem 6, our model (3) can be derived as the conditional distribution of Z_1 given $Z_1 = Z_2$, where Z_1 and Z_2 are independent wrapped Cauchy variables. Invertibility of the transformations $\mathcal{M}$ implies that $Z_1 = Z_2$ if and only if $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$. However, as is known as the Borel paradox (e.g., Pollard [17], Section 5.5), the conditional distribution of Z_1 given $Z_1 = Z_2$ is generally not the same as the conditional distribution of Z_1 given $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$. Consequently, the Möbius transformation of the conditional distribution of Z_1 given $Z_1 = Z_2$, which is the Möbius transformation of model (3), is not the same as the conditional distribution of $\mathcal{M}(Z_1)$ given $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$, which is model (3). In a similar manner one can see that the extended model (10) with $k \geq 2$ is not closed under the Möbius transformation.

6. Symmetric cases

So far, we have mostly considered the full family of distributions with densities (2) and (3). In this section, we focus on the symmetric special cases of the proposed model. Some properties, which the general family does not have, hold for the symmetric cases. In Section 6.1, model (2) with $\mu_1 = \mu_2$ or $\mu_2 + \pi$ is discussed. In Section 6.2, we briefly consider another symmetric model, namely, model (2) with $\rho_1 = \rho_2$.

6.1. Symmetric case I: $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$

Model (2) with $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$ is essentially the same as the distribution with density

$$f(\theta) = \frac{1 - \rho_1 \rho_2}{2\pi(1 + \rho_1 \rho_2)} \frac{1 - \rho_1^2}{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu)} \frac{1 - \rho_2^2}{1 + \rho_2^2 - 2\rho_2 \cos(\theta - \mu)},$$
$$-\pi \leq \theta < \pi,$$

(11)

where $-1 < \rho_1 < 1$, $0 \leq \rho_2 < 1$, and $-\pi \leq \mu < \pi$. Note the extension in the range of ρ: this density corresponds to (2) with $\mu_1 = \mu_2$ when $\rho_1 \geq 0$ in (11), while the density is equivalent to (2) with $\mu_1 = \mu_2 + \pi$ when $\rho_1 < 0$ in (11). Clearly, the density is symmetric about $\theta = \mu$ and $\theta = \mu + \pi$.

One might, however, wish to restrict interest to the case where $0 \leq \rho_1 < 1$ because then density (11) is unimodal. To see this, note that in this case $a_0 = a_1 = a_4 = 0$ in (5) so that stationary points of the density occur at $\sin\theta = 0$ and, potentially, also at $\cos\theta = -a_2/a_3 = (\rho_1 + \rho_2)(1 + \rho_1 \rho_2)/4\rho_1 \rho_2$. However, the latter expression can easily be proved to not be less than unity, showing that (11) is unimodal. On the other hand, the symmetric model (11) with negative ρ_1 has a relationship with mixtures of two wrapped Cauchy distributions. The density for this submodel can be expressed as

$$f(\theta) = p \frac{1}{2\pi} \frac{1 - \rho_1^2}{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu)} + (1 - p) \frac{1}{2\pi} \frac{1 - \rho_2^2}{1 + \rho_2^2 - 2\rho_2 \cos(\theta - \mu)},$$

where $p = -\rho_1 (1 - \rho_2^2)/\{(1 + \rho_1 \rho_2)(\rho_2 - \rho_1)\}$. Inter alia, this representation allows random numbers from density (11) with $\rho_1 \leq 0$ to be more easily generated than they were for the general case in Kato and Jones [10], Section 5.3.

6.2. Symmetric case II: $\rho_1 = \rho_2$

In this section we discuss the submodel of density (2) with $\rho_1 = \rho_2 = \rho$ ($\in [0, 1)$). The density for this submodel reduces to

$$f(\theta) = \frac{(1 - \rho^2)\{1 + \rho^4 - 2\rho^2 \cos(\mu_1 - \mu_2)\}}{2\pi(1 + \rho^2)}$$
$$\times [\{1 + \rho^2 - 2\rho \cos(\theta - \mu_1)\}\{1 + \rho^2 - 2\rho \cos(\theta - \mu_2)\}]^{-1}, \qquad -\pi \leq \theta < \pi.$$

The model is symmetric about $\theta = \overline{\mu}$ and $\overline{\mu} + \pi$, where $\overline{\mu} = \arg(\sum_{j=1}^{2} \cos \mu_j + \mathrm{i} \sum_{j=1}^{2} \sin \mu_j)$. As displayed in Figure 2(d), this submodel can be unimodal or bimodal depending on the choice of the parameters. The condition for unimodality can be simplified for this submodel as follows:

$$-2 \arccos\left(\frac{2\rho}{1+\rho^2}\right) \leq \mu_1 - \mu_2 \leq 2 \arccos\left(\frac{2\rho}{1+\rho^2}\right).$$

This result was used in Section 3.2 to find the region of Figure 1(b) in which the discriminant takes negative values, corresponding to unimodality, when $\rho_1 = \rho_2$ and $\mu_1 - \mu_2 = 2\pi/3$.

7. A generalization on the sphere

As described in Section 2.1, the proposed model (2) can be derived using Brownian motion. By adopting a multi-dimensional Brownian motion instead of the two-dimensional one, we can extend model (2) to a distribution on the unit sphere. The generalized model is defined as follows.

Definition 2. *Let $\{B_t; t \geq 0\}$ be $\mathbb{R}^d$-valued Brownian motion starting at $B_0 = \rho_1 \eta_1$, where $d \geq 2$, $0 \leq \rho_1 < 1$, $\eta_1 \in S^d$ and $S^d = \{x \in \mathbb{R}^d; \|x\| = 1\}$. Assume $\tau_1 = \inf\{t; \|B_t\| = 1\}$ and $\tau_2 = \inf\{t; \|B_t\| = \rho_2^{-1}\}$, where $0 < \rho_2 < 1$. Then the proposed model is defined by the conditional distribution of B_{τ_1} given $B_{\tau_2} = \rho_2^{-1} \eta_2$, where $\eta_2 \in S^d$.*

For simplicity, write $X = B_{\tau_1}$. The probability density function of this extended model is available, and it is given in the following theorem.

Theorem 10. *The conditional distribution of X given $B_{\tau_2} = \rho_2^{-1} \eta_2$ is of the form*

$$f(x) = \frac{1}{A_{d-1}} \frac{(1 + \rho_1^2 \rho_2^2 - 2\rho_1 \rho_2 \eta_1' \eta_2)^{d/2}}{1 - \rho_1^2 \rho_2^2} \frac{1 - \rho_1^2}{\|x - \rho_1 \eta_1\|^d} \frac{1 - \rho_2^2}{\|x - \rho_2 \eta_2\|^d}, \qquad x \in S^d, \quad (12)$$

where A_{d-1} is the surface area of S^d, that is, $A_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$.

See Appendix C of Kato and Jones [10] for the proof. Note that density (12) reduces to the circular case (2) if $d = 2$, $x = (\cos\theta, \sin\theta)'$, $\eta_1 = (\cos\mu_1, \sin\mu_1)'$ and $\eta_2 = (\cos\mu_2, \sin\mu_2)'$. It might be appealing that density (12), which is not rotationally symmetric in general, can be expressed in a relatively simple form.

When $\rho_2 = 0$, the distribution becomes the so-called "exit" distribution on the sphere, whose density is given by

$$f(x) = \frac{1}{A_{d-1}} \frac{1 - \rho_1^2}{\|x - \rho_1 \eta_1\|^d}, \qquad x \in S^d. \quad (13)$$

We write $X \sim \mathrm{Exit}_d(\rho_1 \eta_1)$ if a random variable X has density (13). This model is rotationally symmetric about $x = \eta_1$. See, for example, Durrett [3], Section 1.10, for details about the distribution. In particular, when $d = 2$, model (13) becomes the wrapped Cauchy distribution. It is

noted that this distribution is a submodel of Jones and Pewsey's [7] family of distributions on the sphere with density

$$f(x) = \frac{1}{A_{d-1}} \frac{|\sinh(\kappa\psi)|^{d/2-1}}{2^{2/d-1}\Gamma(d/2)} \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi)x'\mu\}^{1/\psi}}{P_{1/\psi+d/2-1}^{1-d/2}\{\cosh(\kappa\psi)\}}, \qquad x \in S^d, \qquad (14)$$

where $P_\alpha^\beta(z)$ is the associated Legendre function of the first kind of degree α and order β (Gradshteyn and Ryzhik [5], Sections 8.7, 8.8). It follows from equation (8.711.1) of Gradshteyn and Ryzhik [5] and equation (2) of McCullagh [14] that density (14) reduces to density (13) if $\kappa = d\log\{(1+\rho_1)/(1-\rho_1)\}/2$, $\mu = \eta_1$ and $\psi = -2/d$. Also, the extended model (12) includes the model with density

$$f(x) = \frac{1}{A_{d-1}} \frac{(1-\rho_1^2)^{d+1}}{1+\rho_1^2} \frac{1}{\|x - \rho_1\eta_1\|^{2d}}, \qquad x \in S^d, \qquad (15)$$

when $\rho_1 = \rho_2$, $\eta_1 = \eta_2$. It can be seen that this model is another submodel of Jones and Pewsey's [7] family (14) by putting $\psi = -1/d$ and $\kappa = d\log\{(1+\rho_1)/(1-\rho_1)\}$. Also, notice that, if $d = 2$, density (15) corresponds to Case 2, (4), of circular submodels in Section 2.4. In addition to these two submodels, the multivariate model (12) contains the uniform distribution ($\rho_1 = \rho_2 = 0$), one-point distribution ($\rho_1 = 0$ and $\rho_2 \to 1$, or $\rho_2 = 0$ and $\rho_1 \to 1$), and two-point distribution ($\rho_1 = \rho_2$, $\mu_1 \neq \mu_2$ and $\rho_1 \to 1$).

In a similar manner to Theorem 7, it is easy to prove that model (12) can be derived from Bayesian analysis of the exit distribution.

Theorem 11. *Let $X|\xi$ be distributed as the exit distribution* $\mathrm{Exit}(\rho_1\xi)$ *with known ρ_1. Suppose that the prior distribution of ξ is* $\mathrm{Exit}(\rho_2\eta_2)$. *Then the posterior distribution of ξ given $X = \eta_1$ is given by density (12).*

From this result and the random number generator of the exit distribution given by Kato [8], Section 3.3, it is possible to generate random samples from (12) through the Markov chain Monte Carlo method in a similar manner to Kato and Jones [10], Section 5.3.

Acknowledgements

References

[1] Abramowitz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*. New York: Dover Press.

[2] Batschelet, E. (1981). *Circular Statistics in Biology*. London: Academic Press. MR0659065

[3] Durrett, R. (1984). *Brownian Motion and Martingales in Analysis*. *Wadsworth Mathematics Series*. Belmont, CA: Wadsworth International Group. MR0750829

[4] Gatto, R. and Jammalamadaka, S.R. (2007). The generalized von Mises distribution. *Stat. Methodol.* **4** 341–353. MR2380560

[5] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals*, *Series*, *and Products*, 7th ed. Amsterdam: Elsevier/Academic Press. Translated from the Russian. Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger. MR2360010

[6] Jones, M.C. (2004). The Möbius distribution on the disc. *Ann. Inst. Statist. Math.* **56** 733–742. MR2126808

[7] Jones, M.C. and Pewsey, A. (2005). A family of symmetric distributions on the circle. *J. Amer. Statist. Assoc.* **100** 1422–1428. MR2236452

[8] Kato, S. (2009). A distribution for a pair of unit vectors generated by Brownian motion. *Bernoulli* **15** 898–921. MR2555204

[9] Kato, S. and Jones, M.C. (2010). A family of distributions on the circle with links to, and applications arising from, Möbius transformation. *J. Amer. Statist. Assoc.* **105** 249–262. MR2656051

[10] Kato, S. and Jones, M.C. (2011). An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion. Technical Report 11/02, Statistics Group, The Open Univ. Available at http://statistics.open.ac.uk.

[11] Kent, J.T. and Tyler, D.E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Statist.* **15** 247–254.

[12] Maksimov, V.M. (1967). Necessary and sufficient statistics for a family of shifts of probability distributions on continuous bicompact groups. *Teor. Verojatnost. i Primenen.* **12** 307–321 (in Russian). English translation: *Theor. Probab. Appl.* **12** 267–280. MR0214175

[13] Mardia, K.V. (1972). *Statistics of Directional Data*. London: Academic Press. MR0336854

[14] McCullagh, P. (1989). Some statistical properties of a family of continuous univariate distributions. *J. Amer. Statist. Assoc.* **84** 125–129. MR0999670

[15] McCullagh, P. (1996). Möbius transformation and Cauchy parameter estimation. *Ann. Statist.* **24** 787–808. MR1394988

[16] Pewsey, A. (2008). The wrapped stable family of distributions as a flexible model for circular data. *Comput. Statist. Data Anal.* **52** 1516–1523. MR2422752

[17] Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability*. *Cambridge Series in Statistical and Probabilistic Mathematics* **8**. Cambridge: Cambridge Univ. Press. MR1873379

[18] Rudin, W. (1987). *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill. MR0924157

[19] Uspensky, J.V. (1948). *Theory of Equations*. New York: McGraw-Hill.

[20] Yfantis, E.A. and Borgman, L.E. (1982). An extension of the von Mises distribution. *Comm. Statist. Theory Methods* **11** 1695–1706.