Symmetric unimodal models for directional data motivated by inverse stereographic projection

Toshihiro Abe[1], Kunio Shimizu[2] and Arthur Pewsey[3]

[1] Keio University, Japan, [2] Keio University, Japan, [3] University of Extremadura, Spain.

1. Introduction

Minh and Farnum (2003) used inverse stereographic projection, or equivalently bilinear (Möbius) transformation, of distributions defined on the real line to induce distributions on the circle. Their construction based on this approach proceeds as follows. Let $f(x)$ denote the probability density function (pdf) of the t-distribution on the real line with m (a positive integer) degrees of freedom, i.e.

$$f(x) = \frac{\Gamma((m+1)/2)}{\sqrt{\pi m}\,\Gamma(m/2)(1+x^2/m)^{(m+1)/2}}, \quad -\infty < x < \infty. \tag{1}$$

Then, applying inverse stereographic projection, defined by the (generally) one-to-one mapping,

$$x = u + v\frac{\sin\theta}{1+\cos\theta} = u + v\tan\left(\frac{\theta}{2}\right), \quad -\pi \le \theta < \pi, \tag{2}$$

with $u = 0$ and $v = \sqrt{m}$, and writing $m = 2n+1$, for $n = 0, 1, \ldots$, leads to the pdf on the circle of radius v,

$$f(\theta) = \frac{\Gamma(n+1)}{2^{n+1}\sqrt{\pi}\Gamma(n+1/2)}(1+\cos\theta)^n. \tag{3}$$

Given its construction, the distribution with this density might be referred to as a type of t-distribution on the circle, or the circular t-distribution induced by inverse stereographic projection if one were being more precise. However, the distribution with density (3) already has a name in the literature, namely Cartwright's power-of-cosine distribution. When $n = 0$ (i.e. $m = 1$), the circular uniform distribution is obtained, clearly induced by inverse stereographic projection of the Cauchy distribution on the real line.

The three-parameter family of symmetric circular distributions proposed by Jones and Pewsey (2005) has density

$$f(\theta) = \frac{(\cosh(\kappa\psi))^{1/\psi}(1+\tanh(\kappa\psi)\cos(\theta-\mu))^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}, \tag{4}$$

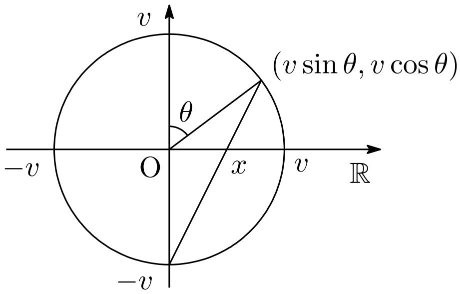
where $-\pi \leq \mu < \pi$ is a location parameter (equal, in general, to the mean/modal/median direction), $\kappa \geq 0$ is a concentration parameter, $\psi \in \mathbb{R}$ is a shape index, and $P_{1/\psi}$ is the associated Legendre function of the first kind of degree $1/\psi$ and order 0.

As special cases, Jones–Pewsey's distribution includes von Mises ($\psi \to 0$), cardioid ($\psi = 1$), wrapped Cauchy ($\psi = -1$) as well as Minh and Farnum's (2003) (or, equivalently, Cartwright's power-of-cosine) ($\psi > 0, \kappa \to \infty$), and is can be seen to be closely related to the pdf of the t-distribution on the circle proposed by Shimizu and Iida (2002).

The distribution with density (3) has no continuity, in the sense that (1) tends to the standard normal pdf as $m \to \infty$ but a degenerate distribution concentrated at the point $\theta = 0$ is obtained when $n \to \infty$ in (3).

In order to resolve this problem we employ the transformation (2) with $u = 0$ and $v > 0$ independent of n.

$x = v \tan\left(\dfrac{\theta}{2}\right), \quad \theta \in [-\pi, \pi).$



Now inverse stereographic projection of the t-distribution on the real line yields a different circular t-distribution from that given in (3), with density

$$f(\theta) = \frac{v}{2\sqrt{m}B(m/2, 1/2)} \times \frac{(1 + \tan^2(\theta/2))}{(1 + v^2\tan^2(\theta/2)/m)^{(m+1)/2}}, \tag{5}$$

where $m = 2n+1$ and $B(p, q)$ denotes the beta function, defined by $B(p, q) = \int_0^1 t^{p-1}(1-t)^{q-1}dt$. In this paper, we shall refer to the distribution with density (5) as the *modified Minh–Farnum distribution*. As $m \to \infty$, the pdf in (5) tends to

$$f(\theta) = \frac{v}{2\sqrt{2\pi}}\left(1 + \tan^2\left(\frac{\theta}{2}\right)\right)\exp\left(-\frac{v^2}{2}\tan^2\left(\frac{\theta}{2}\right)\right). \tag{6}$$

This is the density of the distribution obtained when the modified inverse stereographic projection described above is applied to the standard normal distribution on the real line. Thus, the transformation (2) with $v > 0$ independent of n has continuity from the t to the normal distributions on the circle.

The density (6) is unimodal if $v \geq \sqrt{2}$, and is bimodal if $v < \sqrt{2}$. An alternative representation of it is

$$f(\theta) = \frac{v_0}{B(n+1/2, 1/2)(1 + v_0{}^2)^{n+1}} \times \frac{(1 + \cos\theta)^n}{(1 - z\cos\theta)^{n+1}}, \tag{7}$$

where $v_0 = v/\sqrt{2n+1}$ and $z = ({v_0}^2-1)/({v_0}^2+1)$. Finally, we note that the Jones–Pewsey density tends to that of a Minh–Farnum distribution when $\kappa \to \infty$.

2. A new extended family of unimodal symmetric circular distributions

In this section we introduce the new family of unimodal symmetric distributions on the circle which extends both the Jones–Pewsey family of distributions as well as the unimodal distributions contained within the modified Minh–Farnum family. Its density results on combining the defining structure of the pdf in (7), derived by inverse stereographic projection, with the transformation technique based on the hyperbolic tangent function employed in the construction of Jones and Pewsey (2005). The resulting density is given by

$$f(\theta) = C_\psi(\kappa_1, \kappa_2) \frac{(1+\tanh(\kappa_1\psi)\cos(\theta-\mu))^{1/\psi}}{(1-\tanh(\kappa_2\psi)\cos(\theta-\mu))^{1/\psi+1}}, \tag{8}$$

where $-\pi \le \mu < \pi$ is a location parameter (which, in general, corresponds to the distribution's mean/modal/median direction), κ_1, $\kappa_2 \ge 0$

are concentration parameters and $\psi > 0$ is an index which also effects the peakedness and tails of the distribution. The overall shape of the distribution is determined by the last three of these parameters. The normalizing constant $C_\psi(\kappa_1, \kappa_2)$ is the reciprocal of the integral

$$\begin{aligned} &\int_{-\pi}^{\pi} \frac{(1+\tanh(\kappa_1\psi)\cos\theta)^{1/\psi}}{(1-\tanh(\kappa_2\psi)\cos\theta)^{1/\psi+1}} d\theta \\ &= \frac{2(1+\tanh(\kappa_1\psi))^{1/\psi}}{(1-\tanh(\kappa_2\psi))^{1/\psi+1}} \int_0^1 \frac{(1-z_1t)^{1/\psi}}{(1+z_2t)^{1/\psi+1}} t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}} dt, \end{aligned} \tag{9}$$

where $z_1 = 2\tanh(\kappa_1\psi)/(1+\tanh(\kappa_1\psi))$ and $z_2 = -2\tanh(\kappa_2\psi)/(1-\tanh(\kappa_2\psi))$. Here, the transformation $\cos\theta = 1-2t$ has been used to obtain the right-hand side of (9). The last integral is expressible in terms of Appell's function, F_1, and the Gauss hypergeometric function, ${}_2F_1$, as

$$\begin{aligned} \int_0^1 \frac{(1-z_1t)^{1/\psi}}{(1+z_2t)^{1/\psi+1}} t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}} dt &= \pi F_1\left(\frac{1}{2}; -\frac{1}{\psi}, 1+\frac{1}{\psi}; 1; z_1, -z_2\right) \\ &= \pi(1+z_2)^{-1/2} {}_2F_1\left(\frac{1}{2}, -\frac{1}{\psi}; 1; \frac{z_1+z_2}{1+z_2}\right), \end{aligned}$$

due to the relation (Appell and Kampé de Fériet, 1926, Eq. (28), p. 24) $F_1(a; b_1, b_2; b_1+b_2; z_1, z_2) = (1-z_2)^{-a} {}_2F_1(a, b_1; b_1+b_2; (z_1-z_2)/(1-z_2))$, where

$$F_1(a; b_1, b_2; c; z_1, z_2) = \frac{1}{B(a, c-a)} \int_0^1 \frac{t^{a-1}(1-t)^{c-a-1}}{(1-z_1t)^{b_1}(1-z_2t)^{b_2}} dt.$$

Note how, as ψ and κ_2 increase, κ_1 has increasingly less effect on the shape of the density.

3. Basic properties

3.1 Unimodality

Given that the two concentration parameters are constrained so that κ_1, $\kappa_2 \geq 0$, it follows that distributions with density (8) are always unimodal with $f(\mu) > f(\mu \pm \pi)$, apart from the special case of the circular uniform distribution which, formally, has no mode. The circular uniform distribution is obtained when $\kappa_1 = \kappa_2 = 0$ or $\psi \to \infty$ and κ_1 and κ_2 are finite.

To show that, for all other cases, distributions from the proposed family are unimodal with mode at μ, we set, without loss of generality, μ equal to 0 and consider the derivative of (8) with respect to θ, i.e.

$$\frac{d}{d\theta}\left(\frac{(1+\tanh(\kappa_1\psi)\cos\theta)^{1/\psi}}{(1-\tanh(\kappa_2\psi)\cos\theta)^{1/\psi+1}}\right) = -\frac{2\sin\theta(1+\tanh(\kappa_1\psi)\cos\theta)^{1/\psi-1}}{(1-\tanh(\kappa_2\psi)\cos\theta)^{1/\psi+2}}$$
$$\times \tanh(\kappa_1\psi)\tanh(\kappa_2\psi)\left(\cos\theta + \frac{(1+\psi)\tanh(\kappa_2\psi)+\tanh(\kappa_1\psi)}{\psi\tanh(\kappa_1\psi)\tanh(\kappa_2\psi)}\right).$$

It follows from the constraints on the parameters that the density is unimodal if $|[(1+\psi)\tanh(\kappa_2\psi)+\tanh(\kappa_1\psi)]/[\psi\tanh(\kappa_1\psi)\tanh(\kappa_2\psi)]| > 1$. Distributions with density (8) are either circular uniform or, more generally, unimodal with mode at μ. The fact that they are also symmetric about μ follows directly from (8).

3.2 Special cases

Case 1: $\kappa_2 = 0$, the Jones–Pewsey distribution with positive power $1/\psi$.
Case 2: $\kappa_1 = 0$, the Jones–Pewsey distribution with negative power $-(1 + 1/\psi)$.
Case 3: $\kappa_1 \to \infty$, the unimodal distributions contained in the modified Minh–Farnum family with density (5).
Case 4: $\kappa_2 \to +\infty$, a new distribution with a pole at 0.

3.3 Trigonometric moments

The trigonometric moments of the distribution are given by $\{\phi_m : m = 1, 2, 3, ...\}$, where $\phi_m = \alpha_m + i\beta_m$, with $\alpha_m = E(\cos m\Theta)$ and $\beta_m = E(\sin m\Theta)$ being the mth order cosine and sine moments of the random angle Θ, respectively. Because the distribution is symmetric about $\mu = 0$, it follows that the sine moments are all zero. Thus, $\phi_m = \alpha_m$. To obtain the cosine moments, we will make use of the relationships that exist between them and the (standard) moments of $\cos\Theta$. The latter can be expressed as integral expressions involving Appell's function, F_1, or the Gauss hypergeometric function, ${}_2F_1$. Specifically, in terms of Appell's function,

$$
\begin{aligned}
& E(\cos^m \Theta) = 2C_\psi(\kappa_1,\kappa_2)\int_0^\pi \frac{(\cos\theta)^m(1+\tanh(\kappa_1\psi)\cos\theta)^{1/\psi}}{(1-\tanh(\kappa_2\psi)\cos\theta)^{1/\psi+1}}d\theta \\
= \; & \frac{1}{\pi F_1\left(1/2,-1/\psi,1/\psi+1;1,z_1,-z_2\right)}\sum_{k=0}^{m} {}_mC_k(-2)^k \\
& \times F_1\left(k+\frac{1}{2},-\frac{1}{\psi},1+\frac{1}{\psi};k+1,z_1,-z_2\right)B\left(k+\frac{1}{2},\frac{1}{2}\right),
\end{aligned}
$$

for all $m = 0, 1, 2, \ldots$, and, in terms of the Gauss hypergeometric function,

$$
\begin{aligned}
& E(\cos^m \Theta) \\
= \; & \frac{2C_\psi(\kappa_1,\kappa_2)(1+\tanh(\kappa_1\psi))^{1/\psi}}{(1-\tanh(\kappa_2\psi))^{1/\psi+1}}\sum_{n=0}^{\infty}\sum_{k=0}^{m}(-1/\psi)_n\frac{(-2)^k z_1^n}{n!}{}_mC_k \\
& \times {}_2F_1\left(1+\frac{1}{\psi},n+k+\frac{1}{2};n+k+1;-z_2\right)B\left(n+k+\frac{1}{2},\frac{1}{2}\right).
\end{aligned}
$$

Using these expressions, the cosine moments of Θ can be obtained.

3.4 Simulation

Lacking any obvious direct construction which leads to the distribution with density (8), here we can present two approaches to simulating pseudo-random variates from it based on the more generally applicable *inversion* and *acceptance-rejection* methods.

4. Parameter estimation (Maximum likelihood)

Using the results presented in Section 2, the log-likelihood function can be expressed as

$$
\begin{aligned}
& l(\mu,\kappa_1,\kappa_2,\psi) \\
= \; & -n\log(2\pi) - n\log(1+\tanh(\kappa_1\psi))/\psi + n\log(1+\tanh(\kappa_2\psi))/2 \\
& -n\log\left({}_2F_1\left(1/2,-1/\psi;1;\frac{2(\tanh(\kappa_1\psi)+\tanh(\kappa_2\psi))}{(1+\tanh(\kappa_1\psi))(1+\tanh(\kappa_2\psi))}\right)\right) \\
& +n(1/\psi+1/2)\log(1-\tanh(\kappa_2\psi)) \\
& +(1/\psi)\sum_{i=1}^{n}\log(1+\tanh(\kappa_1\psi)\cos(\theta_i-\mu)) \\
& -(1+1/\psi)\sum_{i=1}^{n}\log(1-\tanh(\kappa_2\psi)\cos(\theta_i-\mu)).
\end{aligned}
$$

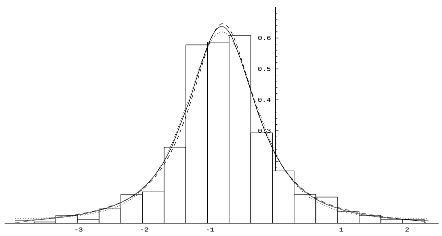
Maximum likelihood (ML) point estimation then reduces to the constrained maximization over the points in the parameter space.

5. An example

As our illustrative example, we present an analysis of the same grouped data set of $n = 714$ observations as was considered in Jones and Pewsey (2005). The data consist of the 'vanishing angles' of non-migratory British mallard ducks taken from Table 1.1 of Mardia and Jupp (1999). As reported in Jones and Pewsey (2005), the test of Pewsey (2002) provides no evidence against circular reflective symmetry for these data (p-value=0.124).

The results obtained from fitting the new family of distributions and the Jones–Pewsey and modified Minh–Farnum submodels are presented in the table below. A histogram of the data together with the densities for the three fits are presented in Figure below. From the latter it can be seen that the density of the fit for the full family is bounded above and below by the densities of the fits for the two limiting submodels. Moreover, there is evidence, particularly around the mode, that the fit for the full family is more similar to the fit for the modified Minh–Farnum distribution than that for the Jones–Pewsey. This visual impression is confirmed by the results for the maximum values of the log-likelihood presented in the table below. The maximum likelihood solution for the full family is an interior point of the parameter space, and so applying standard asymptotic distribution theory for the likelihood ratio test, there is some evidence that it offers a significant improvement in fit over the Jones–Pewsey family (p-value $= 0.070$), but not over the modified Minh–Farnum family (p-value $= 0.123$).

Histogram of the vanishing angles of 714 British mallard ducks together with the maximum likelihood fits for the proposed model (solid) and the modified Minh–Farnum (dashed) and Jones–Pewsey (dotted) distributions. The data and densities are plotted on approximately $(\hat{\mu} - \pi, \hat{\mu} + \pi)$.



Distribution	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\mu}$	$\hat{\psi}$	*MLL*	*AIC*	*BIC*
Jones–Pewsey	0	1.19	−0.82	0.54	−869.36	1744.72	1747.28
Proposed Model	0.74	0.42	−0.82	2.35	−867.72	1743.44	1746.85
Modified Minh–Farnum	∞	0.35	−0.80	3.05	−868.91	1743.82	1746.38

Focusing on the results for the two model selection criteria included in the Table, the AIC identifies the full family as providing a slightly better fit than the modified Minh–Farnum submodel. The BIC penalizes parameter-heavy models more and the results based on it reverse the order of preference of these two models. Under both criteria, the fit for the Jones–Pewsey submodel is clearly identified as the worst of the three fits.

Chi-squared goodness-of-fit tests at the 5% significance level marginally rejected the fit of the Jones–Pewsey submodel (p-value $= 0.047$), but did not reject the fit of the modified Minh–Farnum submodel (p-value $= 0.120$)

nor that of the full family (p-value $= 0.082$). Again, these p-values provide more evidence of the superior fit of the modified Minh–Farnum submodel for these data.

Finally, approximate 95% confidence intervals for μ, κ_1, κ_2 and ψ, obtained from their respective profile log-likelihood functions, are $(-0.87, -0.77)$, $(0, \infty)$, $(-1.35, -0.28)$ and $(2.35, 2.40)$. Of course, the interpretation of these confidence intervals is hampered by the fact that they only provide information about each individual parameter, not their relations with the others. Nevertheless, the intervals for μ, κ_2 and ψ give fairly tight information regarding the probable values for these three parameters. The interval for κ_1 simply indicates that models across the whole range from the Jones–Pewsey to the modified Minh–Farnum are potential generators of the data. Obviously, confidence regions for different combinations of the parameters would provide insight as to the probable joint values of the parameters.

References

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

[2] Minh, D.L.P. and Farnum, N.R. (2003), “Using bilinear transformations to induce probability distributions,” Communications in Statistics–Theory and Methods, 32, 1–9.

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