THE MÖBIUS DISTRIBUTION ON THE DISC

M. C. Jones

Department of Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, U.K.,
e-mail: m.c.jones@open.ac.uk

Abstract. A simple new family of distributions is proposed which has support the unit disc in two dimensions. The density functions of the family are unimodal, monotonic or uniantimodal. The bivariate symmetric beta distributions, which include the uniform distribution, are special cases, but many members of the family are skew. The distributions have three parameters, one controlling orientation, one controlling degree of concentration and the third controlling skewness, or more precisely off-centredness. Importantly, these parameters are globally orthogonal. An illustrative example of fitting the model to data is given. Conditional and marginal distributions are considered. The new distributions are compared favourably with an earlier suggestion of the same author.

Key words and phrases: Beta distribution, bivariate distribution, circular law, Möbius transformation, Pearson type II distribution.

1. Introduction

How many probability distributions can you think of with support the unit disc? The uniform distribution probably comes to mind immediately. The only other distributions with support the unit disc that have any degree of familiarity are the bivariate spherically symmetric beta (or Pearson type II) distributions with density

$$(1.1) \qquad f_\gamma(x,y) = \frac{\gamma}{\pi}(1 - x^2 - y^2)^{\gamma-1}, \qquad \gamma > 0, \qquad 0 \le x^2 + y^2 \le 1.$$

These distributions include the uniform distribution on the disc (sometimes called the 'circular law') when $\gamma = 1$ and are all spherically symmetric with mode at zero when $\gamma > 1$ and antimode at zero when $\gamma < 1$. Marginals are symmetric beta distributions on $[-1, 1]$ with parameter $\gamma + 1/2$. See Kotz (1975), Johnson (1987) and Fang *et al.* (1990) for details.

But these distributions are all spherically symmetric. How about distributions that can also exhibit skewness, i.e., that do not exhibit bivariate symmetry on the disc? The only specific proposal that I know of was made by Jones ((2002), Section 3). But now I can do better! In this paper, I introduce an alternative simple family of distributions on the disc which has many skew members and which also incorporates the beta distribution (1.1) as its spherically symmetric special case. In addition, all density functions in the new family remain unimodal, monotone or uniantimodal. Other advantages include meaningful parameters that are readily estimated by maximum likelihood and which prove to exhibit parameter orthogonality.

The new distribution, which I call the Möbius distribution, and its parameters are described in Section 2; an outline derivation of its density is deferred to Appendix A. Modality properties of the distribution are described in Section 3 and proved in Appendix B. Some plots of the density are presented in Section 4. Section 5 deals with parameter orthogonality. In Section 6, an illustrative example, fitting the model to data of Johnson and Wehrly (1977), is given. Conditional and marginal distributions are considered in Section 7 and the paper closes with Section 8 which gives a brief comparison of the Möbius distribution and that of Jones (2002).

2. The Möbius distribution

The unit disc is most naturally parametrised in terms of polar coordinates or, equivalently, complex numbers. (That said, I work explicitly in terms of complex numbers only in this paragraph and a little in Appendix A!) So, write $z = (x, y)$ and $w = (u, v)$ as complex numbers with, in particular, $w = re^{i\theta}$, $0 \leq r \leq 1$, $-\pi < \theta \leq \pi$. The Möbius transformation is defined as

$$z = M_c(w) = \frac{w - c}{1 - \bar{c}w} \quad \text{with inverse} \quad w = M_{-c}(z) = \frac{z + c}{1 + \bar{c}z}$$

where $c = ae^{i\mu}$, $0 \leq a < 1$ and the bar denotes complex conjugate. The Möbius transformation is essentially the only conformal mapping from the unit disc to itself (e.g., Krantz (1999), Section 6.2.2). It is natural, therefore, to apply either the Möbius transformation or its inverse to a random variable with the symmetric beta distribution to obtain the new class of Möbius distributions. I choose to apply the inverse Möbius transformation to z, so that then the Möbius distribution can be shown to have density

$$(2.1) \qquad f(r, \theta) = f_{\gamma, a, \mu}(r, \theta) = \frac{\gamma(1 - a^2)^{\gamma+1} r(1 - r^2)^{\gamma-1}}{\pi(1 - 2ar\cos(\theta - \mu) + a^2 r^2)^{\gamma+1}}$$

with parameters $\gamma > 0$, $0 \leq a < 1$ and $-\pi < \mu \leq \pi$ on $0 \leq r \leq 1$, $-\pi < \theta \leq \pi$ in polar coordinates, or equivalently

$$(2.2) \qquad f_{\gamma, a, \mu}(u, v) = \frac{\gamma(1 - a^2)^{\gamma+1}(1 - u^2 - v^2)^{\gamma-1}}{\pi(1 - 2au\cos\mu - 2av\sin\mu + a^2(u^2 + v^2))^{\gamma+1}}$$

on $0 \leq u^2 + v^2 \leq 1$ when transformed to Cartesian coordinates. An outline of the derivation of (2.1) from (1.1) is given in Appendix A. Note that application of the Möbius transformation itself to (1.1) would simply result in a rotation of the Möbius distribution through π radians. Note also the simplicity of the form of (2.1) and (2.2), including the normalisation constant: not even a gamma function is involved! The symmetric beta distributions, including the uniform, are recovered when $a = 0$.

The parameter μ controls the orientation of distribution (2.1). The standard form of the distribution takes $\mu = 0$, in which case (2.2) reduces to

$$f_{\gamma, a, 0}(u, v) = \frac{\gamma(1 - a^2)^{\gamma+1}(1 - u^2 - v^2)^{\gamma-1}}{\pi((1 - au)^2 + a^2 v^2)^{\gamma+1}}.$$

It is envisaged that much the most useful members of family (2.1) or (2.2) will be those for which $\gamma > 1$, whose densities are unimodal (Section 3). Speaking in broad general

terms, the parameter γ is a concentration parameter, in the sense that the greater the value of γ the less the variability of the distribution, and the parameter a controls the skewness of the distribution. More precisely, a controls the off-centredness of the distribution, the distance of the mode from the centre of the disc (or antimode when $\gamma < 1$) being a monotone increasing function of a (Section 3). The fact that these parameters are measuring quite different aspects of the distribution is borne out by the theoretical likelihood calculations of Section 5: it is very appealing to find that μ, γ and a are orthogonal parameters.

3. Unimodality

Humans look at distributions in (x, y)-space (Section 4) and hence the modality questions of interest concern the distribution in form (2.2), although it still proves clearer to give the results in terms of polar coordinates.

For $\gamma \neq 1$, the Möbius distribution (2.2) is unimodal with its mode at $(\theta, r) = (\mu, r_-)$ when $\gamma > 1$ and is uniantimodal with its antimode at $(\theta, r) = ((\mu+\pi)\mathrm{mod}(2\pi), r_+)$ when $\gamma < 1$. Here,

$$(3.1) \qquad r_\pm = \frac{\sqrt{(\gamma-1)^2 + 8a^2(\gamma+1)} \pm (\gamma-1)}{4a}.$$

This is so for all $0 \leq a < 1$; it is proved for $0 < a < 1$ in Appendix B and has already been asserted for $a = 0$ in Section 1, when $r_\pm = 0$. Note that, for $\gamma > 1$, r_- increases monotonically from 0 to 1 as a increases from 0 to 1 and, for $a > 0$, reduces monotonically from 1 to a as γ increases from 1 to ∞; for $\gamma < 1$, r_+ increases monotonically from 0 to $(\gamma + 1)/2$ as a increases from 0 to 1 and, for $a > 0$, reduces monotonically from 1 to $(\sqrt{1 + 8a^2} - 1)/4a$ as γ decreases from 1 to 0. When $\gamma = 1$, the distribution has no modal point in the interior of the disc; its maximum, for $a > 0$, is then at $(\theta, r) = (\mu, 1)$. These properties tie in with the fact that

$$f(1, \theta) = \begin{cases} 0 & \text{if} \quad \gamma > 1, \\ \frac{\gamma(1-a^2)^2}{\pi(1-2a\cos(\theta-\mu)+a^2)^2} & \text{if} \quad \gamma = 1, \\ \infty & \text{if} \quad \gamma < 1. \end{cases}$$

4. Pictures of densities

Without loss of generality, take $\mu = 0$. The four frames of Fig. 1 show the essential behaviour of distribution (2.2) for the most important case of $\gamma > 1$. For given $\gamma > 1$, increasing a moves the distribution off-centre (in the direction specified by $\theta = \mu$, here 0). For given a, increasing γ 'tightens up' the distribution about its mode.

The two frames of Fig. 2 exhibit some of the other behaviour of distribution (2.2). In Fig. 2(a) is a case in which $\gamma = 1$, resulting in a monotone density across the disc. Figure 2(b) exhibits a single antimodal case corresponding to $\gamma < 1$.

5. Maximum likelihood estimation

Write a random sample of data in polar form as (r_i, θ_i), $i = 1, \ldots, n$. Based on density (2.1), assuming $a \neq 0$, maximum likelihood estimators $\hat{\mu}$ and $\hat{a}$ of μ and a satisfy
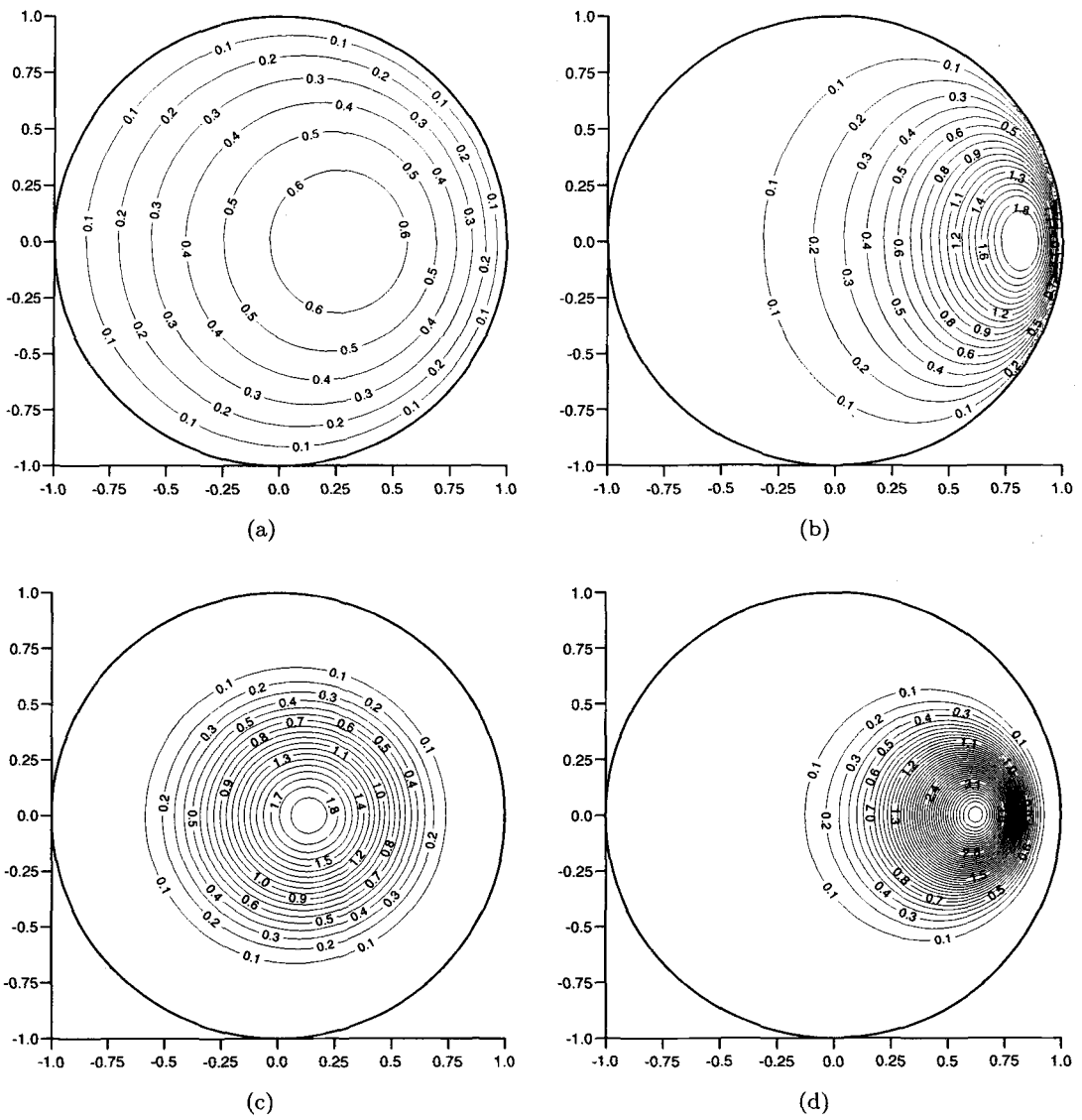
Fig. 1. Density (2.1)/(2.2) for $\mu = 0$ and: (a) $\gamma = 2$, $a = 0.1$; (b) $\gamma = 2$, $a = 0.5$; (c) $\gamma = 6$, $a = 0.1$; (d) $\gamma = 6$, $a = 0.5$.

the estimating equations

$$\sum_{i=1}^{n} \frac{r_i \sin(\theta_i - \hat{\mu})}{1 - 2\hat{a}r_i \cos(\theta_i - \hat{\mu}) + \hat{a}^2 r_i^2} = 0$$

and

$$n^{-1} \sum_{i=1}^{n} \frac{r_i \{\cos(\theta_i - \hat{\mu}) - \hat{a}r_i\}}{1 - 2\hat{a}r_i \cos(\theta_i - \hat{\mu}) + \hat{a}^2 r_i^2} = \frac{\hat{a}}{1 - \hat{a}^2},$$
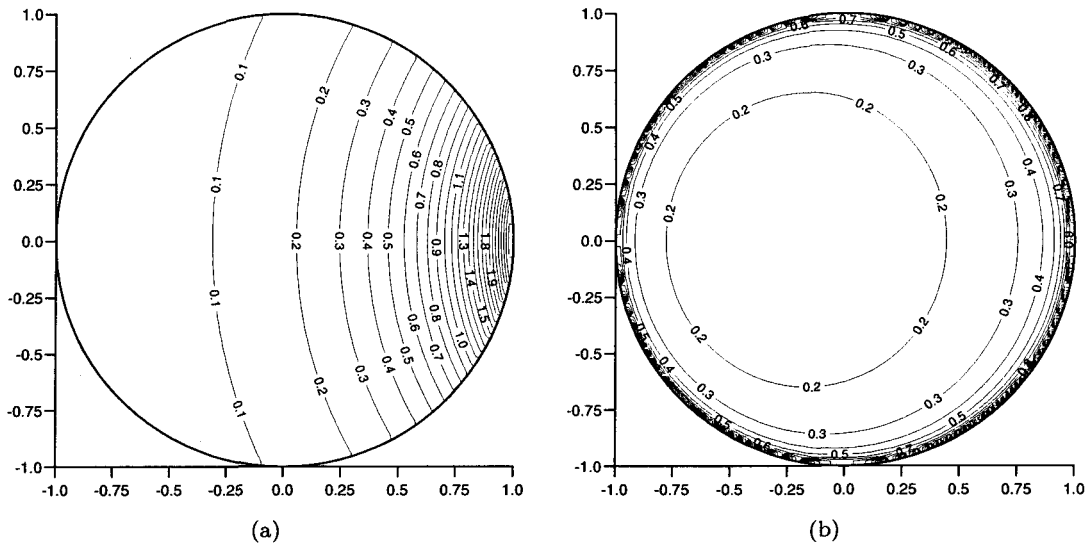
Fig. 2. Density (2.1)/(2.2) for $\mu = 0$ and: (a) $\gamma = 1$, $a = 0.5$; (b) $\gamma = 0.5$, $a = 0.1$.

while $\hat{\gamma}$ is given by

$$\hat{\gamma}^{-1} = n^{-1}\sum_{i=1}^{n}\log\{1 - 2\hat{a}r_i\cos(\theta_i - \hat{\mu}) + \hat{a}^2r_i^2\} - \log(1 - \hat{a}^2) - n^{-1}\sum_{i=1}^{n}\log(1 - r_i^2).$$

In practice, the only log-likelihood terms involving μ and a are $\gamma + 1$ times

$$(5.1) \qquad n\log(1 - a^2) - \sum_{i=1}^{n}\log\{1 - 2ar_i\cos(\theta_i - \mu) + a^2r_i^2\}.$$

I can therefore make a contour plot of the function (5.1) on the finite parameter space $-\pi < \mu \leq \pi$, $0 \leq a < 1$, to greatly assist the search for the global maximum over μ and a. Having ascertained the approximate position of the maximum visually, any standard (local) optimizer can be used to home in on the precise value of the maximum. Also, because

$$\frac{1 - 2\hat{a}r_i\cos(\theta_i - \hat{\mu}) + \hat{a}^2r_i^2}{(1 - \hat{a}^2)(1 - r_i^2)} \geq 1 + \frac{(\hat{a} - r_i)^2}{(1 - \hat{a}^2)(1 - r_i^2)} \geq 1,$$

$\hat{\gamma} > 0$.

Further differentiations of the likelihood equations immediately show that the non-diagonal terms of the expected information matrix are all zero. Assuming $a \neq 0$, this implies that the three parameters of distribution (2.1)/(2.2) are globally orthogonal, and hence that their maximum likelihood estimators are asymptotically independent (by the standard theory of maximum likelihood estimation, e.g., Cox and Hinkley (1974)). The diagonal terms of the expected information matrix can readily be seen to be n times:

$$\iota_{\mu\mu} = 2a(\gamma + 1)E\left[\frac{R\{(1 + a^2R^2)\cos(\Theta - \mu) - 2aR\}}{\{1 - 2aR\cos(\Theta - \mu) + a^2R^2\}^2}\right], \qquad \iota_{\gamma\gamma} = \gamma^{-2}$$

and

$$\iota_{aa} = 2(\gamma + 1) \left(\frac{1 + a^2}{(1 - a^2)^2} + E \left[\frac{R^2 \{1 + 2aR\cos(\Theta - \mu) - 2\cos^2(\Theta - \mu) - a^2 R^2\}}{\{1 - 2aR\cos(\Theta - \mu) + a^2 R^2\}^2} \right] \right).$$

The asymptotic variances of $\hat{\mu}$, $\hat{\gamma}$ and $\hat{a}$ are, of course, n^{-1} times $1/\iota_{\mu\mu}$, $1/\iota_{\gamma\gamma}$ and $1/\iota_{aa}$, respectively. Note, in particular, that the asymptotic variance of $\hat{\gamma}$ is proportional to γ.

The case of uniformity, $a = 0$, $\gamma = 1$, is on the boundary of the parameter space and corresponds to indeterminate μ. If it is desired to test for uniformity on the basis of $\hat{a}$ and $\hat{\gamma}$, then the approach of Davies (1977) could be used.

6. Example

For a brief illustrative example, the small dataset, size $n = 19$, given in Table 1 of Johnson and Wehrly (1977) is considered. These data consist of values, r_i, of ozone concentration which, for the purposes of this example, are divided by a maximal value taken to be 120, and wind direction, θ_i, converted from degrees to radians. The data were collected at a weather station in Milwaukee, U.S.A., in 1975, and are plotted in Fig. 3. They are clearly not uniformly distributed on the disc.

The contour plot (not shown) of that part of the log-likelihood dependent only on μ and a exhibited a clear single maximum at $\hat{\mu} = 0.648$, $\hat{a} = 0.284$. The corresponding value of $\hat{\gamma} = 4.494$. The fitted model is shown in Fig. 4. Given the small sample size, the model appears to fit reasonably well. However, one might alternatively argue that there is a hint of clustering in the data and that the fitted model is something of a compromise solution. Without more data, it cannot be told whether the single component model would really prove adequate or whether, perhaps, a mixture of Möbius distributions might be better.

Continuing with the model as fitted and using values of the *observed* information, approximate 95% confidence intervals for μ, a and γ are $(0.200, 1.097)$, $(0.162, 0.406)$ and $(2.473, 6.514)$, respectively.
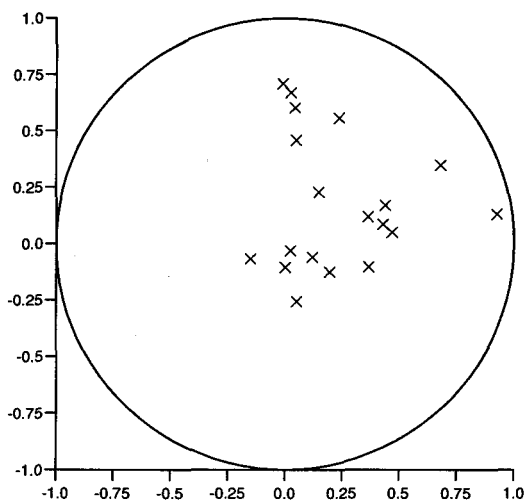
Fig. 3. A scatterplot of the ozone concentration/wind direction data from Johnson and Wehrly (1977).
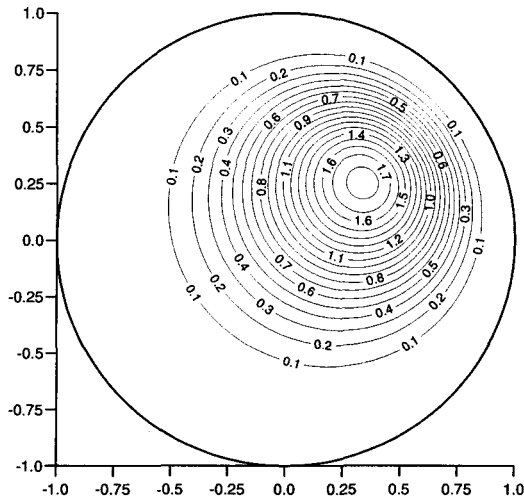
Fig. 4. The version of density (2.1)/(2.2) fitted to the data of Johnson and Wehrly (1977): $\hat{\mu} = 0.648$, $\hat{a} = 0.284$ and $\hat{\gamma} = 4.494$.

7. Conditionals and marginals

The conditional distribution of $\Theta \mid R = r$ is

$$f_{\gamma,a,\mu}(\theta \mid r) = \{2\pi P_\gamma(z)(z - \sqrt{z^2 - 1}\cos(\theta - \mu))^{\gamma+1}\}^{-1}$$

where $z = (1 + a^2 r^2)/(1 - a^2 r^2)$ and $P_\gamma(z)$ is the associated Legendre function (Abramowitz and Stegun (1965), Chapter 8). This is the subset of the wide class of symmetric unimodal circular distributions proposed by Jones and Pewsey (2004) corresponding to their parameter $\psi = -1/(\gamma + 1) \in (-1, 0)$—a set of distributions 'bridging the gap' between the von Mises distribution ($\psi = 0$; $\gamma = \infty$) and the wrapped Cauchy distribution ($\psi = -1$; $\gamma = 0$) (see, e.g., Mardia and Jupp (1999) for the latter two distributions). Formulae for the conditional trigonometric moments of Θ given r are given in terms of ratios of associated Legendre functions by Jones and Pewsey (2004). This set of distributions also arises as the conditional distribution of twice the angle given a fixed radius in a bivariate spherically symmetric t distribution on 2γ degrees of freedom with nonzero location (Shimizu and Iida (2002)).

The marginal distribution of R is then immediately

$$(7.1) \qquad f(r) = 2\gamma(1 - a^2)^{\gamma+1} \frac{r(1 - r^2)^{\gamma-1}}{(1 - a^2 r^2)^{\gamma+1}} P_\gamma\left(\frac{1 + a^2 r^2}{1 - a^2 r^2}\right)$$

on $0 \le r \le 1$. It can be confirmed, using 7.137.9 and 8.755.2 of Gradshteyn and Ryzhik (1994), that $\int_0^1 f(r)dr = 1$ and that formula (7.1) covers the symmetric beta case $a = 0$ because $P_\gamma(1) = 1$.

The conditional density of $R \mid \Theta = \theta$ does not have a particularly attractive or recognisable form, nor have I managed to obtain its normalising constant or, therefore, the marginal distribution of Θ. Neither is there anything very interesting to be said about the conditional or marginal distributions of U or V. Seshadri (1991) applies

a real univariate Möbius transformation to the univariate symmetric beta distribution to obtain an interesting class of univariate distributions with density proportional to $(1 - t^2)^{\lambda-1}/(1 - \alpha t)^{2\lambda}$ on $-1 < t < 1$, for a certain parameter α. It is perhaps surprising that this distribution does not figure in this section.

8. Comparison with alternative distribution

The alternative suggestion of Jones (2002) has density

$$(8.1) \qquad \frac{(1 + x)^{b_1 - \gamma - (1/2)}(1 - x)^{b_2 - \gamma - (1/2)}(1 - x^2 - y^2)^{\gamma-1}}{2^{b_1 + b_2 - 1}B(b_1, b_2)B(\gamma, 1/2)}$$

on $0 \le x^2 + y^2 \le 1$ with parameters $b_1, b_2, \gamma > 0$. Here, $B(\cdot, \cdot)$ is the beta function. The distribution has a Beta(b_1, b_2) X-marginal and a (rescaled) Beta(γ, γ) conditional distribution for $Y \mid X$. It was proposed as a special case of a general method for skewing multivariate distributions and hence does not have especially strong affinity with the polar nature of the disc. Indeed, (8.1) has a specific orientation and therefore, to allow general orientation, needs the addition of a fourth parameter.

Distribution (8.1) has equally pleasant modality properties as distribution (2.2), but is not so simple in terms of interpretation of parameters. In particular, the two parameters b_1 and b_2, or rather a complicated function of them involving their difference, drive skewness, while in (2.1)/(2.2), a does the skewness job in an immediately appealing way; b_1 and b_2 are not orthogonal parameters.

Acknowledgements

Appendix A: Outline demonstration of veracity of (2.1)

Write s and ϕ for the polar coordinates of z. If (x, y) has density (1.1), then the density of (s, ϕ) is $\gamma s(1 - s^2)^{\gamma-1}/\pi$. Recall that we are transforming by the inverse Möbius transformation $w = M_{-c}(z)$. Write

$$B = 1 - 2ar\cos(\theta - \mu) + a^2 r^2.$$

The squared modulus of the Möbius transformation is

$$s^2 = B^{-1}(a^2 - 2ar\cos(\theta - \mu) + r^2)$$

so that

$$\partial s/\partial r = (sB^2)^{-1}(1 - a^2)\{r(1 + a^2) - a(1 + r^2)\cos(\theta - \mu)\}$$

and

$$\partial s/\partial \theta = (sB^2)^{-1}ar(1 - a^2)(1 - r^2)\sin(\theta - \mu).$$

It is also the case that, by multiplying top and bottom of the Möbius transformation by $1 - c\bar{w}$,

$$\tan\phi = \frac{r\sin\theta - a(1 + r^2)\sin\mu - a^2 r\sin(\theta - 2\mu)}{r\cos\theta - a(1 + r^2)\cos\mu + a^2 r\cos(\theta - 2\mu)}.$$

Then,

$$\partial\phi/\partial r = -(sB)^{-2}a(1-a^2)(1-r^2)\sin(\theta-\mu)$$

and

$$\partial\phi/\partial\theta = (sB)^{-2}r(1-a^2)\{r(1+a^2)-a(1+r^2)\cos(\theta-\mu)\}.$$

The reciprocal of the Jacobian then reduces to $(sB^2)^{-1}r(1-a^2)^2$, and with $1-s^2 = B^{-1}(1-a^2)(1-r^2)$, (1.1) can be seen to transform to (2.1).

Appendix B: Proof of unimodality when $\gamma \neq 1$ and $a > 0$

Interest is in the modality of (2.2) but it will be more convenient to work in terms of polar coordinates, noting that this is not the same as studying the modality of (2.1). Consider the derivatives of that part of the log density which depends on r and θ, namely

$$(\gamma-1)\log(1-r^2) - (\gamma+1)\log(1-2ar\cos(\theta-\mu)+a^2r^2).$$

Then $\partial \log f(r,\theta)/\partial\theta = 0$ when $\theta = \mu$, when θ equals whichever of $\mu \pm \pi \in (-\pi, \pi]$ or when $r = 0$. Consider, first, $\theta = \mu$. Then,

$$\frac{\partial \log f(r,\theta)}{\partial r}\Bigg|_{\theta=\mu} = -\frac{2(\gamma-1)r}{(1-r^2)} + \frac{2a(\gamma+1)}{1-ar} = \frac{g(r)}{(1-r^2)(1-ar)}$$

where $g(r) = 2\{-2ar^2 - (\gamma-1)r + a(\gamma+1)\}$. Now, since $g(r) \to -\infty$ as $r \to \pm\infty$, $g(0) = 2a(1+\gamma) > 0$ and $g(1) = 2(1-a)(1-\gamma)$, only when $\gamma > 1$ is there a single r in $(0,1)$ such that $g(r) = 0$, namely r_- given by (3.1). Also,

$$\frac{\partial^2 \log f(r,\theta)}{\partial\theta^2}\Bigg|_{\theta=\mu,r=r_-} = \frac{-2ar_-(\gamma+1)}{(1-ar_-)^2} < 0,$$

$$\frac{\partial^2 \log f(r,\theta)}{\partial r^2}\Bigg|_{\theta=\mu,r=r_-} = -\frac{2(\gamma-1)(1+r_-^2)}{(1-r_-^2)^2} + \frac{2a^2(\gamma+1)}{(1-ar_-)^2}$$

$$= -\frac{2(\gamma-1)(r_-^2 - 2ar_- + 1)}{(1-ar_-)(1-r_-^2)^2},$$

which is negative for $\gamma > 1$, and $\partial^2 \log f(r,\theta)/\partial\theta\partial r\,|_{\theta=\mu,r=r_-} = 0$. It follows that for $\gamma > 1$ there is a maximum at $(\theta, r) = (\mu, r_-)$.

Now take $\theta = \mu \pm \pi$. This gives

$$\frac{\partial \log f(r,\theta)}{\partial r}\Bigg|_{\theta=\mu\pm\pi} = -\frac{2(\gamma-1)r}{1-r^2} - \frac{2a(\gamma+1)}{1+ar} = \frac{-g(-r)}{(1-r^2)(1+ar)}.$$

Since $g(-1) = 2(1+a)(\gamma-1)$, there is a single solution of $g(-r) = 0$ in $(0,1)$ when $\gamma < 1$, and this is r_+ given by (3.1). In this case,

$$\frac{\partial^2 \log f(r,\theta)}{\partial\theta^2}\Bigg|_{\theta=\mu\pm\pi,r=r_+} = \frac{2ar_+(\gamma+1)}{(1+ar_+)^2} > 0,$$

$$\frac{\partial^2 \log f(r,\theta)}{\partial r^2}\Bigg|_{\theta=\mu\pm\pi,r=r_+} = -\frac{2(\gamma-1)(1+r_+^2)}{(1-r_+^2)^2} + \frac{2a^2(\gamma+1)}{(1+ar_+)^2} > 0$$

and $\partial^2 \log f(r,\theta)/\partial\theta\partial r \mid_{\theta=\mu\pm\pi, r=r_+} = 0$. It follows that for $\gamma < 1$ there is a minimum at $(\theta, r) = (\mu \pm \pi, r_+)$.

When $r = 0$ solves $\partial \log f(r,\theta)/\partial\theta = 0$, $\theta = (\mu \pm (2n+1)\pi/2)\mathrm{mod}(2\pi)$, $n = 0, 1, \ldots$, solves $\partial \log f(r,\theta)/\partial r = 0$. This can be shown to correspond to a saddlepoint.

REFERENCES

Abramowitz, M. and Stegun, I. A. (eds.) (1965). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover, New York.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*, Chapman and Hall, London.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, **64**, 247–254.

Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*, Chapman and Hall, London.

Gradshteyn, I. S. and Ryzhik, I. M. (1994). *Table of Integrals, Series, and Products*, 5th ed. (ed. A. Jeffrey), Academic Press, San Diego.

Johnson, M. E. (1987). *Multivariate Statistical Simulation*, Wiley, New York.

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

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

Jones, M. C. and Pewsey, A. R. (2004). A family of distributions on the circle (under consideration).

Kotz, S. (1975). Multivariate distributions at a cross road, *Statistical Distributions in Scientific Work, Volume 1—Models and Structures* (eds. G. P. Patil, S. Kotz and J. K. Ord), 247–270, Reidel, Dordrecht.

Krantz, S. G. (1999). *Handbook of Complex Analysis*, Birkhäuser, Boston, Massachusetts.

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

Seshadri, V. (1991). A family of distributions related to the McCullagh family, *Statistics and Probability Letters*, **12**, 373–378.

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