Biometrika (2002), **89**, 3, *pp*. 719–723

Printed in Great Britain

Miscellanea

Probabilistic model for two dependent circular variables

By HARSHINDER SINGH

Department of Statistics, West Virginia University, Morgantown, West Virginia 26506-6330, U.S.A.

hsingh@stat.wvu.edu

VLADIMIR HNIZDO and EUGENE DEMCHUK

Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, West Virginia 26505-2888, U.S.A.

vhnizdo@cdc.gov eed5@cdc.gov

Summary

Motivated by problems in molecular biology and molecular physics, we propose a five-parameter torus analogue of the bivariate normal distribution for modelling the distribution of two circular random variables. The conditional distributions of the proposed distribution are von Mises. The marginal distributions are symmetric around their means and are either unimodal or bimodal. The type of shape depends on the configuration of parameters, and we derive the conditions that ensure a specific shape. The utility of the proposed distribution is illustrated by the modelling of angular variables in a short linear peptide.

Some key words: Bivariate circular data; Circular random variable; Directional data; Torus; Von Mises distribution

1. Introduction

Inspired by problems in probabilistic modelling of torsional angles in molecules (Demchuk & Singh, 2001), we aim to develop a bivariate distribution on the torus which imbeds naturally the bivariate normal distribution when the range of observations is small. To this end, we propose to replace the quadratic terms $\frac{1}{2}(\theta_i - \mu_i)^2$ and the linear terms $(\theta_i - \mu_i)$, for $i = 1, 2$, in the bivariate normal probability density function by their natural circular analogues $1 - \cos(\theta_i - \mu_i)$ and $\sin(\theta_i - \mu_i)$, respectively. This gives the probability density function of two circular random variables Θ_1 and Θ_2 the form

$$f(\theta_1, \theta_2) = C \exp\{\kappa_1 \cos(\theta_1 - \mu_1) + \kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_1 - \mu_1) \sin(\theta_2 - \mu_2)\}, \quad (1{\cdot}1)$$

for $-\pi \leqslant \theta_1,\ \theta_2 < \pi$, where $\kappa_1, \kappa_2 \geqslant 0$, $-\infty < \lambda < \infty$, $-\pi \leqslant \mu_1,\ \mu_2 < \pi$, and C is the appropriate normalisation constant. The parameter λ accounts for the statistical dependence between Θ_1 and Θ_2: if $\lambda = 0$, then Θ_1 and Θ_2 are independent and each of them assumes the von Mises distribution. The distribution (1·1) is a five-parameter analogue on the torus of the bivariate normal density, to which it reduces when the angular observation range is small. A similar construction on the sphere was given by Kent (1982). The probability density function (1·1) has a natural generalisation that allows multiple modes in the marginal distributions, obtained by changing $(\theta_i - \mu_i)$ in (1·1) to $l_i(\theta_i - \mu_i)$, for $i = 1, 2$, where l_1 and l_2 are positive integers.

Mardia (1975a, b) introduced a class of bivariate distributions with a probability density function

$$c \exp\{\kappa_1 \cos(\theta_1 - \mu_1) + \kappa_2 \cos(\theta_2 - \mu_2) + (\cos\theta_1, \sin\theta_1) A (\cos\theta_2, \sin\theta_2)^{\mathrm{T}}\}, \qquad (1{\cdot}2)$$

where A is a 2×2 matrix. Mardia (1975b) and Jupp & Mardia (1980) considered submodels of (1·2) by restricting A. Another submodel of (1·2) was considered by Rivest (1988), namely

$$c \exp\{\kappa_1 \cos(\theta_1 - \mu_1) + \kappa_2 \cos(\theta_2 - \mu_2) + \alpha \cos(\theta_1 - \mu_1) \cos(\theta_2 - \mu_2) + \beta \sin(\theta_1 - \mu_1) \sin(\theta_2 - \mu_2)\}. \qquad (1{\cdot}3)$$

For $\alpha = 0$ and $\beta = \lambda$, this density reduces to (1·1). This choice of parameters and the two alternative choices $\alpha = \beta$ and $\alpha = -\beta$ in (1·3) all have asymptotic bivariate normal distributions under high concentration. For simplicity we limit attention to (1·1) in this paper. A comprehensive review of directional statistics has been given by Mardia & Jupp (1999).

2. PROPERTIES OF THE PROPOSED DISTRIBUTION

Let (Θ_1, Θ_2) be distributed according to the probability density function (1·1). When the fluctuations in Θ_1 and Θ_2 are sufficiently small, so that

$$\cos(\theta_i - \mu_i) \simeq 1 - \tfrac{1}{2}(\theta_i - \mu_i)^2, \quad \sin(\theta_i - \mu_i) \simeq (\theta_i - \mu_i) \quad (i = 1, 2),$$

then it follows that (Θ_1, Θ_2) has approximately a bivariate normal distribution with parameters

$$\sigma_1^2 = \frac{\kappa_2}{\kappa_1\kappa_2 - \lambda^2}, \quad \sigma_2^2 = \frac{\kappa_1}{\kappa_1\kappa_2 - \lambda^2}, \quad \rho = \frac{\lambda}{\surd(\kappa_1\kappa_2)}. \qquad (2{\cdot}1)$$

For these bivariate-normal parameters to be meaningful, only the condition $\lambda^2 < \kappa_1\kappa_2$ is required.

Define new parameters a and β by $\kappa_2 = a \cos\beta$ and $\lambda \sin(\theta_1 - \mu_1) = a \sin\beta$, so that

$$a = \{\kappa_2^2 + \lambda^2 \sin^2(\theta_1 - \mu_1)\}^{1/2}, \quad \tan\beta = (\lambda/\kappa_2)\sin(\theta_1 - \mu_1).$$

Write $a = a(\theta_1)$ and $\beta = \beta(\theta_1)$ to emphasise the dependence on θ_1. Then the marginal density of Θ_1 is given by

$$f_1(\theta_1) = C\, e^{\kappa_1 \cos(\theta_1 - \mu_1)} \int_{-\pi}^{\pi} e^{a(\theta_1)\cos(\theta_2 - \mu_2 - \beta)}\, d\theta_2 = 2\pi C I_0\{a(\theta_1)\}\, e^{\kappa_1 \cos(\theta_1 - \mu_1)} \quad (-\pi \leqslant \theta_1 < \pi), \qquad (2{\cdot}2)$$

where $I_0(y)$ is the modified Bessel function of order zero. The marginal probability density function of Θ_2 is obtained in a similar way.

Rivest (1988) has derived an expression for the normalisation constant of (1·3) in terms of a doubly infinite series. Here we give an expression for C in terms of a singly infinite series, involving sequences of modified Bessel functions, which can be computed efficiently using the algorithm of Amos (1974). The proof of the theorem is omitted.

THEOREM 1. *In the joint probability density function* (1·1), *the normalisation constant C is given by*

$$\frac{1}{C} = 4\pi^2 \sum_{m=0}^{\infty} \binom{2m}{m} \left(\frac{\lambda^2}{4\kappa_1\kappa_2}\right)^m I_m(\kappa_1) I_m(\kappa_2), \qquad (2{\cdot}3)$$

where $I_m(y)$ is the modified Bessel function I of order m.

Given that $\Theta_1 = \theta_1$, the conditional probability density function of Θ_2 is

$$f(\theta_2 \mid \theta_1) = \frac{1}{2\pi I_0\{a(\theta_1)\}}\, e^{a(\theta_1)\cos(\theta_2 - \mu_2 - \beta)}. \qquad (2{\cdot}4)$$

Thus the conditional distribution of Θ_2, given that $\Theta_1 = \theta_1$, is a von Mises distribution with the concentration parameter $a(\theta_1)$ and mean angle $\mu_2 + \beta$. If $\kappa_2 \to \infty$ and $\lambda \to \infty$ so that $\lambda/\kappa_2 \to \zeta$, then the concentration parameter $a(\theta_1)$ of the conditional von Mises distribution tends to infinity for each given θ_1. Thus, for this case, $\Theta_2 = \mu_2 + \arctan\{\zeta \sin(\Theta_1 - \mu_1)\}$ with probability one. We observe that, when the fluctuations in the angular variables are small, this function is approximately linear, and that it is curvilinear for the case of larger fluctuations in the angular variables.

We now discuss the properties of the marginal distribution of Θ_1. Similar properties hold for the marginal distribution of Θ_2 with κ_1, κ_2 and μ_1, μ_2 interchanged. In the following theorem, we established the circular mean and circular variance of Θ_1.

THEOREM 2. *If* (Θ_1, Θ_2) *is distributed according to the bivariate probability density function* (1·1), *then the following hold*:

(a) $E\{\sin(\Theta_1 - \mu_1)\} = 0$, *implying that* μ_1 *is the circular mean of* Θ_1;

(b) *the circular variance of* Θ_1 *is given by*

$$1 - E\{\cos(\Theta_1 - \mu_1)\} = 1 - \frac{\sum_{m=0}^{\infty} \binom{2m}{m} \left(\frac{\lambda^2}{4\kappa_1\kappa_2}\right)^m I_{m+1}(\kappa_1) I_m(\kappa_2)}{\sum_{m=0}^{\infty} \binom{2m}{m} \left(\frac{\lambda^2}{4\kappa_1\kappa_2}\right)^m I_m(\kappa_1) I_m(\kappa_2)}. \quad (2·5)$$

Next we investigate the shape of the marginal distribution of Θ_1 using the following lemma.

LEMMA 1. *Let* $A(y) = I_1(y)/I_0(y)$, *where* $I_0(y)$ *and* $I_1(y)$ *are the modified Bessel functions I of orders* 0 *and* 1, *respectively. Then* $A(y)/y$ *is a decreasing function in* $(0, \infty)$.

THEOREM 3. *Let* (1·1) *be the bivariate probability density function of* (Θ_1, Θ_2) *with* $\lambda \neq 0$. *Then the marginal distribution of* Θ_1 *is symmetric around* $\theta_1 = \mu_1$ *and unimodal* (*respectively bimodal*) *with the mode at* μ_1 (*respectively with the modes at* $\mu_1 - \theta_1^*$ *and* $\mu_1 + \theta_1^*$) *if and only if*

$$A(\kappa_2) \leqslant (\textit{respectively} >) \kappa_1\kappa_2/\lambda^2, \quad (2·6)$$

where θ_1^* *is given by the equation* $\cos(\theta_1^* - \mu_1) A\{a(\theta_1^*)\}/a(\theta_1^*) = \kappa_1/\lambda^2$.

If we use Lemma 1 it is easily seen that $A(y) \leqslant \frac{1}{2}y$ for all $y \geqslant 0$. Thus, it follows from Theorem 3 that $\lambda^2 \leqslant 2\min(\kappa_1, \kappa_2)$ provides a sufficient condition for both the marginal distributions to be unimodal.

Note that the marginal distribution of Θ_1 is not von Mises. Expanding $\log f_1(\theta_1)$ and log of a von Mises density with circular mean μ_1 and concentration parameter κ in a Taylor series expansion about $\theta_1 = \mu_1$, we see that for large κ_1 and κ both are quadratic functions over the bulk of their support. Thus, for a suitably chosen κ, given by $\kappa = \kappa_1\{1 - A(\kappa_2)\lambda^2/\kappa_1\kappa_2\}$, both densities are asymptotically the same normal density and hence are similar to one another.

3. APPLICATION

We illustrate the utility of model (1·1) by fitting it to 1000 paired observations of the ω and ϕ dihedral angles of the proline residue, taken from a long molecular dynamics trajectory of the AYPYD peptide in water (Demchuk et al., 1997). Using the function-minimisation system MINUIT (James & Roos, 1975), we fitted model (1·1) to these bivariate data, and obtained the following maximum likelihood estimates: $\hat{\kappa}_1 = 35·41$, $\hat{\kappa}_2 = 20·17$, $\hat{\lambda} = -13·70$, $\hat{\mu}_1 = 0·073$ rad and $\hat{\mu}_2 = -1·560$ rad, with 95% confidence intervals of widths 3·12, 1·78, 1·94, 0·012 and 0·016, respectively. The sample circular variances of Θ_1, Θ_2 and the sample signed correlation coefficient of Fisher & Lee (1983) were respectively 0·019, 0·033 and $-0·493$. The approximating formula of Rivest (1988) gave $\hat{\kappa}_1 = 34·57$, $\hat{\kappa}_2 = 19·68$ and $\hat{\lambda} = -12·39$. The chi-squared goodness-of-fit test confirmed the adequacy of the fit.

ACKNOWLEDGEMENT

The authors are grateful to S. R. Jammalamadaka, Dan S. Sharp, James E. Harner, Sidney C. Soderholm and N. Misra for helpful discussions, and to L. Zheng for his help with the preparation of the data. They are also grateful to an anonymous reviewer and to an associate editor for insightful comments that led to an improved presentation. H. Singh was partially supported by the Biostatistics Branch of the Health Effects Laboratory Division, National Institute for Occupational Safety and Health.

APPENDIX

Proofs of Theorems 2 *and* 3 *and Lemma* 1

Proof of Theorem 2. Without loss of generality we assume that $\mu_1 = \mu_2 = 0$. The proof of part (a) is simple and is thus omitted. To prove (b), we observe that

$$E(\cos\Theta_1) = 2\pi C \frac{d}{d\kappa_1}\frac{1}{2\pi C} = C\frac{d}{d\kappa_1}\frac{1}{C}. \tag{A·1}$$

Part (b) of the theorem now follows from (2·3) by using $(d/dy)\{I_m(y)/y^m\} = I_{m+1}(y)/y^m$; see Abramowitz & Stegun (1965, 9.6.28). □

Proof of Lemma 1. Using $(d/dy)\{I_m(y)/y^m\} = I_{m+1}(y)/y^m$, we have that

$$\frac{d}{dy}\frac{A(y)}{y} = \frac{I_0(y)I_2(y) - \{I_1(y)\}^2}{y\{I_0(y)\}^2}.$$

It is sufficient to show that $I_0(y)I_2(y) < \{I_1(y)\}^2$ in order to prove that $A(y)/y$ is a decreasing function of $y > 0$. However,

$$I_0(y)I_2(y) = \frac{y^2}{4}\sum_{s=0}^{\infty}\frac{(2+2s)!}{\{s!(2+s)!\}^2}\left(\frac{y^2}{4}\right)^s, \tag{A·2}$$

$$\{I_1(y)\}^2 = \frac{y^2}{4}\sum_{s=0}^{\infty}\frac{(2+2s)!}{\{(1+s)!\}^2(2+s)!s!}\left(\frac{y^2}{4}\right)^s, \tag{A·3}$$

(Abramowitz & Stegun, 1965, 9.1.14, 9.6.3). The ratio of the coefficient of $(y^2/4)^s$ in (A·2) to that in (A·3) is $(s+1)/(s+2) < 1$, and this implies that $I_0(y)I_2(y) < \{I_1(y)\}^2$. □

Proof of Theorem 3. Clearly $f_1(\theta_1)$ given by (2·2) is symmetric around $\theta_1 = \mu_1$. Without loss of generality we take $\mu_1 = 0$. We have that

$$\log f_1(\theta_1) = \log(2\pi C) + \log I_0\{a(\theta_1)\} + \kappa_1 \cos\theta_1.$$

The derivative of the above equation yields

$$\frac{f_1'(\theta_1)}{f_1(\theta_1)} = \frac{I_1\{a(\theta_1)\}}{I_0\{a(\theta_1)\}a(\theta_1)}\lambda^2 \sin\theta_1 \cos\theta_1 - \kappa_1 \sin\theta_1.$$

Thus

$$\frac{f_1'(\theta_1)}{f_1(\theta_1)} = \lambda^2 \sin\theta_1\left[\frac{A\{a(\theta_1)\}}{a(\theta_1)}\cos\theta_1 - \frac{\kappa_1}{\lambda^2}\right]. \tag{A·4}$$

For $\theta_1 \in [-\pi, -\frac{1}{2}\pi]$, both $\sin\theta_1$ and $\cos\theta_1$ are nonpositive, implying that the right-hand side of equation (A·4) is nonnegative, further implying that $f_1(\theta_1)$ is increasing in $[-\pi, -\frac{1}{2}\pi]$. Similarly, if $\theta_1 \in [\frac{1}{2}\pi, \pi]$, $\sin\theta_1$ is nonnegative whereas $\cos\theta_1$ is nonpositive. Thus, the right-hand side of equation (A·4) is nonpositive, implying that $f_1(\theta_1)$ is decreasing in $[\frac{1}{2}\pi, \pi]$.

Now we investigate the behaviour of $f_1(\theta_1)$ for $\theta_1 \in [0, \frac{1}{2}\pi]$. In this range $a(\theta_1)$ is an increasing function of θ_1. By Lemma 1, $A\{a(\theta_1)\}/a(\theta_1)$ is a nonnegative decreasing function of θ_1 in $[0, \frac{1}{2}\pi]$. Also, $\cos\theta_1$ is a nonnegative decreasing function of θ_1 in $[0, \frac{1}{2}\pi]$, implying that $(\cos\theta_1)A\{a(\theta_1)\}/a(\theta_1)$ is a nonnegative decreasing function of θ_1 in $[0, \frac{1}{2}\pi]$. Also, $\sin\theta_1$ is nonnegative in $[0, \frac{1}{2}\pi]$. Thus, if

$$[A\{a(0)\}/a(0)]\cos 0 \leqslant \kappa_1/\lambda^2,$$

that is if $A(\kappa_2) \leqslant \kappa_1\kappa_2/\lambda^2$, then $f_1(\theta_1)$ is decreasing in $[0, \frac{1}{2}\pi]$, and by the symmetry of $f_1(\theta_1)$ around $\theta_1 = 0$ it follows that $f_1(\theta_1)$ is increasing in $[-\frac{1}{2}\pi, 0]$. Thus, for this case $f_1(\theta_1)$ is a symmetrical and unimodal distribution, with the mode at $\theta_1 = \mu_1 = 0$.

On the other hand, if $A(\kappa_2) > \kappa_1\kappa_2/\lambda^2$, then $f_1(\theta_1)$ is increasing in $[0, \theta_1^*]$ and decreasing in $[\theta_1^*, \frac{1}{2}\pi]$ where θ_1^* is the solution of the equation

$$[A\{a(\theta_1^*)\}/a(\theta_1^*)]\cos\theta_1^* = \kappa_1/\lambda^2.$$

By the symmetry of $f_1(\theta_1)$ around $\theta_1 = 0$, it follows that $f_1(\theta_1)$ is increasing in $[-\frac{1}{2}\pi, -\theta_1^*]$ and decreasing in $[-\theta_1^*, 0]$. Thus, in this case, $f_1(\theta_1)$ is a symmetric bimodal distribution. This establishes the theorem. □

References

Abramowitz, M. & Stegun, I. A. (Ed.) (1965). *Handbook of Mathematical Functions*. New York: Dover.

Amos, D. E. (1974). Computation of modified Bessel functions and their ratios. *Math. Comp.* **28**, 239–57.

Demchuk, E., Bashford, D. & Case, D. A. (1997). Dynamics of a type VI reverse turn in a linear peptide in aqueous solution. *Folding Design* **2**, 35–46.

Demchuk, E. & Singh, H. (2001). Statistical thermodynamics of hindered rotation from computer simulations. *Molec. Phys.* **99**, 627–36.

Fisher, N. I. & Lee, A. J. (1983). A correlation coefficient for circular data. *Biometrika* **70**, 327–32.

James, F. & Roos, M. (1975). MINUIT—A system for function minimization and analysis of the parameter errors and correlations. *Comp. Phys. Commun.* **10**, 343–67.

Jupp, P. E. & Mardia, K. V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika* **67**, 163–73.

Kent, J. T. (1982). The Fisher–Bingham distribution on the sphere. *J. R. Statist. Soc.* B **44**, 71–80.

Mardia, K. V. (1975a). Statistics of directional data (with Discussion). *J. R. Statist. Soc.* B **37**, 349–93.

Mardia, K. V. (1975b). Characterizations of directional distributions. In *Statistical Distributions in Scientific Work*, **3**, Ed. G. P. Patil, S. Kotz and J. K. Ord, pp. 365–85. Dordrecht: Reidel.

Mardia, K. V. & Jupp, P. E. (1999). *Directional Statistics*. New York: Wiley.

Rivest, L. P. (1988). A distribution for dependent unit vectors. *Commun. Statist.* A **17**, 461–83.

[*Received March* 2001. *Revised December* 2001]