Ann Inst Stat Math (2013) 65:833–858
DOI 10.1007/s10463-012-0394-5

Extending circular distributions through transformation of argument

Toshihiro Abe · Arthur Pewsey · Kunio Shimizu

Received: 20 July 2011 / Revised: 8 November 2012 / Published online: 12 January 2013

Abstract This paper considers the general application to symmetric circular densities of two forms of change of argument: one produces extended families of distributions which contain symmetric densities which are more flat-topped, as well as others which are more sharply peaked, than the originals, and the second produces families which are skew. General results for the modality and shape characteristics of the densities which ensue are presented, and maximum likelihood estimation of the parameters of two extensions of the Jones–Pewsey family is discussed. The application of these two particular extended families is illustrated within analyses of data on monthly cases of sudden infant death syndrome in the UK.

Keywords Asymmetry · Batschelet distributions · Flat-toppedness · Jones–Pewsey distribution · Papakonstantinou distributions · Peakedness · Skewness · Symmetry

T. Abe
Department of Management Science, Faculty of Engineering, Tokyo University of Science,
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-0825, Japan
e-mail: abe@ms.kagu.tus.ac.jp

A. Pewsey (✉)
Department of Mathematics, Escuela Politécnica, University of Extremadura,
Avenida de la Universidad s/n, 10003 Cáceres, Spain
e-mail: apewsey@unex.es

K. Shimizu
Department of Mathematics, Keio University, 3-14-1 Hiyoshi,
Kohoku-ku, Yokohama 223-8522, Japan
e-mail: shimizu@math.keio.ac.jp

1 Introduction

The cardioid and von Mises distributions are two of the most widely known unimodal symmetric circular distributions (see Mardia and Jupp 1999, Sect. 3.5 or Jammalamadaka and SenGupta 2001, Sect. 2.2, for example). The density of the cardioid distribution with modal direction μ, $-\pi \leq \mu < \pi$, is

$$f_{\mathrm{C}}(\theta) = (2\pi)^{-1}(1 + 2\rho\cos(\theta - \mu)), \quad -\pi \leq \theta < \pi, \tag{1}$$

where $0 \leq \rho \leq 1/2$ is a concentration parameter, whilst the analogous von Mises density is

$$f_{\mathrm{vM}}(\theta) = \{2\pi I_0(\kappa)\}^{-1} \exp(\kappa\cos(\theta - \mu)), \quad -\pi \leq \theta < \pi, \tag{2}$$

where $\kappa \geq 0$ is a concentration parameter and $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero (Abramowitz and Stegun 1972, p. 376, Eq. 9.6.19). Note that in both densities the argument $\theta - \mu$ is transformed using the cosine function, which is even.

To extend the shape characteristics of the cardioid and von Mises distributions, respectively, Papakonstantinou (1979) and Batschelet (1981, Sects. 15.6 and 15.7) effectively made use of two extensions of the $\cos(\theta - \mu)$ transformation. The first involves replacing $\cos(\theta - \mu)$ by

$$\cos(\theta - \mu + \nu\sin(\theta - \mu)), \quad -\infty < \nu < \infty, \tag{3}$$

whilst the second replaces it by

$$\cos(\theta - \mu + \nu\cos(\theta - \mu)), \quad -\infty < \nu < \infty. \tag{4}$$

Clearly, (3) and (4) are cosine functions of two different transformations of the argument $\theta - \mu$. Substituting $\phi + \pi/2$ for θ in (4) gives

$$\sin(\phi - \mu + \nu\sin(\phi - \mu)), \quad -\infty < \nu < \infty, \tag{5}$$

so the second transformation can be expressed in terms of either two sines or two cosines. If $\cos(\theta - \mu)$ is replaced by either (3) or (4), the normalising constants in (1) and (2) generally no longer apply and must be recalculated. Obviously, the original density is unchanged if $\nu = 0$. The resulting densities are unimodal if $-1 < \nu < 1$, and are multimodal otherwise.

Substituting (3) for $\cos(\theta - \mu)$ in (1) or (2) results in densities which are also symmetric but more flat-topped (sharply peaked) than the original cardioid or von Mises densities if $-1 < \nu < 0$ ($0 < \nu < 1$). Thus, for this transformation of argument, $-1 < \nu < 1$ acts as a kurtosis regulating parameter. The density obtained by substituting (3) for $\cos(\theta - \mu)$ within the cardioid density (1), which in order to

reflect its origins (Papakonstantinou 1979) we shall refer to as Papakonstantinou's symmetric extended cardioid (SEC) density, is

$$f_{\mathrm{SEC}}(\theta)=\frac{1}{2\pi\{1-k\,J_1(\nu)\}}\{1+k\,\cos(\theta-\mu+\nu\sin(\theta-\mu))\}\,,\quad -\pi\le\theta<\pi,\qquad(6)$$

where $0\le k\le 1$ is a concentration parameter and, more generally, $J_p(z)$ denotes the Bessel function of the first kind and order p (Gradshteyn and Ryzhik 2007, p. 912, Eq. 8.411.1), whose integral representation is given by

$$J_p(z)=\frac{1}{\pi}\int_0^{\pi}\cos(p\theta-z\sin\theta)\mathrm{d}\theta,\quad p=0,\pm1,\pm2,\ldots.\qquad(7)$$

Abe et al. (2009) study this distribution in depth and explain its origins. Substituting (3) for $\cos(\theta-\mu)$ within the von Mises density (2) results in what we shall refer to as Batschelet's symmetric extended von Mises (SEvM) density,

$$f_{\mathrm{SEvM}}(\theta)=c_{\kappa,\nu}^{-1}\exp\{\kappa\cos(\theta-\mu+\nu\sin(\theta-\mu))\},\quad -\pi\le\theta<\pi.$$

The normalising constant, $c_{\kappa,\nu}^{-1}$, generally has no closed form and must be calculated using numerical integration. Pewsey et al. (2011) provide an in-depth treatment of this distribution introduced by Batschelet (1981, p. 288).

For $0<|\nu|<1$, substituting (4), or equivalently (5), for $\cos(\theta-\mu)$ within densities (1) or (2) leads to unimodal densities which are asymmetric. They are skewed clockwise (anticlockwise) if $-1<\nu<0$ ($0<\nu<1$), $-1<\nu<1$ thus playing the role of a skewness parameter. Substituting (5) within the cardioid density (1) results in what we shall refer to as Papakonstantinou's asymmetric extended cardioid (AEC) density (Papakonstantinou 1979),

$$f_{\mathrm{AEC}}(\theta)=\frac{1}{2\pi}\{1+k\sin(\theta-\mu+\nu\sin(\theta-\mu))\}\,,\qquad(8)$$

where $0\le k\le 1$ is a concentration parameter. Note the closed form of the normalising constant. Substituting (5) for $\cos(\theta-\mu)$ within the von Mises density (2) leads to what, in order to reflect its origins, we shall call Batschelet's asymmetric extended von Mises (AEvM) density (Batschelet 1981, p. 286),

$$f_{\mathrm{AEvM}}(\theta)=d_{\kappa,\nu}^{-1}\exp\{\kappa\sin(\theta-\mu+\nu\sin(\theta-\mu))\},\quad -\pi\le\theta<\pi.\qquad(9)$$

As was the case for its symmetric counterpart, the normalising constant, here $d_{\kappa,\nu}^{-1}$, generally has no closed form and must be calculated using quadrature.

Recently, Jones and Pewsey (2005) proposed a family of unimodal symmetric circular distributions with density

$$f_{\mathrm{JP}}(\theta)\propto\{1+\tanh(\kappa\psi)\cos(\theta-\mu)\}^{1/\psi}\,,\quad -\pi\le\theta<\pi,\qquad(10)$$

where $-\pi \leq \mu < \pi, \kappa \geq 0$ and $-\infty < \psi < \infty$. As before, μ is the modal direction and κ is a concentration parameter, whilst ψ is referred to as a shape index. Despite this terminology, the overall shape of the distribution is actually controlled by both κ and ψ. A major appeal of the Jones–Pewsey family is that it contains the uniform ($\kappa = 0$ or finite κ and $\psi \to \pm\infty$), cardioid ($\psi = 1$), von Mises ($\psi = 0$), wrapped Cauchy ($\psi = -1$), Cartwright's power-of-cosine ($\psi > 0$ and $\kappa \to \infty$) and circular t (Shimizu and Iida 2002) ($-1 < \psi < 0$) distributions. Consequently, it provides an important over-arching family of distributions which, together with likelihood-ratio testing, can be used to explore the fits of many of the classical models of circular statistics. Note that in (10) also, the argument $\theta - \mu$ is transformed using the cosine function.

In Sect. 2.1 we consider the general application of the change of argument in (3) to symmetric circular densities and provide results for the modality and shape properties of the symmetric densities which ensue. In Sect. 2.2 we propose an extension to the Jones–Pewsey family based on the application of (3) and consider special cases of, and maximum likelihood estimation for, the resulting symmetric family of distributions. Sect. 3.1 focuses on the general application of (5) to symmetric circular densities and presents results for the modality of the ensuing densities which are generally skewed. In Sect. 3.2 we propose a skew extension of the Jones–Pewsey family based on the application of (5). Results are provided for the shape properties of the densities within the new family and consideration is given to maximum likelihood estimation of its parameters. We also study the properties of, and method of moments estimation for, its Papakonstantinou AEC subclass. In Sect. 4 we illustrate the application of the two new extended Jones–Pewsey families within analyses of data on monthly cases of sudden infant death syndrome in the UK. Concluding remarks are made in Sect. 5. Four R scripts used in the illustrative analyses of Sect. 4 are presented in an Appendix.

2 Extending the shape characteristics of symmetric distributions

2.1 General construction

Here we generalise the approach of Papakonstantinou (1979) and Batschelet (1981, Sect. 15.7) based on the transformation of argument in (3) to one which can be used to extend the shape characteristics of a wide family of base symmetric circular densities. Detailed results are presented for the modality and shape properties of the densities which ensue.

2.1.1 Definition

Without loss of generality, we will assume in what follows that the location parameter, μ, equals 0. Let $g(\cos\theta)$ ($-\pi \leq \theta < \pi$) denote a symmetric circular density which is a function of $\cos\theta$. Now consider the circular density

$$f_\nu(\theta) = c_\nu^{-1} g(\cos(\theta + \nu \sin\theta)), \quad -\pi \leq \theta < \pi, \tag{11}$$

where $-\infty < \nu < \infty$ and c_ν^{-1} is a normalising constant. Clearly, when $\nu = 0$, $f_\nu(\theta)$ reduces to the original density, i.e., $f_0(\theta) = g(\cos\theta)$, and so we can write $f_\nu(\theta) = c_\nu^{-1} f_0(\theta + \nu \sin\theta)$, where $c_\nu = \int_{-\pi}^{\pi} f_0(\theta + \nu \sin\theta)\mathrm{d}\theta$. Because the cosine function is even, and the sine function is odd, the density (11) is also symmetric about zero.

2.1.2 Modality

Assume now that the symmetric circular density $f_0(\theta) = g(\cos\theta)$ is unimodal with continuous first derivative $f_0'(\theta)$ and that $f_0'(0) = 0$. $f_0(\theta)$ will have a mode at $\theta = 0$ and an antimode at $\theta = -\pi$.

Proposition 1 *Density* (11) *is unimodal with a mode at* $\theta = 0$ *and an antimode at* $\theta = -\pi$ *if and only if* $|\nu| \leq 1$. *For* $|\nu| > 1$, (11) *has* $(2|m| + 1)$ *modes if* $m\pi < h(\nu) \leq (m+1)\pi$ *for* $m = 1, 2, \ldots$ *or* $m\pi \leq h(\nu) < (m+1)\pi$ *for* $m = -1, -2, \ldots$, *where* $h(\nu) = \cos^{-1}(-1/\nu) + \nu\sqrt{1 - 1/\nu^2}$.

Proof The proof proceeds as in the proof of Proposition 1 in Abe et al. (2009) and is therefore omitted. □

2.1.3 Shape properties

In addition to the assumptions of Sect. 2.1.2, here we further assume that the second derivative of $f_0(\theta)$, $f_0''(\theta)$, exists and is continuous and that $f_0(\theta) > 0$ for all θ $(-\pi \leq \theta < \pi)$. Since $f_\nu'(0) = 0$, the curvature of $f_\nu(\theta)$ at $\theta = 0$, defined as $f_\nu''(0)/[1 + \{f_\nu'(0)\}^2]^{3/2}$, equals $f_\nu''(0)$.

Proposition 2 *For* $-1 \leq \nu_1 < \nu_2 \leq 1$*:*

(a) $f_{\nu_1}(0) < f_{\nu_2}(0)$ *and* $f_{\nu_1}(-\pi) < f_{\nu_2}(-\pi)$,
(b) $f_{\nu_1}''(0) > f_{\nu_2}''(0)$ *and* $f_{\nu_1}''(-\pi) < f_{\nu_2}''(-\pi)$.

Proof (a) First note that $0 < \theta + \nu_i \sin\theta < \pi$ $(i = 1, 2)$ and thus $\theta + \nu_1 \sin\theta < \theta + \nu_2 \sin\theta$ for all θ in the interval $(0, \pi)$. Since $f_0(\theta)$ has a unique mode at $\theta = 0$ and is strictly decreasing on the interval $(0, \pi)$, it holds that

$$c_{\nu_1} = 2\int_0^{\pi} f_0(\theta + \nu_1 \sin\theta)\mathrm{d}\theta > 2\int_0^{\pi} f_0(\theta + \nu_2 \sin\theta)\mathrm{d}\theta = c_{\nu_2}.$$

Hence,

$$f_{\nu_1}(0) = c_{\nu_1}^{-1} f_0(0) < c_{\nu_2}^{-1} f_0(0) = f_{\nu_2}(0).$$

Similarly, $f_{\nu_1}(-\pi) < f_{\nu_2}(-\pi)$. (b) As

$$f_\nu''(\theta) = c_\nu^{-1}\left\{-\nu \sin\theta\, f_0'(\theta + \nu \sin\theta) + (1 + \nu\cos\theta)^2 f_0''(\theta + \nu \sin\theta)\right\}$$

and $f_0''(0) < 0$, it follows that

$$f_{\nu_1}''(0) = c_{\nu_1}^{-1}(1+\nu_1)^2 f_0''(0) > c_{\nu_2}^{-1}(1+\nu_2)^2 f_0''(0) = f_{\nu_2}''(0).$$

Similarly, $f_{\nu_1}''(-\pi) < f_{\nu_2}''(-\pi)$. □

Given the two results above, ν is clearly a shape parameter which controls the peakedness/flat-toppedness properties of the density. If conditions (a) and (b) are satisfied, we will say that f_{ν_1} is more flat-topped than f_{ν_2}; or, equivalently, f_{ν_2} is more sharply peaked than f_{ν_1}. Obviously, if $\nu < 0$, then f_ν is more flat-topped than the original base density f_0, and if $\nu > 0$, then f_ν is more sharply peaked than the original base density f_0.

2.2 A symmetric extension of the Jones–Pewsey family

Here we apply the general construction of Sect. 2.1 to produce a symmetric extension of the Jones–Pewsey family which has even more shape flexibility. Various important special cases of it are identified, and maximum likelihood estimation of its parameters is discussed.

2.2.1 Definition and special cases

Applying the general construction of Sect. 2.1.1 to the Jones–Pewsey density (10) results in what we will refer to as the symmetric extended Jones–Pewsey (SEJP) family of distributions with density

$$f_{\text{SEJP}}(\theta) = c_{\kappa,\psi,\nu}^{-1}\{1+\tanh(\kappa\psi)\cos(\theta-\mu+\nu\sin(\theta-\mu))\}^{1/\psi}, \quad -\pi \le \theta < \pi, \quad (12)$$

where $-\pi \le \mu < \pi$, $\kappa \ge 0$, $-\infty < \psi < \infty$, $-\infty < \nu < \infty$ and $c_{\kappa,\psi,\nu} = \int_{-\pi}^{\pi}\{1+\tanh(\kappa\psi)\cos(\theta+\nu\sin\theta)\}^{1/\psi}\mathrm{d}\theta$ must generally be computed numerically. This density reduces to a Jones–Pewsey density when $\nu = 0$, a Papakonstantinou SEC density with $k = \tanh(\kappa)$ when $\psi = 1$ and a Batschelet SEvM density when $\psi = 0$. When $\psi = -1$, a symmetric extended wrapped Cauchy (SEWC) distribution is obtained with density

$$f_{\text{SEWC}}(\theta) = c_{\kappa,-1,\nu}^{-1}\{1-\tanh(\kappa)\cos(\theta-\mu+\nu\sin(\theta-\mu))\}^{-1}, \quad -\pi \le \theta < \pi.$$

The shape flexibility of unimodal SEJP densities is apparent from the densities portrayed in Fig. 1. Contour plots of the circular kurtosis, γ_2, (see Sect. 3.2.3 for a general definition) for unimodal members of the SEJP family are presented in Fig. 2. The achievable levels of circular kurtosis for the base Jones–Pewsey distribution are those associated with the value $\nu = 0$. The extra levels of kurtosis offered by the SEJP family are associated with $|\nu|$-values close to 1 and, generally, ψ-values close to 0.

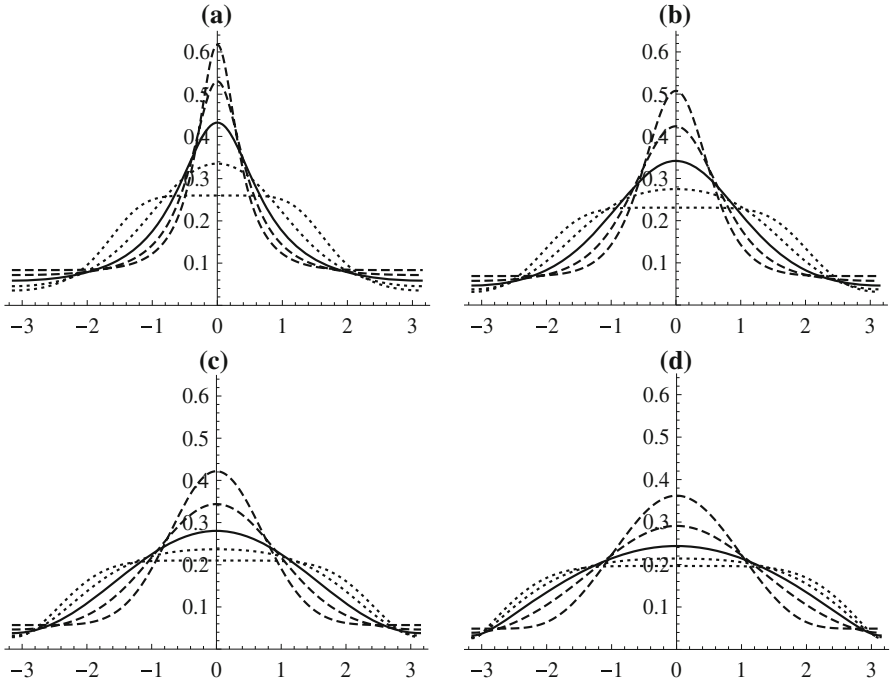

Fig. 1 Unimodal SEJP densities with $\mu = 0$, $\kappa = 1$ and (**a**) $\psi = -1$ (SEWC), (**b**) $\psi = 0$ (SEvM), (**c**) $\psi = 1$ (SEC) and (**d**) $\psi = 2$. The line types, in order of increasing height at the origin, are *dotted* for $\nu = -1, -0.5$, *solid* for $\nu = 0$ and *dashed* for $\nu = 0.5, 1$. The densities are plotted on $[-\pi, \pi)$ using a common vertical scale

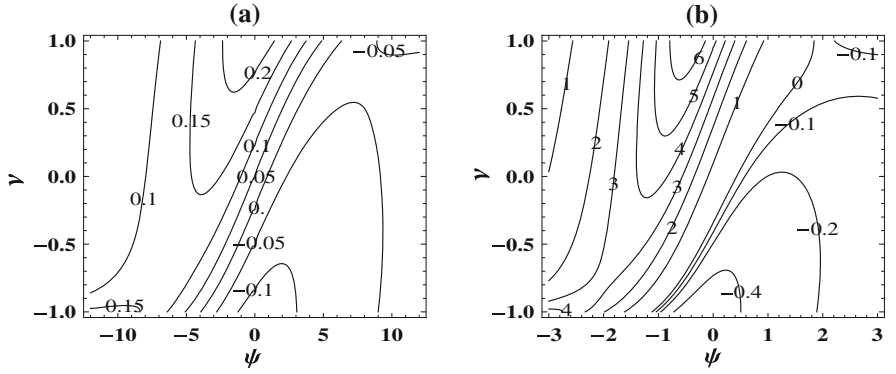

Fig. 2 Contour plots of the circular kurtosis, γ_2, for unimodal members of the SEJP family as a function of $-6/\kappa < \psi < 6/\kappa$ and $-1 \leq \nu \leq 1$, for (**a**) $\kappa = 0.5$ and (**b**) $\kappa = 2$

2.2.2 Maximum likelihood estimation

Let $\theta_1, \ldots, \theta_n$ denote a random sample of size n drawn from a SEJP distribution with density (12). We consider the most relevant practical inferential scenario in which all four parameters, μ, κ, ψ and ν, are unknown. The log-likelihood function is

$$\ell(\mu, \kappa, \psi, \nu) = -n \log c_{\kappa,\psi,\nu} + \frac{1}{\psi} \sum_{i=1}^{n} \log\{1 + \tanh(\kappa\psi) \cos(\theta_i - \mu + \nu \sin(\theta_i - \mu))\}. \quad (13)$$

The first- and second-order partial derivatives of (13) with respect to each of the parameters, used in the score vector and information matrices, are somewhat involved and to save on space we do not reproduce them here. For those interested in obtaining them, we would recommend the use of symbolic mathematical computing packages, such as *Mathematica* or the open source *Sage* software.

For a random sample of SEJP data grouped into the m class intervals $[\theta_0, \theta_1)$, $[\theta_1, \theta_2)$,...,$[\theta_{m-1}, \theta_m)$, where $\theta_0 = \theta_m$, with n_j observations in the jth interval and thus a total of $n = n_1 + n_2 + \cdots + n_m$ observations, the log-likelihood function is given by

$$\ell^{\dagger}(\mu, \kappa, \psi, \nu) = -n \log c_{\kappa,\psi,\nu} + \sum_{j=1}^{m} n_j \log \int_{\theta_{j-1}}^{\theta_j} \{1 + \tanh(\kappa\psi) \cos(\theta - \mu + \nu \sin(\theta - \mu))\}^{1/\psi} \mathrm{d}\theta. \quad (14)$$

In general, there are no closed-form expressions for the maximum likelihood estimates (MLEs) and numerical methods of optimisation are required to identify them. We have developed a suite of scripts for the open source R statistical software with which to perform likelihood based inference for the SEJP family. Scripts RS1 and RS2, specifically designed for use with the first data set analysed in Sect. 4, are reproduced in the Appendix. The data in question are grouped but we also have analogous scripts for use with continuous data, as well as others for computing the profile log-likelihood functions of the other three parameters. The full suite of R scripts is freely available from the authors upon request. The most important R functions used in the various scripts are `integrate` and `optim` for computing $c_{\kappa,\psi,\nu}$ and minimizing the negative log-likelihood, respectively. The `optim` function is particularly powerful as it allows the user to try different starting values for the optimization process and includes an option for computing the Hessian matrix evaluated at the maximum likelihood solution. When it exists, the latter is easily inverted to obtain the observed information matrix and asymptotic correlation matrix for the parameter estimates, and hence asymptotic normal theory based confidence sets for the parameters. The `optim` function also allows the user to impose box-constraints on the parameters via the use of the "L-BFGS-B" method of optimisation due to Byrd et al. (1995). We make use of this optimisation method in the illustrative examples of Sect. 4.

3 Skewing symmetric distributions

3.1 General construction

Here we generalise the approach of Papakonstantinou (1979) and Batschelet (1981, Sect. 15.7) based on the transformation (4) (or, equivalently, (5)) to one which can

be used to produce skew extensions of symmetric circular densities. Results are also given for the modality of the densities which ensue.

3.1.1 Definition

We will assume here too, without loss of generality, that the location parameter, μ, equals 0. Suppose that $g(\sin\theta)$ $(-\pi \leq \theta < \pi)$ denotes a circular density which is symmetric about $\pi/2$. Then

$$f_\nu(\theta) = d_\nu^{-1} g(\sin(\theta + \nu \sin\theta)), \quad -\pi \leq \theta < \pi, \tag{15}$$

where $-\infty < \nu < \infty$ and d_ν^{-1} is a normalising constant, is also a circular density. Clearly, when $\nu = 0$, $f_\nu(\theta)$ reduces to the original density, i.e., $f_0(\theta) = g(\sin\theta)$, and so we can write $f_\nu(\theta) = d_\nu^{-1} f_0(\theta + \nu \sin\theta)$ where $d_\nu = \int_{-\pi}^{\pi} f_0(\theta + \nu \sin\theta)\mathrm{d}\theta$ and $d_\nu = d_{-\nu}$. Density (15) is symmetric about $\pi/2$ if $\nu = 0$; otherwise it is asymmetric. Setting $\phi = \theta - \pi/2$, (15) is transformed into $f_\nu(\phi) = d_\nu^{-1} g(\cos(\phi + \nu \cos\phi))$, where $-3\pi/2 \leq \phi < \pi/2$ and $f_{-\nu}(\phi) = f_\nu(-\phi)$.

3.1.2 Modality

In what follows we make use of the representation (15) because of its mathematical tractability. In that representation, the function $H_\nu(\theta) = \theta + \nu \sin\theta : [-\pi, \pi] \to [-\pi, \pi]$ is monotone increasing if $|\nu| \leq 1$. Thus, for $|\nu| \leq 1$, $\theta_\nu^* = H_\nu^{-1}(\pi/2)$ is defined uniquely.

We will assume that the density $f_0(\theta) = g(\sin\theta)$ is unimodal with mode at $\theta = \pi/2$ and antimode at $\theta = -\pi/2$. We further assume that it has a continuous first derivative and that $f_0'(\pi/2) = 0$.

Proposition 3 *Density* (15) *is unimodal with a mode at* $\theta = 0$ *and an antimode at* $\theta = -\pi$ *if and only if* $|\nu| \leq 1$. *For* $|\nu| > 1$, (15) *has* $(2|m| + 1)$ *modes if* $m\pi < h(\nu) \leq (m+1)\pi$ *for* $m = 1, 2, \ldots$ *or* $m\pi \leq h(\nu) < (m+1)\pi$ *for* $m = -1, -2, \ldots$, *where* $h(\nu) = \cos^{-1}(-1/\nu) + \nu\sqrt{1 - 1/\nu^2}$.

Proof The derivative of (15) with respect to θ is

$$f_\nu'(\theta) \propto (1 + \nu \cos\theta) \cos(H_\nu(\theta)) g'(\sin H_\nu(\theta)),$$

which is 0 for $\theta = \theta_\nu^*$ and $\theta = -\theta_\nu^*$. It is also 0 if $1 + \nu\cos\theta = 0$ or $\cos H_\nu(\theta) = 0$. If $|\nu| \leq 1$, then $f_\nu'(\theta)$ is non-zero for all θ such that $\theta \neq \theta_\nu^*$ or $\theta \neq -\theta_\nu^*$. As $f_\nu(\theta_\nu^*) > f_\nu(-\theta_\nu^*)$, $f_\nu(\theta)$ is unimodal with mode θ_ν^* and antimode $-\theta_\nu^*$. Now assume that $|\nu| > 1$ and consider the number of solutions to the equation $H_\nu'(\theta) = 1 + \nu\cos\theta = 0$ or $H_\nu(\theta) = \theta + \nu\sin\theta = (j + 1/2)\pi$ $(j = \pm 1, \pm 2, \ldots)$ in the interval $(0, \pi)$ since $H_\nu(\theta)$ is an odd function of θ. Let $\theta_1 = \cos^{-1}(-1/\nu)$ $(0 < \theta_1 < \pi)$, which is a unique solution to the equation $1 + \nu\cos\theta = 0$. Note that, for $\nu > 1$ $(\nu < -1)$, $1 + \nu\cos\theta$ is positive (negative) if $0 < \theta < \theta_1$ and is negative (positive) if $\theta_1 < \theta < \pi$. Since $H_\nu'(\theta) = 1 + \nu\cos\theta$, $H_\nu(\theta)$ takes an extreme value at $\theta = \theta_1$. Set $h(\nu) = H_\nu(\theta_1) = \cos^{-1}(-1/\nu) + \nu\sqrt{1 - 1/\nu^2}$. Then h is a monotone increasing

function since $h'(\nu) = \sqrt{1-1/\nu^2} > 0$, and thus $h(\nu) > h(1) = \pi$ if $\nu > 1$ and $h(\nu) < h(-1) = 0$ if $\nu < -1$. From the above, for $-\pi \le \theta < \pi$, $f_\nu(\theta)$ has two modes if $\pi < h(\nu) \le 3\pi/2$ or $-\pi/2 \le h(\nu) < 0$ and $2|m|$ modes if $(m-1/2)\pi < h(\nu) \le (m+1/2)\pi$ $(m = 2, 3, \ldots)$ for $\nu > 1$ or $(m+1/2)\pi \le h(\nu) < (m+3/2)\pi$ $(m = -2, -3, \ldots)$ for $\nu < -1$. □

3.2 A unimodal skew Jones–Pewsey family

Here we apply the general construction of Sect. 3.1 to produce a skew unimodal extension of the Jones–Pewsey family. We identify important special cases of it, provide results for the shape characteristics of densities within it and discuss maximum likelihood estimation of its parameters. We also study the properties of, and method of moments estimation for, its AEC subclass.

3.2.1 Definition

Due to the periodicity of circular distributions, the Jones–Pewsey density can also be represented as

$$f_{\mathrm{JP}}(\theta) \propto \{1 + \tanh(\kappa\psi)\sin(\theta-\mu)\}^{1/\psi}, \quad -\pi \le \theta < \pi.$$

As an application of (15), we propose the asymmetric extended Jones–Pewsey (AEJP) distribution with density

$$f_{\mathrm{AEJP}}(\theta) = d^{-1}_{\kappa,\psi,\nu}\{1+\tanh(\kappa\psi)\sin(\theta-\mu+\nu\sin(\theta-\mu))\}^{1/\psi}, \quad -\pi \le \theta < \pi, \quad (16)$$

where $-\pi \le \mu < \pi$, $\kappa \ge 0$, $-\infty < \psi < \infty$ and, to ensure unimodality, $-1 \le \nu \le 1$. The normalising constant is the inverse of $d_{\kappa,\psi,\nu} = \int_{-\pi}^{\pi}\{1+\tanh(\kappa\psi)\sin(\theta+\nu\sin\theta)\}^{1/\psi}\mathrm{d}\theta$ which must generally be computed numerically.

As special cases, (16) reduces to Papakonstantinou's AEC density (8) when $\psi = 1$ and $k = \tanh(\kappa)$, and to Batschelet's AEvM density (9) when $\psi = 0$. When $\psi = -1$, an asymmetric extended wrapped Cauchy (AEWC) distribution, with density

$$f_{\mathrm{AEWC}}(\theta) = d^{-1}_{\kappa,-1,\nu}\{1-\tanh(\kappa)\sin(\theta-\mu+\nu\sin(\theta-\mu))\}^{-1}, \quad -\pi \le \theta < \pi,$$

is obtained. The densities portrayed in Fig. 3 provide an indication of the flexibility of the AEJP family. The achievable levels of circular skewness, γ_1, (see Sect. 3.2.3 for a general definition) for densities within the AEJP family are pictured in Fig. 4. The highest levels of circular skewness are associated with $|\nu|$-values close to 1 and ψ-values close to 0.

3.2.2 Shape properties

Here we present results which summarise how the parameters ν and ψ control the peakedness/flat-toppedness of AEJP densities. To save on space, we will denote an

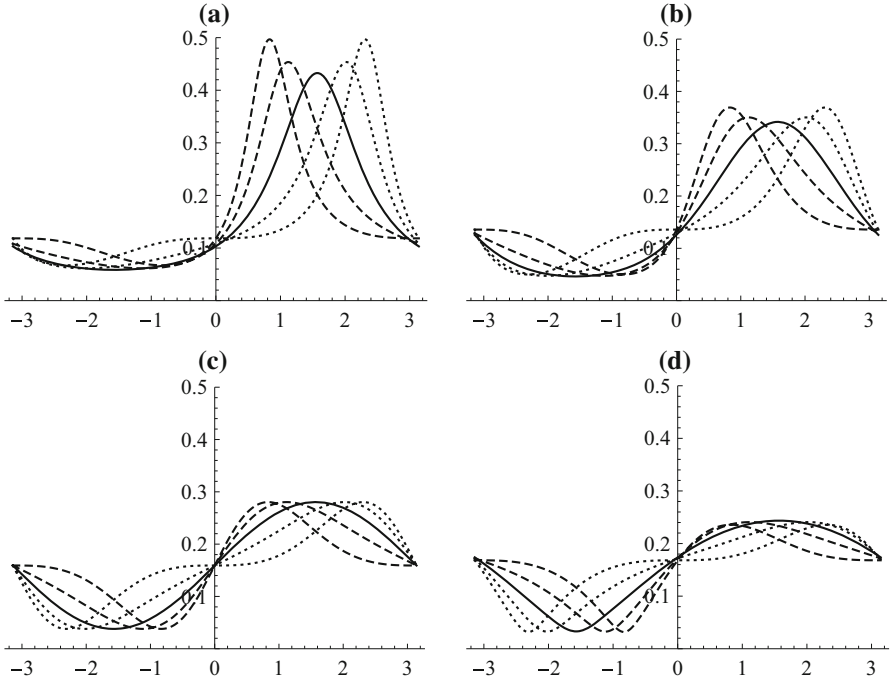

Fig. 3 AEJP densities with $\mu = 0, \kappa = 1$ and (**a**) $\psi = -1$ (AEWC), (**b**) $\psi = 0$ (AEvM), (**c**) $\psi = 1$ (AEC) and (**d**) $\psi = 2$. The line types, in order of increasing height at $\theta = 0.5$, are *dotted* for $\nu = -1, -0.5$, *solid* for $\nu = 0$ and *dashed* for $\nu = 0.5, 1$. The densities are plotted on $[-\pi, \pi)$ using a common vertical scale

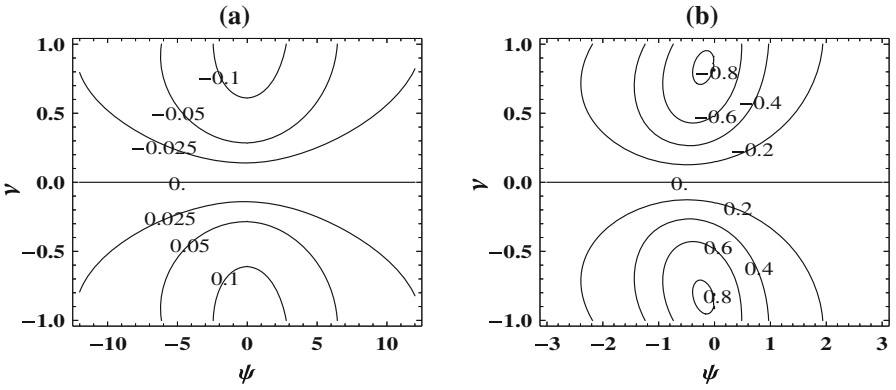

Fig. 4 Contour plots of the circular skewness, γ_1, for members of the AEJP family as a function of $-6/\kappa < \psi < 6/\kappa$ and $-1 \le \nu \le 1$, for (**a**) $\kappa = 0.5$ and (**b**) $\kappa = 2$

AEJP density with parameter ν by $f_\nu(\theta)$. As in Sect. 3.1.2, we denote the mode of such a density by θ^*_ν. Since $f'_\nu(\theta^*_\nu) = d^{-1}_{\kappa,\psi,\nu}(1 + \nu\cos\theta^*_\nu) f'_0(\pi/2) = 0$, the curvature of $f_\nu(\theta)$ at the mode, defined as $f''_\nu(\theta^*_\nu)/[1 + \{f'_\nu(\theta^*_\nu)\}^2]^{3/2}$, equals $f''_\nu(\theta^*_\nu)$.

Proposition 4 *For* $0 \le |\nu_1| < |\nu_2| \le 1$*, it holds that*

(a) $f_{\nu_1}(\theta^*_{\nu_1}) < f_{\nu_2}(\theta^*_{\nu_2})$, $f_{\nu_1}(-\theta^*_{\nu_1}) < f_{\nu_2}(-\theta^*_{\nu_2})$, $f_{\nu_1}(0) < f_{\nu_2}(0)$ *and* $f_{\nu_1}(-\pi) < f_{\nu_2}(-\pi)$ *if* $\psi < 1$*;*
(b) $f_{\nu_1}(\theta^*_{\nu_1}) = f_{\nu_2}(\theta^*_{\nu_2})$, $f_{\nu_1}(-\theta^*_{\nu_1}) = f_{\nu_2}(-\theta^*_{\nu_2})$, $f_{\nu_1}(0) = f_{\nu_2}(0)$ *and* $f_{\nu_1}(-\pi) = f_{\nu_2}(-\pi)$ *if* $\psi = 1$*. Moreover, these properties hold for all* $\nu_1, \nu_2 \in (-\infty, \infty)$*;*
(c) $f_{\nu_1}(\theta^*_{\nu_1}) > f_{\nu_2}(\theta^*_{\nu_2})$, $f_{\nu_1}(-\theta^*_{\nu_1}) > f_{\nu_2}(-\theta^*_{\nu_2})$, $f_{\nu_1}(0) > f_{\nu_2}(0)$ *and* $f_{\nu_1}(-\pi) > f_{\nu_2}(-\pi)$ *if* $\psi > 1$*;*
(d) $f''_{\nu_1}(\theta^*_{\nu_1}) > f''_{\nu_2}(\theta^*_{\nu_2})$ *and* $f''_{\nu_1}(-\theta^*_{\nu_1}) < f''_{\nu_2}(-\theta^*_{\nu_2})$.

Proof (a) First, assume that $0 \le \nu \le 1$, $\psi < 1$ ($\psi \ne 0$). Consider the derivative of $d_{\kappa,\psi,\nu}$ with respect to ν, i.e.,

$$\begin{aligned}
\frac{\partial\, d_{\kappa,\psi,\nu}}{\partial \nu} &= \frac{\tanh(\kappa\psi)}{\psi}\int_{-\pi}^{\pi} \sin\theta\cos H_\nu(\theta)\{1+\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi-1}\mathrm{d}\theta \\
&= \frac{\tanh(\kappa\psi)}{\psi}\int_{0}^{\pi} \sin\theta\cos H_\nu(\theta)[\{1+\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi-1} \\
&\quad -\{1-\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi-1}]\mathrm{d}\theta \\
&= \frac{\tanh(\kappa\psi)}{\psi}\left(\int_0^{\theta^*_\nu}+\int_{\theta^*_\nu}^{\pi}\right)\sin\theta\cos H_\nu(\theta)[\{1+\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi-1} \\
&\quad -\{1-\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi-1}]\mathrm{d}\theta \\
&= \frac{\tanh(\kappa\psi)}{\psi}\int_{-\pi/2}^{0}\left(\frac{\sin H_\nu^{-1}(-u)}{1+\nu\cos H_\nu^{-1}(-u)}-\frac{\sin H_\nu^{-1}(u+\pi)}{1+\nu\cos H_\nu^{-1}(u+\pi)}\right) \\
&\quad \times\cos u[\{1-\tanh(\kappa\psi)\sin u\}^{1/\psi-1}-\{1+\tanh(\kappa\psi)\sin u\}^{1/\psi-1}]\mathrm{d}u.
\end{aligned}$$

Define the function $q(u)$ as

$$q(u) = \frac{\sin H_\nu^{-1}(-u)}{1+\nu\cos H_\nu^{-1}(-u)} - \frac{\sin H_\nu^{-1}(u+\pi)}{1+\nu\cos H_\nu^{-1}(u+\pi)}.$$

We show that $q(u)$ is negative for all $u \in (-\pi/2, 0)$. Since $\sin x/(1+\nu\cos x)$ is a monotone increasing function of $x \in (0, \pi/2)$, it follows that the first term of $q(u)$, $\sin H_\nu^{-1}(-u)/\{1+\nu\cos H_\nu^{-1}(-u)\}$, is monotone decreasing. The second term of $q(u)$, $\sin H_\nu^{-1}(u+\pi)/\{1+\nu\cos H_\nu^{-1}(u+\pi)\}$, is minimum when $u = -\pi/2$ or $u = 0$. So, $q(u)$ is negative for all $u \in [-\pi/2, 0]$ because the maximum of $q(u)$ is $q(-\pi/2) = q(0) = 0$. From the fact that $\{1-\tanh(\kappa\psi)\sin u\}^{1/\psi-1} - \{1+\tanh(\kappa\psi)\sin u\}^{1/\psi-1} > 0$ when $\psi < 1$ ($\psi \ne 0$), it follows that $\partial d_{\kappa,\psi,\nu}/\partial\nu < 0$ and $d_{\kappa,\psi,\nu}$ is a monotone decreasing function of $\nu \in [0, 1]$. Since $d_{\kappa,\psi,\nu} = d_{\kappa,\psi,-\nu}$, we conclude that $d_{\kappa,\psi,\nu_1} > d_{\kappa,\psi,\nu_2}$ if $0 \le |\nu_1| < |\nu_2| \le 1$ and $\psi < 1$ ($\psi \ne 0$). The same inequality holds for the AEvM distribution, obtained when $\psi = 0$. Hence,

$$f_{\nu_1}(\theta^*_{\nu_1}) = d^{-1}_{\kappa,\psi,\nu_1} f_0(\pi/2) < d^{-1}_{\kappa,\psi,\nu_2} f_0(\pi/2) = f_{\nu_2}(\theta^*_{\nu_2})$$

if $\psi < 1$. Similarly, we obtain the inequalities $f_{\nu_1}(-\theta^*_{\nu_1}) < f_{\nu_2}(-\theta^*_{\nu_2})$, $f_{\nu_1}(0) < f_{\nu_2}(0)$ and $f_{\nu_1}(-\pi) < f_{\nu_2}(-\pi)$.

(b) The assertion is trivial.
(c) Noting that $\{1-\tanh(\kappa\psi)\sin u\}^{1/\psi-1}-\{1+\tanh(\kappa\psi)\sin u\}^{1/\psi-1}<0$ when $\psi>1$, then $\partial d_{\kappa,\psi,\nu}/\partial\nu>0$, and the assertion follows immediately.
(d) When $\psi\le 1$ and $0\le|\nu_1|<|\nu_2|\le 1$, it follows from

$$f_\nu''(\theta)=d_{\kappa,\psi,\nu}^{-1}\left[-\nu\sin\theta\, f_0'(\theta+\nu\sin\theta)+(1+\nu\cos\theta)^2 f_0''(\theta+\nu\sin\theta)\right]$$

and $f_0''(\pi/2)<0$ that

$$\begin{aligned}f_{\nu_1}''(\theta_{\nu_1}^*)&=d_{\kappa,\psi,\nu_1}^{-1}(1+\nu_1\cos\theta_{\nu_1}^*)^2 f_0''(\pi/2)\\&>d_{\kappa,\psi,\nu_2}^{-1}(1+\nu_2\cos\theta_{\nu_2}^*)^2 f_0''(\pi/2)=f_{\nu_2}''(\theta_{\nu_2}^*).\end{aligned}$$

Note that $1+\nu_1\cos\theta_{\nu_1}^*<1+\nu_2\cos\theta_{\nu_2}^*$ and $1-\nu\cos\theta_{-\nu}^*=1+\nu\cos\theta_\nu^*$. Similarly, we have $f_{\nu_1}''(-\theta_{\nu_1}^*)<f_{\nu_2}''(-\theta_{\nu_2}^*)$.

When $\psi>1$, the results follow if $d_{\kappa,\psi,\nu}/(1+\nu\cos\theta_\nu^*)^2$ is a monotone decreasing function of ν. This is proved as follows: let $0<\nu\le 1$. Note that $\theta_\nu^{*\prime}=-\sin\theta_\nu^*/(1+\nu\cos\theta_\nu^*)$ and consider the derivative of $d_{\kappa,\psi,\nu}/(1+\nu\cos\theta_\nu^*)^2$ with respect to ν, i.e.,

$$\begin{aligned}&-\frac{2(\nu+\cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^4}d_{\kappa,\psi,\nu}+\frac{1}{(1+\nu\cos\theta_\nu^*)^2}d'_{\kappa,\psi,\nu}\\&=\frac{1}{(1+\nu\cos\theta_\nu^*)^2}\int_0^\pi\left(-\frac{2(\nu+\cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2}-\frac{\nu+\cos\theta}{(1+\nu\cos\theta)^2}\right)\\&\quad\times(\{1+\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi}+\{1-\tanh(\kappa\psi)\sin H_\nu(\theta)\}^{1/\psi})\mathrm{d}\theta\\&=\frac{1}{(1+\nu\cos\theta_\nu^*)^2}\int_0^\pi\left\{-\frac{2(\nu+\cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2}-\frac{\nu+\cos H_\nu^{-1}(t)}{(1+\nu\cos H_\nu^{-1}(t))^2}\right\}\\&\quad\times[\{1+\tanh(\kappa\psi)\sin t\}^{1/\psi}+\{1-\tanh(\kappa\psi)\sin t\}^{1/\psi}]\frac{\mathrm{d}t}{1+\nu\cos H_\nu^{-1}(t)}.\end{aligned}$$

In the last equality, we made use of the change of variable $t=H_\nu(\theta)$. Now define the function $s(u)$ as

$$s(u)=-\frac{2(\nu+\cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2}-\frac{\nu+\cos u}{(1+\nu\cos u)^2}.$$

Then there exists a unique $\delta\in(0,\pi)$ such that $s(\delta)=0$. Noting that $\theta_\nu^*+\theta_{-\nu}^*=\pi$, we have

$$s(\theta_{-\nu}^*)=-\frac{2(\nu+\cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2}-\frac{\nu-\cos\theta_\nu^*}{(1-\nu\cos\theta_\nu^*)^2}.$$

Since $\nu - \cos\theta_\nu^*$ is a monotone decreasing function of ν and its minimum is attained at 0, we obtain $s(\theta_{-\nu}^*) < 0$. Moreover, we have $s(0) < 0$ and

$$s(\pi) > -2(\nu + \cos\theta_\nu^*) + \frac{1}{1-\nu} > -4\nu + \frac{1}{1-\nu} = \frac{(2\nu-1)^2}{1-\nu} \geq 0.$$

From these facts, we have $(\pi/2 <)\ \theta_{-\nu}^* < \delta$. On the other hand,

$$\begin{aligned}
&\int_0^\pi \left(-\frac{2(\nu + \cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2} - \frac{\nu + \cos H_\nu^{-1}(t)}{(1+\nu\cos H_\nu^{-1}(t))^2} \right) \frac{\mathrm{d}t}{1+\nu\cos H_\nu^{-1}(t)} \\
&= \int_0^\pi \left(-\frac{2(\nu + \cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2} - \frac{\nu + \cos\theta}{(1+\nu\cos\theta)^2} \right) \mathrm{d}\theta \\
&= \int_0^\pi s(\theta)\mathrm{d}\theta = -\frac{2\pi(\nu + \cos\theta_\nu^*)}{(1+\nu\cos\theta_\nu^*)^2} < 0,
\end{aligned}$$

and taking into account the fact that $\{1+\tanh(\kappa\psi)\sin t\}^{1/\psi}$ and $\{1-\tanh(\kappa\psi)\sin t\}^{1/\psi}$ are symmetric about $\pi/2$, we see that the derivative of $d_{\kappa,\psi,\nu}/(1+\nu\cos\theta_\nu^*)^2$ with respect to ν is negative. Therefore, $d_{\kappa,\psi,\nu}/(1+\nu\cos\theta_\nu^*)^2$ is a monotone decreasing function of ν and the results hold. □

3.2.3 Properties of the AEC distribution

In what follows we describe the properties of the subclass of the AEJP family formed by the Papakonstantinou asymmetric extended cardioid distributions with density (8). Throughout we assume, without loss of too much generality, that $\mu = 0$. In Sect. 3.2.4 we consider method of moments estimation for the parameters of an AEC distribution.

Clearly, (8) tends to the circular uniform density as $k \to 0$ and to a cardioid density as $\nu \to 0$.

Abe et al. (2009) provide the Fourier series expansion of Papakonstantinou's symmetric extended cardioid density (6) expressed in terms of Bessel functions of the first kind (see (7)). To obtain the Fourier series expansion of (8) we make use of the fact that $\sin(\theta + \nu\sin\theta)$ can be represented in terms of Bessel functions of the first kind as

$$\begin{aligned}
\sin(\theta + \nu\sin\theta) &= \sum_{p=1}^{\infty} \{J_{p-1}(\nu) + (-1)^p J_{p+1}(\nu)\} \sin p\theta \\
&= \sum_{p=1}^{\infty} \left\{ 2J'_{2p-1}(\nu)\sin(2p-1)\theta + \frac{4p}{\nu} J_{2p}(\nu)\sin 2p\theta \right\}.
\end{aligned}$$

The second equality in the above is obtained using the first and second formulae in (Abramowitz and Stegun 1972, p. 361, Eq. 9.1.27). Hence,

$$f_{AEC}(\theta) = \frac{1}{2\pi}\{1 + k\,\sin(\theta + \nu\sin\theta)\}$$
$$= \frac{1}{2\pi} + \frac{k}{\pi}\sum_{p=1}^{\infty}\left\{J'_{2p-1}(\nu)\sin(2p-1)\theta + \frac{2p}{\nu}J_{2p}(\nu)\sin 2p\theta\right\}.$$

For $p = 1, 2, \ldots$, the pth cosine and sine moments of a random variable Θ with an AEC distribution, $\alpha_p = E(\cos p\Theta)$ and $\beta_p = E(\sin p\Theta)$, are therefore given by

$$\alpha_p = 0, \quad \beta_{2p-1} = kJ'_{2p-1}(\nu), \quad \beta_{2p} = \frac{2pk}{\nu}J_{2p}(\nu),$$

from which the pth trigonometric moment, $\phi_p = E(e^{ip\Theta})$, follows immediately. Writing $\phi_p = \rho_p e^{i\mu_p^0}$ ($\rho_p > 0$), then $\mu_p^0 = \arg(\alpha_p + i\beta_p)$, with $\mu^0 = \mu_1^0 = \pi/2$ as a special case. Setting $\bar{\phi}_p = E(e^{ip(\Theta-\mu^0)})$, $\bar{\alpha}_p = E[\cos p(\Theta - \mu^0)]$ and $\bar{\beta}_p = E[\sin p(\Theta - \mu^0)]$, we obtain $\bar{\phi}_p = (-i)^p\phi_p$, $\bar{\alpha}_{2p-1} = (-1)^{p-1}kJ'_{2p-1}(\nu)$, $\bar{\alpha}_{2p} = 0$, $\bar{\beta}_{2p-1} = 0$, $\bar{\beta}_{2p} = (-1)^p 2pkJ_{2p}(\nu)/\nu$.

The mean resultant length is given by $\rho = |\phi_1| = kJ'_1(\nu)$. For $0 \le k \le 1$ and $-1 \le \nu \le 1$, ρ ranges from 0 to 0.5 with its maximum value occurring when $\nu = 0$. The circular variance, V, and circular standard deviation, σ, are thus

$$V = 1 - \rho = 1 - kJ'_1(\nu)$$

and

$$\sigma = \{-2\log(1 - V)\}^{1/2} = \left\{-2\log kJ'_1(\nu)\right\}^{1/2}.$$

Finally, the circular skewness, γ_1, and circular kurtosis, γ_2, are given by

$$\gamma_1 = \frac{\bar{\beta}_2}{V^{3/2}} = -\frac{2kJ_2(\nu)}{\nu\{1 - kJ'_1(\nu)\}^{3/2}}$$

and

$$\gamma_2 = \frac{\bar{\alpha}_2 - \rho^4}{V^2} = -\frac{\rho^4}{(1-\rho)^2} = -\frac{k^4\{J'_1(\nu)\}^4}{\{1 - kJ'_1(\nu)\}^2}.$$

Note that the circular kurtosis is always negative.

3.2.4 Method of moments estimation for the AEC distribution

Method of moments (MM) estimation for the three parameters of Papakonstantinou's AEC distribution proceeds by first estimating the mean direction, μ, by $\bar{\theta} - \pi/2 \pmod{2\pi}$, where $\bar{\theta}$ is the sample mean direction (see, for example, Mardia and Jupp 1999, p. 15 for a formal definition of $\bar{\theta}$). Then k and ν can be estimated by equating the first two sample mean resultant lengths, $\bar{R}_1 = n^{-1}\{(\sum_{i=1}^n \cos\theta_i)^2 + (\sum_{i=1}^n \sin\theta_i)^2\}^{1/2}$

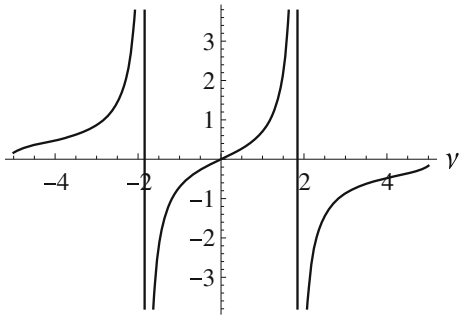

Fig. 5 The function $2J_2(\nu)/(\nu J_1'(\nu))$ for $-5 \le \nu \le 5$. Note that $2J_2(1)/J_1'(1) = 0.7068$

and $\bar{R}_2 = n^{-1}\sum_{i=1}^{n}\{(\sum_{i=1}^{n}\cos 2\theta_i)^2+(\sum_{i=1}^{n}\sin 2\theta_i)^2\}^{1/2}$, with their corresponding population moments, $\rho_1 = kJ_1'(\nu)$ and $\rho_2 = 2kJ_2(\nu)/\nu$ and solving for k and ν.

Equating ρ_2/ρ_1 with $\bar{R}_2/\bar{R}_1$, a method of moments estimate for ν, $\tilde{\nu}$, is obtained as the solution to the equation $2J_2(\tilde{\nu})/(\tilde{\nu}J_1'(\tilde{\nu})) = \bar{R}_2/\bar{R}_1$. From a consideration of the behaviour of the odd function $2J_2(\nu)/(\nu J_1'(\nu))$, portrayed in Fig. 5, there will be a unique solution for $\tilde{\nu}$ if $\tilde{\nu}$ is restricted to taking values in $(-1.8412, 1.8412)$. Moreover, the MM solution will correspond to an underlying unimodal distribution (with $|\tilde{\nu}| \le 1$) if $|\bar{R}_2/\bar{R}_1| \le 0.7068$.

Equating ρ_1 with $\bar{R}_1$, substituting $\tilde{\nu}$ for ν and solving for k, a method of moments estimate for k is given by $\tilde{k} = \bar{R}_1/J_1'(\tilde{\nu})$. Note that, for $\nu \in (-1.8412, 1.8412)$, the minimum value of $J_1'(\nu) = (J_0(\nu) - J_2(\nu))/2$ is 0. Thus, as $0 \le \bar{R}_1 \le 1$, $\tilde{k}$ must be non-negative. However, $\tilde{k}$ can take values greater than 1.

The method of moments estimates of the parameters of an assumed underlying SEC distribution can provide useful starting values for maximum likelihood estimation, which we consider next.

3.2.5 Maximum likelihood estimation for the AEJP distribution

Let $\theta_1, \ldots, \theta_n$ denote a random sample of size n drawn from an AEJP distribution with density (16). We assume here also that all four of the parameters are unknown. The log-likelihood function is

$$\ell(\mu, \kappa, \psi, \nu) = -n \log d_{\kappa,\psi,\nu} + \frac{1}{\psi}\sum_{i=1}^{n}\log\{1 + \tanh(\kappa\psi)\sin(\theta_i - \mu + \nu\sin(\theta_i - \mu))\}. \quad (17)$$

As was the case for the SEJP distribution, the first- and second-order partial derivatives of (17) are rather involved and we do not reproduce them here. Once more, we would recommend the use of symbolic mathematical computing packages like *Sage* or *Mathematica* to those interested in obtaining them.

For a random sample of AEJP data grouped into the m class intervals $[\theta_0, \theta_1)$, $[\theta_1, \theta_2)$,...,$[\theta_{m-1}, \theta_m)$, where $\theta_0 = \theta_m$, with n_j observations in the jth interval and thus a total of $n = n_1 + n_2 + \cdots + n_m$ observations, the log-likelihood function is given by

$$\ell^{\dagger}(\mu,\kappa,\psi,\nu) = -n\log d_{\kappa,\psi,\nu} + \sum_{j=1}^{m} n_j \log \int_{\theta_{j-1}}^{\theta_j} \{1+\tanh(\kappa\psi)\sin(\theta-\mu+\nu\sin(\theta-\mu))\}^{1/\psi}\mathrm{d}\theta. \quad (18)$$

As for the SEJP model, in general numerical methods must be used to identify the maximum likelihood estimates. In the Appendix we provide the R scripts RS3 and RS4 used to perform maximum likelihood based inference for the second grouped data set analysed in Sect. 4. Again, scripts for use with continuous data and for calculating the profile log-likelihood functions of the other three parameters, are available from the authors upon request.

4 Illustrative examples

Mooney et al. (2003) provide the monthly totals of sudden infant death syndrome (SIDS) cases in England, Wales, Scotland and Northern Ireland for the years 1983–1998 and fit circular models to the data for each year. This period covers the introduction of the 'Back to Sleep' campaign in the early 1990s which led to a considerable reduction in the number of SIDS cases. Here we present analyses of the data for the years 1989 and 1985 which illustrate the application of the SEJP and AEJP distributions, respectively. When fitting both models we allowed for grouping and different month lengths using the log-likelihood functions (14) and (18). Four R scripts used in fitting the two models are presented in the Appendix.

4.1 Fitting the SEJP distribution to the 1989 SIDS data

In 1989 there was a total of 1526 SIDS cases. Positioning each case at the centre of the month in which it occurred, the p-value for Pewsey's large-sample test for circular reflective symmetry was 0.97. Also, the histogram of the data portrayed in Fig. 6 suggests the underlying distribution to be unimodal. Thus, we use the 1989 data to illustrate how unimodal members of the SEJP family and its SEC, SEvM, SEWC and JP subclasses can be fitted using constrained maximum likelihood estimation. As can be appreciated from the R script RS1 in the Appendix, the unimodality constraint, i.e. $-1 \leq \nu \leq 1$, is imposed through the `lower` and `upper` modifiers of `optim`'s L-BFGS-B method of optimisation.

Results for the five unimodal fits are presented in Table 1. We consider the use of chi-squared goodness-of-fit testing to be justified in this context because (Mooney et al. 2003) data reporting unequivocally determines 12 class intervals and there is no need to combine class intervals because the expected number of cases in each of them is far in excess of 5. The p-values for the chi-squared goodness-of-fit test indicate that all five submodels provide quite adequate fits to the data.

As the maximum likelihood estimate of ν occurs on the boundary of the parameter space constrained to ensure unimodality (i.e. with $-1 \leq \nu \leq 1$), `optim` is unable to compute the Hessian matrix. Thus, instead of using standard asymptotic normal theory to calculate nominally 95 % confidence intervals for the individual parameters we

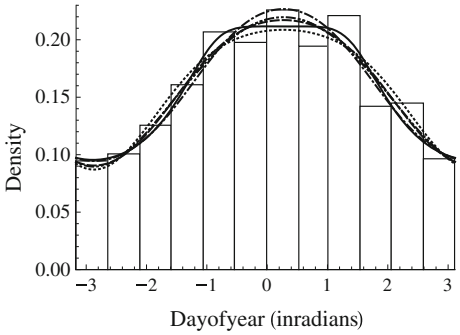

Fig. 6 Linear histogram of the monthly numbers of SIDS cases in 1989, with the day of the year represented in radians lying between −3.167 (start of July) and 3.116 (end of June). Each *bar* represents a month, its width being proportional to the length of that month in days. The superimposed densities correspond to the best fitting unimodal members of the SEJP family (*solid*) and its SEC (*dot*), SEvM (*dash*), SEWC (*dash-dot*) and JP (*dash-dot-dot*) subclasses

Table 1 Maximum likelihood estimates (MLEs) for the fits to the 1989 SIDS data of unimodal members of the SEJP family and its SEC ($\psi = 1$), SEvM ($\psi = 0$), SEWC ($\psi = -1$) and JP ($\nu = 0$) submodels

Distribution	MLE				MLL	AIC	p (g-o-f)
	μ	κ	ψ	ν			
SEJP	0.25	0.40	−4.05	−1.00	−3728.73	7465.45	0.21
SEC	0.27	0.44	1	−0.16	−3729.02	7464.04	0.25
SEvM	0.27	0.44	0	−0.38	−3728.94	7463.88	0.26
SEWC	0.27	0.43	−1	−0.58	−3728.87	7463.75	0.27
JP	0.27	0.45	1.64	0	−3729.07	7464.15	0.24

The maximised log-likelihood (MLL) and AIC values, and p-value for the chi-squared goodness-of-fit test, are included as fit diagnostics. BIC values have not been included because the data are grouped

employed profile log-likelihood based methods. The script RS2 of the Appendix was the one we used to compute the confidence interval for ν. It is based on the large-sample theory of Self and Liang (1987) for the limiting distribution of the maximum likelihood estimator when the true parameter value may be on the boundary of the parameter space. According to that theory, the asymptotic distribution of the likelihood-ratio test statistic for testing the value of an individual parameter is not the usual chi-squared on 1 degree of freedom obtained using standard asymptotic theory. Instead it is that of $Y = Z^2 I(Z > 0)$, where Z denotes a standard normal random variable and $I(Z > 0)$ is the indicator function which takes the value 1 when $Z > 0$ and 0 otherwise. Trivially then, $P(Y > y) = P(Z^2 > y)P(Z > 0) = \frac{1}{2}P(X > y)$, where $X = Z^2$ is a chi-squared random variable on 1 degree of freedom. Hence, a nominally $100(1-\alpha)\%$ confidence interval for ν is given by the set of those ν-values for which the difference between their profile log-likelihood values and that for the maximum likelihood estimate of ν (on the boundary of the parameter space) is less than $\frac{1}{2}\chi_1^2(2\alpha)$, where $\chi_1^2(2\alpha)$ denotes the upper 2α quantile of the chi-squared distribution with 1 degree of freedom. The nominally 95 % confidence for ν, calculated from its profile

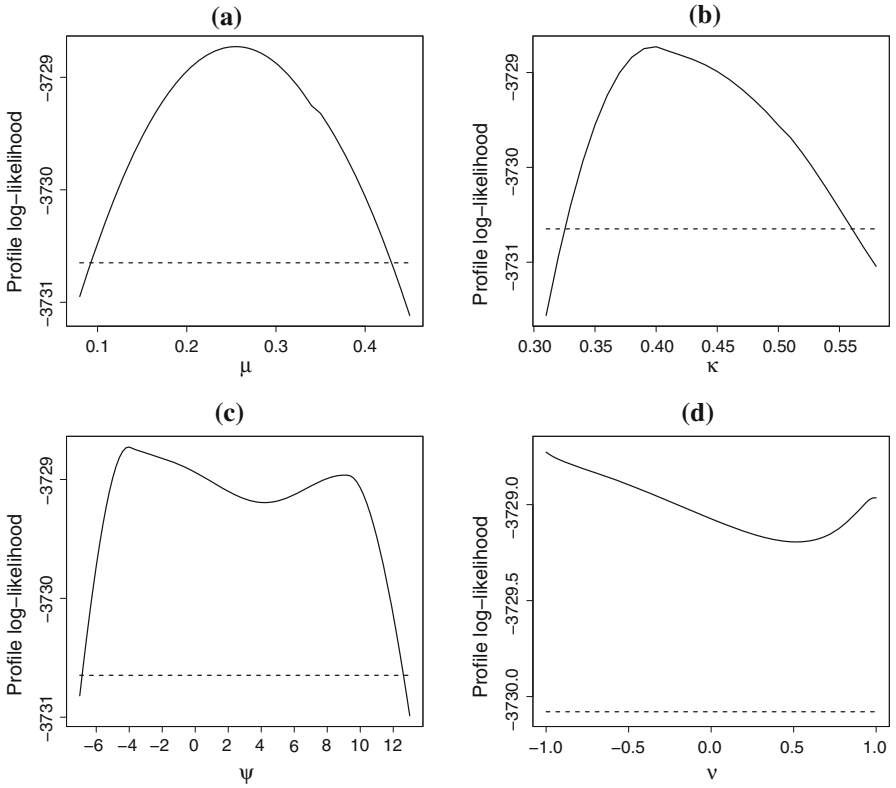

Fig. 7 Profile log-likelihood functions of (**a**) μ, (**b**) κ, (**c**) ψ and (**d**) ν for the 1989 SIDS data. The *dashed horizontal line* in each plot delimits the cut-point for the construction of a nominally 95 % confidence interval for the parameter in question. The confidence interval contains those parameter values with profile log-likelihood values above the cut-point

log-likelihood function displayed in Fig. 7d, is $[-1, 1]$; all the profile log-likelihood values lying above the cut-point are delimited by the dashed horizontal line. Thus, for any ν-value in $[-1, 1]$, there is a unimodal SEJP distribution whose log-likelihood value is not significantly different, at the 5 % significance level, from that of the maximum likelihood unimodal SEJP fit.

The maximum likelihood estimates of the parameters μ, κ and ψ are interior points of the parameter space and hence individual profile log-likelihood based confidence intervals for them can be calculated using the usual chi-squared distribution on 1 degree of freedom for the distribution of the likelihood-ratio test statistic. The R scripts for doing so have much in common with the script RS4 presented in the Appendix. Their nominally 95 % confidence intervals, calculated from their profile log-likelihood functions displayed in Fig. 7, are $(0.09, 0.43)$, $(0.33, 0.56)$ and $(-6.86, 12.61)$, respectively. The relatively wide intervals for ψ and ν reflect the fact that their profile log-likelihood functions are fairly flat over sizeable parts of their ranges. The shape of the profile log-likelihood function for μ is far closer to being quadratic. In keeping with these findings, likelihood-ratio tests fail to identify any significant improvement

in fit of the full SEJP family over any of its four unimodal submodels considered in Table 1.

The AIC identifies the SEWC fit as providing the most parsimonious model for the data. This fit is the one for which the chi-squared goodness-of-fit test is least significant. The densities of the five fits are superimposed upon a linear histogram of the data in Fig. 6. The major differences in fit occur around the peak and in the shoulders. The SEWC density, judged to provide the best fit, is the most peaked of the five flat-topped fits.

4.2 Fitting the AEJP distribution to the 1985 SIDS data

In 1985, again prior to the 'Back to Sleep' campaign, there was a total of 1504 SIDS cases; about the same number as in 1989. Pewsey's test, with a p-value of 0.0001, emphatically rejects an underlying symmetric circular distribution and so we explored the fit of the AEJP family and its AEC, AEvM and AEWC subclasses. Analogous results to those presented previously in Table 1 appear in Table 2. For these data, $\bar{\theta} = 0.1651$, $\bar{R}_1 = 0.2269$ and $\bar{R}_2 = 0.0782$, and the MM estimates for the parameters of the AEC subclass are $\tilde{\mu} = -1.41$, $\tilde{\kappa} = 0.59$ and $\tilde{\nu} = 0.61$, close to their ML counterparts.

The p-values for the chi-squared goodness-of-fit test indicate that only the full AEJP model and its AEC submodel provide adequate fits to the data. Approximate 95 % confidence intervals for the parameters μ, κ, ψ and ν of the full AEJP family, calculated from their individual profile log-likelihood functions together with standard chi-squared theory, are $(-1.56, -1.30)$, $(0.50, 0.74)$, $(0.15, 2.22)$ and $(0.34, 0.83)$, whilst those calculated using the observed information matrix and standard asymptotic normal theory are $(-1.56, -1.30)$, $(0.49, 0.73)$, $(0.15, 2.13)$ and $(0.37, 0.86)$. The similarity of the two sets of confidence intervals reflects the fact that the profile log-likelihoods of the four parameters (not shown) are all close to quadratic in shape. Both intervals for ψ contain the value 1 but neither 0 nor -1, supporting the previous findings regarding the lack of fit of the AEvM and AEWC models. The fit for the full AEJP model is very close to that for its AEC submodel, and the likelihood-ratio test, with a p-value of 0.78, leads to the conclusion that it does not provide a

Table 2 Maximum likelihood estimates (MLEs) for the fits to the 1985 SIDS data of the full AEJP family and its AEC ($\psi = 1$), AEvM ($\psi = 0$) and AEWC ($\psi = -1$) submodels

Distribution	MLE				MLL	AIC	p (g-o-f)
	μ	κ	ψ	ν			
AEJP	−1.43	0.61	1.14	0.61	−3650.22	7308.45	0.07
AEC	−1.43	0.60	1	0.62	−3650.26	7306.53	0.11
AEvM	−1.43	0.55	0	0.63	−3652.71	7311.43	0.02
AEWC	−1.44	0.50	−1	0.58	−3657.53	7321.10	0.00

The maximised log-likelihood (MLL) and AIC values, and p-value for the chi-squared goodness-of-fit test, are included as fit diagnostics. BIC values have not been included because the data are grouped

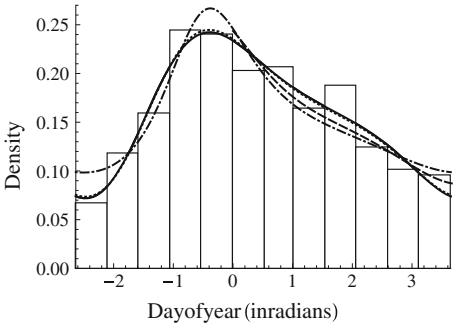

Fig. 8 Linear histogram of the monthly numbers of SIDS cases in 1985, with the day of the year represented in radians lying between −2.634 (start of August) and 3.649 (end of July). Each *bar* represents a month, its width being proportional to the length of that month in days. The superimposed densities correspond to the best fitting members of the AEJP family (*solid*) and its AEC (*dot*), AEvM (*dash*) and AEWC (*dash-dot*) subclasses

significant improvement over the AEC submodel. The AIC also identifies the AEC fit as providing the most parsimonious model for the data. The densities of the four fits are superimposed upon a linear histogram of the data in Fig. 8. This figure highlights the similarity of the full AEJP and AEC fits as well as the difference between them and their AEvM and AEWC counterparts. The R script RS4 of the Appendix was the one used to compute the profile log-likelihood function for ψ. The scripts for computing the profile log-likelihood functions for the other three parameters are very similar to it in structure and have therefore not been reproduced. The code used to carry out the other forms of inference reported above is contained in the R script RS3 of the Appendix.

5 Concluding remarks

Over the years, various extensions of symmetric circular distributions have been proposed within the literature; the motivation for doing so having been to provide distributions with greater modelling flexibility. Here, greater modelling flexibility refers to the ability to model wider ranges of peakedness/flat-toppedness and/or asymmetry and/or multimodality. Of these extensions we would highlight the generalised von Mises models (Maksimov 1967; Gatto and Jammalamadaka 2007), Beran (1979) exponential family models, Azzalini (1985) type perturbation based models (Umbach and Jammalamadaka 2009; Abe and Pewsey 2011) and the models of Kato and Jones (2010) related to Möbius transformation. The generalisation of Papakonstantinou and Batschelet's transformation of argument approach, considered here, provides an alternative. One of its appealing features is that the argument in the base symmetric density is simply replaced by a transformation of that same argument. However, as we have seen, there is a price to be paid for this simplicity, namely that the normalising constant has to be recomputed, generally numerically as usually there will be no closed-form expression for it. Nevertheless, as the R scripts in the Appendix demonstrate, the quadrature involved in its computation is easily carried out in R.

The sine and cosine functions employed in the transformations of argument in (3) and (4), respectively, are appealing because they are periodic, continuous, differentiable everywhere and exist for all θ. In an attempt to build other flexible models, one could try replacing the sine and cosine functions by some other function, say $t(\theta)$. Whilst $t(\theta)$ need not necessarily be periodic itself, to obtain well-behaved circular densities $\cos(\theta + \nu t(\theta))$ should be periodic, continuous, differentiable everywhere and exist for all θ. These four requirements rule out, for instance, the use of the tangent function and all three inverse trigonometric functions. In related work, Jones and Pewsey (2012) propose "inverse Batschelet" distributions based on "transformation of scale". One of their approaches to skewing a symmetric base model has the appeal that the normalising constant remains unchanged, although at the expense of having to invert a function numerically.

In our applications of the transformation of argument approach we chose the Jones–Pewsey family as the base symmetric model. The Jones–Pewsey family is an appealing choice because of its inclusion of numerous well-known symmetric circular distributions. The modelling flexibility of the resulting SEJP and AEJP families is illustrated in the examples of Sect. 4 and by the densities in Figs. 1 and 3. The analyses performed in Sect. 4 show how likelihood-based inference provides a powerful tool for investigating the fit of the two families and their submodels. When defining the AEJP family we constrained the densities within it to being unimodal by imposing the restriction $-1 \leq \nu \leq 1$. Letting ν take values in $(-\infty, \infty)$ further extends the family's flexibility by admitting multimodality but at the expense of ν losing its appealing interpretation as a skewness parameter.

Mooney et al. (2003, 2006) refer to a generic asymmetric annual SIDS case signature with a rapid increase in cases during the autumn, a flat peak over the winter months and a slower decline during the spring and summer. Whilst the AEC fit for the 1985 SIDS data is consistent with this description, the SEWC fit for the 1989 data clearly is not. Moreover, Mooney et al. (2003, 2006) identify bimodal distributions providing best fits to the data for some of the years after the 'Back to Sleep' campaign. As mentioned previously, that campaign led to a pronounced decrease in the numbers of SIDS cases. However, the decrease was more pronounced in some months than it was in others, producing what appears to have been a general tendency towards uniformity in the distribution of SIDS cases.

Appendix: R Scripts

Here we provide four of the R scripts employed when fitting the SEJP family to the 1989 SIDS data, and the AEJP family to the 1985 SIDS data, using likelihood based methods. Both data sets are grouped. For the 1989 SIDS data, the script RS2 for computing the profile log-likelihood function for ν is founded upon asymptotic theory for maximum likelihood estimation on the boundary of the parameter space due to Self and Liang (1987). For the 1985 SIDS data, the script RS4 for computing the profile log-likelihood function for ψ employs standard asymptotic chi-squared theory.

RS1: Maximum likelihood estimation for the unimodal SEJP family: 1989 SIDS grouped data

```
# Data entry and class intervals
breaks=c(0, 0.53218, 1.01564, 1.54928, 2.06570, 2.59935, 3.11577, 3.64941,
4.18305, 4.69948, 5.23312, 5.74954, 6.28319)
freqs=c(184, 143, 180, 112, 118, 76, 77, 82, 99, 131, 163, 161)
nfreq=length(freqs) ; n=sum(freqs)

# Define negative log-likelihood
func <- function(p){
mu = p[1] ; kappa = p[2] ; psi = p[3] ; nu = p[4]
intgrnd1 = function(y){(1+tanh(kappa*psi)*cos(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*cos(y-mu+nu*sin(y-mu)))^(1/psi)}
nll <- 0
for (j in 1:nfreq) {
nll = nll-freqs[j]*log(integrate(intgrnd2,breaks[j],breaks[j+1])$value)}
nll = nll+n*log(q1)
return(nll)
}

# Minimize negative log-likelihood
out <- optim(par=c(0.25,0.40,-4,-0.5), fn=func, gr=NULL, method="L-BFGS-B",
lower=c(-pi,0,-Inf,-1), upper=c(pi,Inf,Inf,1), control=list(), hessian=T)

# Maximized value of log-likelihood
-out$value

# AIC
npar=4 ; 2*(out$value+npar)

# Maximum likelihood estimates
muhat = out$par[1] ; muhat ; kaphat = out$par[2] ; kaphat
psihat = out$par[3] ; psihat ; nuhat = out$par[4] ; nuhat

# Chi-squared goodness-of-fit test
mu = muhat ; kappa = kaphat ; psi = psihat ; nu = nuhat
intgrnd1 = function(y){(1+tanh(kappa*psi)*cos(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*cos(y-mu+nu*sin(y-mu)))^(1/psi)}
chistat <- 0
for (j in 1:nfreq) {
q2 = integrate(intgrnd2,breaks[j],breaks[j+1])$value)
expfreq = n*q2/q1
chistat = chistat + ((freqs[j]-expfreq)^2)/expfreq
}
dfchis = nfreq-npar-1
pval = pchisq(chistat, df=dfchis, lower.tail = FALSE) ; pval
```

RS2: Profile log-likelihood function of ν for the unimodal SEJP family: 1989 SIDS grouped data

```
# Specify NU-values for profile log-likelihood
nuseq=seq(from=-1, to=1, by=0.02) ; nnu=length(nuseq)

# Do loop to compute profile log-likelihood values
pll=0
for (i in 1:nnu) {
nu=nuseq[i]
```

```
# Negative log-likelihood
func <- function(p){
mu = p[1] ; kappa = p[2] ; psi = p[3]
intgrnd1 = function(y){(1+tanh(kappa*psi)*cos(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*cos(y-mu+nu*sin(y-mu)))^(1/psi)}
nll <- 0
for (j in 1:nfreq) {
nll = nll-freqs[j]*log(integrate(intgrnd2,breaks[j],breaks[j+1])$value)}
nll = nll+n*log(q1)
return(nll)
}
if (nu==-1) {muval=0.25 ; kapval=0.4 ; psival = -4.05}
else {muval=out$par[1] ; kapval=out$par[2] ; psival = out$par[3]}

# Minimize negative log-likelihood and compute profile log-likelihood
out <- optim(par=c(muval,kapval,psival), fn=func, gr=NULL, method="L-BFGS-B",
lower=c(-pi,0,-Inf), upper=c(pi,Inf,Inf), control=list(), hessian=T)
pll[i] <- -out$value
}

# Cut-point for confidence interval construction (Self & Liang 1987)
lolim = max(pll)-qchisq(p=0.1, df=1, lower.tail = FALSE)/2 ; lolim

# Output profile log-likelihood
ploglike <- cbind(nuseq, pll) ; ploglike

# Graphical representation of profile log-likelihood
xlolim=nuseq[1] ; xlolim[2]=nuseq[nnu] ; ylolim=lolim ; ylolim[2]=lolim
plot(nuseq, pll, cex=0, lab=c(8,6,10), xlab=expression(nu),
ylab="Profile log-likelihood")
lines(nuseq,pll) ; lines(xlolim,ylolim,lty=2)
```

RS3: Maximum likelihood estimation for the AEJP family: 1985 SIDS grouped data

```
# Specify NU-values for profile log-likelihood
nuseq=seq(from=-1, to=1, by=0.02) ; nnu=length(nuseq)

# Do loop to compute profile log-likelihood values
pll=0
for (i in 1:nnu) {
nu=nuseq[i]

# Negative log-likelihood
func <- function(p){
mu = p[1] ; kappa = p[2] ; psi = p[3]
intgrnd1 = function(y){(1+tanh(kappa*psi)*cos(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*cos(y-mu+nu*sin(y-mu)))^(1/psi)}
nll <- 0
for (j in 1:nfreq) {
nll = nll-freqs[j]*log(integrate(intgrnd2,breaks[j],breaks[j+1])$value)}
nll = nll+n*log(q1)
return(nll)
}
if (nu==-1) {muval=0.25 ; kapval=0.4 ; psival = -4.05}
else {muval=out$par[1] ; kapval=out$par[2] ; psival = out$par[3]}

# Minimize negative log-likelihood and compute profile log-likelihood
out <- optim(par=c(muval,kapval,psival), fn=func, gr=NULL, method="L-BFGS-B",
lower=c(-pi,0,-Inf), upper=c(pi,Inf,Inf), control=list(), hessian=T)
pll[i] <- -out$value
}
```

```
# Cut-point for confidence interval construction (Self & Liang 1987)
lolim = max(pll)-qchisq(p=0.1, df=1, lower.tail = FALSE)/2 ; lolim

# Output profile log-likelihood
ploglike <- cbind(nuseq, pll) ; ploglike

# Graphical representation of profile log-likelihood
xlolim=nuseq[1] ; xlolim[2]=nuseq[nnu] ; ylolim=lolim ; ylolim[2]=lolim
plot(nuseq, pll, cex=0, lab=c(8,6,10), xlab=expression(nu),
ylab="Profile log-likelihood")
lines(nuseq,pll) ; lines(xlolim,ylolim,lty=2)
```

RS4: Profile log-likelihood function of ψ for the AEJP family: 1985 SIDS grouped data

```
# Data entry and class intervals
breaks=c(0, 0.53218, 1.01564, 1.54928, 2.06570, 2.59935, 3.11577, 3.64941,
4.18305, 4.69948, 5.23312, 5.74954, 6.28319)
freqs=c(163, 150, 132, 146, 100, 79, 77, 54, 92, 128, 190, 193)
nfreq=length(freqs) ; n=sum(freqs)

# Define negative log-likelihood
func <- function(p){
mu = p[1] ; kappa = p[2] ; psi = p[3] ; nu = p[4]
intgrnd1 = function(y){(1+tanh(kappa*psi)*sin(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*sin(y-mu+nu*sin(y-mu)))^(1/psi)}
nll <- 0
for (j in 1:nfreq) {
nll = nll-freqs[j]*log(integrate(intgrnd2,breaks[j],breaks[j+1])$value)}
nll = nll+n*log(q1)
return(nll)
}

# Minimize negative log-likelihood
out <- optim(par=c(-1.5,0.5,1,0.5), fn=func, gr=NULL, method="L-BFGS-B",
lower=c(-pi,0,-Inf,-1), upper=c(pi,Inf,Inf,1), control=list(), hessian=T)

# Maximized value of log-likelihood
-out$value

# AIC
npar=4 ; 2*(out$value+npar)

# Observed information matrix and asymptotic correlation matrix
infmat = solve(out$hessian) ; infmat
corrmat = cov2cor(infmat) ; corrmat

# Standard errors of MLEs
standerr = sqrt(diag(infmat)) ; standerr

# Maximum likelihood estimates
muhat = out$par[1] ; muhat ; kaphat = out$par[2] ; kaphat
psihat = out$par[3] ; psihat ; nuhat = out$par[4] ; nuhat

# Individual asymptotic normal theory based nominal 95% confidence intervals
muint=c(muhat-1.96*standerr[1], muhat+1.96*standerr[1]) ; muint
kapint=c(kaphat-1.96*standerr[2], kaphat+1.96*standerr[2]) ; kapint
psiint=c(psihat-1.96*standerr[3], psihat+1.96*standerr[3]) ; psiint
nuint=c(nuhat-1.96*standerr[4], nuhat+1.96*standerr[4]) ; nuint

# Chi-squared goodness-of-fit test
mu = muhat ; kappa = kaphat ; psi = psihat ; nu = nuhat
intgrnd1 = function(y){(1+tanh(kappa*psi)*sin(y+nu*sin(y)))^(1/psi)}
q1 <- integrate(intgrnd1,-pi,pi)$value
intgrnd2 = function(y){(1+tanh(kappa*psi)*sin(y-mu+nu*sin(y-mu)))^(1/psi)}
chistat <- 0
for (j in 1:nfreq) {
q2 = integrate(intgrnd2,breaks[j],breaks[j+1])$value)
expfreq = n*q2/q1
chistat = chistat + ((freqs[j]-expfreq)^2)/expfreq
}
dfchis = nfreq-npar-1
pval = pchisq(chistat, df=dfchis, lower.tail = FALSE) ; pval
```

Acknowledgments We are most grateful to three referees for their careful reading of the paper and suggestions towards improving its content and presentation. Financial support for the research which led to the production of this paper was received by Pewsey in the form of grant MTM2010-16845 from the Spanish Ministry of Science and Education and grant GR10064 from the Junta de Extremadura.

References

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

Abe, T., Pewsey, A., Shimizu, K. (2009). On Papakonstantinou's extension of the cardioid distribution. *Statistics and Probability Letters*, *79*, 2138–2147.

Abramowitz, M., Stegun, I. A. (Eds.). (1972). *Handbook of mathematical functions*. New York: Dover.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*, 171–178.

Batschelet, E. (1981). *Circular statistics in biology*. London: Academic Press.

Beran, R. (1979). Exponential models for directional data. *Annals of Statistics*, *7*, 1162–1178.

Byrd, R. H., Lu, P., Nocedal, J., Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, *16*, 1190–1208.

Gatto, R., Jammalamadaka, S. R. (2007). The generalized von Mises distribution. *Statistical Methodology*, *4*, 341–353.

Gradshteyn, I. S., Ryzhik, I. M. (2007). *Tables of integrals, series, and products* (7th ed.). London: Academic Press.

Jammalamadaka, S. R., SenGupta, A. (2001). *Topics in circular statistics*. Singapore: World Scientific.

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

Jones, M. C., Pewsey, A. (2012). Inverse Batschelet distributions for circular data. *Biometrics*, *68*, 183–193.

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

Maksimov, V. M. (1967). Necessary and sufficient conditions for the family of shifts of probability distributions on continuous bicompact groups (in Russian). *Theoria Verojatna*, *12*, 307–321.

Mardia, K. V., Jupp, P. E. (1999). *Directional statistics*. Chichester: Wiley.

Mooney, J. A., Helms, P. J., Jolliffe, I. T. (2003). Fitting mixtures of von Mises distributions: a case study involving sudden infant death syndrome. *Computational Statistics and Data Analysis*, *41*, 505–513.

Mooney, J. A., Jolliffe, I. T., Helms, P. J. (2006). Modelling seasonally varying data: a case study for sudden infant death syndrome (SIDS). *Journal of Applied Statistics*, *33*, 535–547.

Papakonstantinou, V. (1979). Beiträge zur zirkulären Statistik. Ph.D. thesis, University of Zurich, Switzerland.

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

Pewsey, A., Shimizu, K., de la Cruz, R. (2011). On an extension of the von Mises distribution due to Batschelet. *Journal of Applied Statistics*, *38*, 1073–1085.

Self, S. G., Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *82*, 605–610.

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

Umbach, D., Jammalamadaka, S. R. (2009). Building asymmetry into circular distributions. *Statistics and Probability Letters*, *79*, 659–663.