The Korean Journal of Applied Statistics (2010) DOI: 10.5351/KJAS.2010.23.6.1125
23*(6), 1125–1145*

A Projected Exponential Family for Modeling Semicircular Data

Hyoung-Moon Kim[1]

[1]Department of Applied Statistics, Konkuk University

(Received August 2010; accepted September 2010)

Abstract

For modeling(skewed) semicircular data, we derive a new exponential family of distributions. We extend it to the l-axial exponential family of distributions by a projection for modeling any arc of arbitrary length. It is straightforward to generate samples from the l-axial exponential family of distributions. Asymptotic result reveals that the linear exponential family of distributions can be used to approximate the l-axial exponential family of distributions. Some trigonometric moments are also derived in closed forms. The maximum likelihood estimation is adopted to estimate model parameters. Some hypotheses tests and confidence intervals are also developed. The Kolmogorov-Smirnov test is adopted for a goodness of fit test of the l-axial exponential family of distributions. Samples of orientations are used to demonstrate the proposed model.

Keywords: Uniformly minimum variance unbiased estimator, maximum likelihood estimator, skewed l-axial data, Kolmogorov-Smirnov test, delta method.

1. Introduction

In linear statistics, the exponential family of distributions (Lehmann and Casella, 1998; Lehmann and Romano, 2005; Casella and Berger, 2002) is a big family that contains the continuous families-normal, gamma, χ^2, beta, log-normal and the discrete families-binomial, Poisson, and negative binomial. It has many nice mathematical and statistical properties. Contrary to linear statistics, the exponential family of distributions that can be used for modeling circular, semicircular, and l-axial data is not clearly shown in most textbooks (Fisher, 1993; Jammalamadaka and SenGupta, 2001; Mardia and Jupp, 2000).

Many useful circular models may be generated by a variety of mechanisms from known probability distributions on the real line or on the plane. A few general methods include:

(1) a wrapping method by wrapping a linear distribution around the unit circle

(2) a method through characterizing properties such as maximum entropy

This work was supported by Konkuk University.

[1]Associate Professor, Department of Applied Statistics, Konkuk University, 1 Hwayang-Dong, Gwangjin-Gu, Seoul 143-701, Korea. E-mail: hmkim@konkuk.ac.kr

(3) an offset method

(4) a stereographic projection method that identifies points on the real line with those on the circle circumference

However, none of these methods and models concentrate on the semicircular or the axial data. Sometimes the angular data are given as modulo π. Some examples are as follows:

(1) the long axis of particles in sediments or the optical axis of a crystal(rather than a direction)

(2) orientations of core samples

(3) a sea turtle example that a sea turtle emerges from the ocean in search of a nesting site on dry land. Therefore, we do not need full circular model in these data and is noted by Guardiola (2004) and Jones (1968). Guardiola (2004) proposes a simple projection method to obtain the semicircular normal distribution.

Most of those models are symmetric. Even recent models appearing in Jones and Pewsey (2005), Pewsey *et al.* (2007) are symmetric. Pewsey (2002, 2004) considers the testing of problems where the underlying distribution is reflectively symmetric about an unknown central direction and about a median axis, respectively. Recently some skewed circular models have been developed using a wrapping method by Pewsey (2000, 2006, 2008), and Jammalamadaka and Kozubowski (2003, 2004); however, none of these models concentrate on semicircular data. Note that the exponential family of distributions contains symmetric and skewed distributions. In this sense we need to develop a new exponential family of distributions for modeling(skewed) semicircular data.

This article is organized as follows. Section 2 defines a new exponential family of distributions for modeling l-axial data. Semicircular, circular, and 4-axial exponential family of distributions are obtained as special cases of the l-axial exponential family of distributions. We derive the trigonometric moments of the semicircular exponential family of distributions. We estimate the parameters of the l-axial exponential family of distributions by a maximum likelihood method in Section 3. Some hypothesis tests and confidence intervals are also developed in the same section. Samples of orientations of termite mounds of Amitermes laurensis at 10^{th} site in Cape York Peninsula, North Queensland are employed to demonstrate the proposed model in Section 4. The conclusion is formed in Section 5.

2. A New Exponential Family of Distributions

2.1. Definition

An important family of distributions is the exponential family, defined by probability densities of the form

$$f(x:\gamma) = a(x)b(\gamma)\exp\left(\sum_{j=1}^{k} c_j(x)d_j(\gamma)\right) \tag{2.1}$$

with respect to a σ-finite measure ν over a Euclidean sample space, where $x \in \Re$, $a(x) \geq 0$ and $c_1(x), \ldots, c_k(x)$ are real-valued functions of the observation x(they cannot depend on γ), and $b(\gamma) \geq 0$ and $d_1(\gamma), \ldots, d_k(\gamma)$ are real-valued functions of the possibly vector-valued parameter γ(they cannot depend on x). These include the continuous families-normal, gamma, and beta as well as the discrete families-binomial, Poisson, and negative binomial. The specific form of

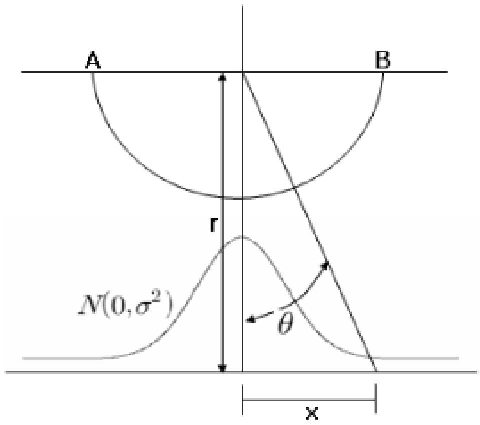

Figure 2.1. Projection from a normal distribution to a semicircle.

(2.1) implies that exponential families have many nice mathematical properties. However, more importantly for a statistical model, the form of (2.1) implies many nice statistical properties, which will be discussed throughout the remainder of the article.

To deal with angular data, we concern about the continuous exponential family of distributions. We concentrate on $k = 1$ and call it the one parameter exponential family(OPEF) since usually a location parameter is set to 0 and the following projection(transformation) is applied. Let $x = r\tan(\theta)$, then $dx = r\sec^2(\theta)d\theta$ and the probability density function(it pdf) of θ is given by

$$f(\theta;\gamma) = a^*(r\tan(\theta))b(\gamma)\exp\left\{c(r\tan(\theta))d(\gamma)\right\},$$

where $a^*(r\tan(\theta)) = r\sec^2(\theta)a(r\tan(\theta))$. See the Figure 2.1 to get an intuition of the projection(transformation). This is the projection from a normal distribution, $N(0,\sigma^2)$, to a semicircle. If a linear random variable X has a support on $\Re$, then θ has a support on $(-\pi/2, \pi/2)$. The support of X is $\Re^+$, then the support of θ is $(0, \pi/2)$. These mean that, after the projection is applied, we can handle semicircular data if support of X is $\Re$ and we can deal with 4-axial data when support of X is $\Re^+$.

Hence, we need to extend it to l-axial data. Occasionally, measurements result in any arc of arbitrary length, say $2\pi/l, l \in \mathbb{N}$, where $\mathbb{N}$ denotes a set of natural number. Apply $\theta^* = 2\theta/l$ for $(-\pi/2, \pi/2)$ and apply $\theta^* = 4\theta/l$ for $(0, \pi/2)$, then the pdfs of θ^* are given by

$$f(\theta^*;\gamma) = a^*\left(r\tan\frac{l\theta^*}{c_p}\right)b(\gamma)\exp\left\{c\left(r\tan\frac{l\theta^*}{c_p}\right)d(\gamma)\right\}, \quad l \in \mathbb{N}, \tag{2.2}$$

where $a^*(r\tan(l\theta^*/c_p)) = r\sec^2(l\theta^*/c_p)a(r\tan(l\theta^*/c_p))$, and $c_p, p = 1, 2$, are $c_1 = 2$ and $c_2 = 4$ depending on the transformation. Notation c_p will be used throughout the remainder of this article. Furthermore $\theta^* \in (-\pi/l, \pi/l)$ for X on $\Re$ and $\theta^* \in (0, \pi/l)$ for X on $\Re^+$. Both projection methods now can handle any arc of arbitrary length say $2\pi/l$ for $l \in \mathbb{N}$. When $l = 1$, (2.2) is a circular OPEF. If $l = 2$, then (2.2) is a semicircular OPEF. (2.2) is a 4-axial OPEF when $l = 4$.

Note that r is not a parameter. It is a known constant since, geometrically it is the distance between the radius and the support of the OPEF density. Therefore, without loss of generality we can assume that $r = 1$ if necessary. Sometimes we can absorb the effect of r into the γ together making a new parameter. For example, the l-axial normal distribution and the semicircular Laplace distribution have $\varphi = \sigma/r$ commonly. See the following examples.

For above *pdf* (2.2), we introduce a location parameter $\mu \in (-\pi, \pi)$. So plug in $\theta^* - \mu$ instead of θ^*, the *pdf* is as follows:

$$f(\theta^*; \mu, \gamma) = a^* \left(r \tan \frac{(\theta^* - \mu)}{c_p} \right) b(\gamma) \exp \left\{ c \left(r \tan \frac{l(\theta^* - \mu)}{c_p} \right) d(\gamma) \right\}, \tag{2.3}$$

where $a^*(r \tan(l(\theta^* - \mu)/c_p)) = r \sec^2(l(\theta^* - \mu)/c_p) a(r \tan(l(\theta^* - \mu)/c_p))$ and $l \in \mathbb{N}$. Now we can handle any l-axial data with a location parameter. This is the reason that we assume a location parameter 0 for linear exponential family and concentrate on OPEF. Unfortunately this family of distributions with a location parameter μ is not a member of exponential family.

2.2. Some basic properties

We consider some basic properties of l-axial exponential family of distributions. The cumulative distribution function(CDF) of l-axial exponential family is given by

$$F_{\theta^*}(\theta^*) = F_X \left(r \tan \left(\frac{l(\theta^* - \mu)}{c_p} \right) \right), \tag{2.4}$$

using the projection method, where $F_X(\cdot)$ is the cdf of a linear random variable X. Simulation from the l-axial exponential family is also straightforward as follows:

$$\theta^* = \mu + \frac{c_p}{l} \tan^{-1} \left(\frac{x}{r} \right) \tag{2.5}$$

by inverting the transformation. So we first generate samples from a linear random variable X and then use the stochastic relationship (2.5).

The following Lemma (Gradshteyn and Ryzhik, 2007) will be used continuously.

Lemma 2.1.

$$\tan(x) = \sum_{k=1}^{\infty} \frac{2^{2k}(2^{2k} - 1)}{(2k)!} |B_{2k}| x^{2k-1}, \quad x^2 < \frac{\pi^2}{4},$$
$$\sec(x) = \sum_{k=0}^{\infty} \frac{|E_{2k}|}{(2k)!} x^{2k}, \quad x^2 < \frac{\pi^2}{4},$$

where the number B_n, representing the coefficients of $t^n/n!$ in the expansion of the function

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}, \quad 0 < |t| < 2\pi,$$

are called Bernoulli numbers. In addition, the numbers E_n, representing the coefficients of $t^n/n!$ in the expansion of the function

$$\frac{1}{\cosh t} = \sum_{n=0}^{\infty} E_n \frac{t^n}{n!}, \quad |t| < \frac{\pi}{2},$$

are known as the Euler numbers.

Asymptotic distribution of l-axial exponential family for sufficiently small γ is given by

$$a^{**}(r\gamma y_p)b(\gamma)\exp\{c(r\gamma y_p)d(\gamma)\},$$

where $a^{**}(r\gamma y_p) = c_p a(r\gamma y_p)r\gamma/l$ and $y_p = l(\theta^* - \mu)/c_p\gamma$, $p = 1, 2$. Support of y_1 is $\Re$ and support of y_2 is $\Re^+$. Note that this is a member of OPEF most of the cases since $a^{**}(r\gamma y_p)$ and $c(r\gamma y_p)$ can be separated as a function of y_p and a function of γ when μ is known. So a linear OPEF can approximate the l-axial OPEF for sufficiently small γ. This can be done first by $y_p = l(\theta^* - \mu)/(c_p\gamma)$, $p = 1, 2$ and then by Lemma 2.1 if we use only up to the first order terms. For all following examples, $a^{**}(r\gamma y_p)$ and $c(r\gamma y_p)$ can be separated a function of y_p and a function of γ. Furthermore the distributions of all examples are members of OPEF when μ is known. Some examples are as follows:

Example 2.1. The *pdf* of l-axial normal(LAN) distribution is given by

$$f(\theta^*; \mu, \varphi) = \frac{l}{2\sqrt{2\pi}\varphi}\sec^2\left(\frac{l(\theta^* - \mu)}{2}\right)\exp\left(-\frac{\tan^2(l(\theta^* - \mu)/2)}{2\varphi^2}\right),$$

where $-\pi/l + \mu < \theta^* < \pi/l + \mu$, $\varphi = \sigma/r$ and $-\pi < \mu < \pi$. This *pdf* is derived using the semicircular normal(SCN) distribution (Guardiola, 2004) and the transformation $\theta^* = 2\theta/l + \mu$. We shall say that θ^* follows LAN(μ, φ^2). Note that $l = 1$ gives us the circular normal(CN) distribution, $l = 2$ suggests the SCN distribution, and $l = 4$ is the 4-axial normal(4AN) distribution. Asymptotic distribution of LAN distribution for sufficiently small φ is the standard normal distribution after the transformation $Y = l(\theta^* - \mu)/2\varphi$ and by Lemma 2.1 if we use only up to the first order terms.

Example 2.2. The l-axial Laplace(LAL) distribution (Ahn and Kim, 2008) is defined by

$$f(\theta^*; \mu, \varphi) = \frac{l}{4\varphi}\sec^2\frac{l(\theta^* - \mu)}{2}\exp\left(-\frac{|\tan(l(\theta^* - \mu)/2)|}{\varphi}\right), \tag{2.6}$$

where $\varphi = \sigma/r$, $-\pi/l + \mu < \theta^* < \pi/l + \mu$, $-\pi < \mu < \pi$. Then, we say that θ is an LAL random variable with parameters μ and φ; for brevity, we shall also say that θ is LAL(μ, φ). Note that when $l = 2$, the *pdf* (2.6) is the semicircular Laplace(SCL) *pdf*. When $l = 1$, it becomes the *pdf* of a circular Laplace(CL) distribution. $l = 4$ is the case of 4-axial Laplace(4AL) *pdf*. We consider the asymptotic behavior of the LAL distribution when $\varphi \to 0$. Suppose θ follows LAL(μ, φ). Let $Y = l(\theta^* - \mu)/2\varphi$, and then use the change of variable technique. For sufficiently small φ, by Lemma 2.1 with only up to the first order terms, the distribution of Y becomes Laplace$(0, 1)$. So, for sufficiently small φ, the LAL distribution can be approximated by a (linear) Laplace distribution.

Example 2.3. We derive the *pdf* of l-axial Gamma(LAG) distribution for known δ as follows:

$$f(\theta^*; \mu, \varphi) = \frac{l\sec^2(l(\theta^* - \mu)/4)}{4\Gamma(\delta)\varphi^\delta}\left\{\tan\frac{l(\theta^* - \mu)}{4}\right\}^{\delta - 1}\exp\left(-\frac{\tan(l(\theta^* - \mu)/4)}{\varphi}\right), \tag{2.7}$$

where $\varphi = \beta/r$, $\mu < \theta^* < 2\pi/l + \mu$ and $-\pi < \mu < \pi$. Then, we say that θ^* follows LAG(μ, φ). Similarly, $l = 1$ gives us the circular gamma(CG) distribution. $l = 2$ suggest the semicircular Gamma(SCG) distribution. The 4-axial gamma(4AG) distribution is obtained when $l = 4$. An asymptotic distribution when $\varphi \to 0$ is $\Gamma(\delta, 1)$ after the transformation $Y = l(\theta^* - \mu)/4\varphi$ and by Lemma 2.1 if we use only up to the first order terms. Hence, linear gamma distribution can be used to approximate the LAG distribution. This distribution is also skewed to the right. Note that when $\delta = 1$, this LAG distribution contains the l-axial exponential distributions as a special case.

Remark 2.1. For unknown δ, (2.7) is still a member of an exponential family when μ is known. For $\delta = \nu/2$ and $\beta = 2$, (2.7) becomes the *pdf* of l-axial χ^2 distribution. However $\delta = \nu/2$ is not known in this case.

Example 2.4. The *pdf* of l-axial χ^2 distribution(LAC2) is defined by

$$f(\theta^*; \mu, \varphi) = \frac{lr\sec^2(l(\theta^* - \mu)/4)}{4\Gamma\left(\nu/2\right) 2^{\frac{\nu}{2}}} \left\{ r\tan\frac{l(\theta^* - \mu)}{4} \right\}^{\frac{\nu}{2}-1} \exp\left(-\frac{r\tan(l(\theta^* - \mu)/4)}{2} \right), \tag{2.8}$$

where $\mu < \theta^* < 2\pi/l + \mu, \ -\pi < \mu < \pi$. Then, we say that θ^* is an LAC2 random variable with parameters μ and ν; for brevity, we shall also say that θ^* follows LAC2(μ, ν). Note that $l = 1$ gives us the circular χ^2(CC2) distribution, $l = 2$ suggests the semicircular χ^2(SCC2) distribution, and $l = 4$ is the 4-axial χ^2(4AC2) distribution. We consider the asymptotic behavior of an LAC2(μ, ν) when ν goes to 0. For the density (2.8), take a transformation $Y = l(\theta^* - \mu)/4\nu$. And then, for sufficiently small ν, by Lemma 2.1 if we use only up to the first order terms, the *pdf* of Y is given by

$$f(y; \nu) = \frac{r\nu}{\Gamma\left(\nu/2\right) 2^{\frac{\nu}{2}}} (r\nu y)^{\frac{\nu}{2}-1} \exp\left(-\frac{r\nu y}{2} \right).$$

Therefore the density of Y is approximately $\chi^2(\nu)/(r\nu)$ which means that the LAC2 distribution can be approximated by 'linear' χ^2 distribution for sufficiently small ν. Note that this distribution is obviously skewed to the right, so we can handle skewed l-axial data.

2.3. Trigonometric moments

Unlike usual linear distributions, we need to derive the trigonometric moments of the l-axial exponential family. Similar to those of any circular density, trigonometric moments of l-axial distribution are defined as follows: $\phi_p = Ee^{ip\theta^*} = \alpha_p + i\beta_p = E\cos(p\theta^*) + iE\sin(p\theta^*)$, $p = 0, \pm 1, \pm 2, \ldots$. In particular, two functions of the first trigonometric moments play the most prominent role defined as

$$\rho = \sqrt{\alpha_1^2 + \beta_1^2} \quad \text{and} \quad \tau = \arctan^*\left(\frac{\beta_1}{\alpha_1}\right),$$

where

$$\tau = \arctan^*\left(\frac{\beta_1}{\alpha_1}\right) = \begin{cases} \tan^{-1}\dfrac{\beta_1}{\alpha_1}, & \text{if} \quad \alpha_1 > 0, \ \beta_1 \geq 0, \\ \dfrac{\pi}{2}, & \text{if} \quad \alpha_1 = 0, \ \beta_1 > 0, \\ \tan^{-1}\dfrac{\beta_1}{\alpha_1} + \pi, & \text{if} \quad \alpha_1 < 0, \\ \tan^{-1}\dfrac{\beta_1}{\alpha_1} + 2\pi, & \text{if} \quad \alpha_1 \geq 0, \ \beta_1 < 0, \\ \text{undefined}, & \text{if} \quad \alpha_1 = 0, \ \beta_1 = 0. \end{cases}$$

The length ρ and when it is non-zero, the direction τ are used to provide theoretical or population measures of the concentration and the mean direction of angular data, respectively. It can be seen that ρ lies between 0 and 1. The larger ρ, *i.e.*, the closer it is to 1, the more the concentration towards the mean direction τ. For the following examples, it is easy to derive these measures using obtained trigonometric moments.

To get the trigonometric moments, we take a projection such as $x = \tan(\theta^*)$ and then calculate those moments treating those as usual integration. This is because most books containing integration formulas are described based on real or complex domains instead of angular domains. To do this we need to change $\cos(p\theta^*)$ and $\sin(p\theta^*)$ in terms of $\tan(\theta^*)$. This can be done by the multiple-angle formulas that are,

$$\cos(p\,\theta^*) = \sum_{k=0}^{p} \binom{p}{k} \cos^k(\theta^*) \sin^{p-k}(\theta^*) \cos\left[\frac{(p-k)\pi}{2}\right],$$
$$\sin(p\,\theta^*) = \sum_{k=0}^{p} \binom{p}{k} \cos^k(\theta^*) \sin^{p-k}(\theta^*) \sin\left[\frac{(p-k)\pi}{2}\right],$$

when $p \in \mathbb{N}$. This multiple-angle formulas established only using the Euler formula and binomial theorem.

Lemma 2.2. *Using $x = \tan(\theta^*)$, above multiple-angle formulas are given in terms of x by*

$$\cos(p\,\theta^*) = \sum_{k=0}^{p} \binom{p}{k} c^1_{p-k} x^{p-k} (1+x^2)^{-\frac{p}{2}},$$
$$\sin(p\,\theta^*) = \sum_{k=0}^{p} \binom{p}{k} c^2_{p-k} x^{p-k} (1+x^2)^{-\frac{p}{2}},$$

where $\sin(\theta^*) = x/\sqrt{1+x^2}$, $\cos(\theta^*) = 1/\sqrt{1+x^2}$,

$$\cos\left[\frac{(p-k)\pi}{2}\right] = c^1_{p-k} = \begin{cases} 1, & \text{if } \; p-k = 4m, \\ 0, & \text{if } \; p-k = 2m+1, \\ -1, & \text{if } \; p-k = 4m+2 \end{cases}$$

and

$$\sin\left[\frac{(p-k)\pi}{2}\right] = c^2_{p-k} = \begin{cases} 1, & \text{if } \; p-k = 4m+1, \\ 0, & \text{if } \; p-k = 2m, \\ -1, & \text{if } \; p-k = 4m+3, \end{cases}$$

where $m = 0, 1, 2, \ldots$.

Specific trigonometric moments depend on the forms of the l-axial exponential family of distributions. To find trigonometric moments of SCN distribution, see Guardiola (2004). SCL trigonometric moments can be found at Ahn and Kim (2008). For all these papers, only up-to the fourth (or the second) trigonometric moments are derived so we need to derive a general formula of all the trigonometric moments.

Without loss of generality, we assume that a location parameter μ is 0. In general, the k^{th} cosine moment, $\alpha_k = E\cos(k\theta^*)$, of the semicircular exponential family of distributions is the same as the $2k^{th}$ cosine moment, $\alpha_{2k} = E\cos(2k\theta^*)$, of the 4-axial exponential family of distributions. The k^{th} cosine moment, $\alpha_k = E\cos(k\theta^*)$, of the circular exponential family of distributions is the same as the $2k^{th}$ cosine moment, $\alpha_{2k} = E\cos(2k\theta^*)$, of the semicircular exponential family of distributions. If the l-axial exponential family of distributions is symmetric, then $\beta_p = E\sin(p\theta^*)$, $p = 0, \pm 1, \pm 2, \ldots$, are 0 like any other symmetric circular density. If not, the similar relationship also exist for the sine moments. That is, the k^{th} sine moment, $\beta_k = E\sin(k\theta^*)$, of the semicircular exponential family of distributions is the same as the $2k^{th}$ sine moment, $\beta_{2k} = E\sin(2k\theta^*)$, of the 4-axial exponential family of distributions because of the transformation we use. Furthermore the k^{th} sine moment, $\beta_k = E\sin(k\theta^*)$, of the circular exponential family of distributions is the same as the $2k^{th}$ sine moment, $\beta_{2k} = E\sin(2k\theta^*)$, of the semicircular exponential family of distributions.

Based on the relationship between the trigonometric moments and our goal for developing models for semicircular data, we concentrate on the trigonometric moments of the semicircular exponential family of distributions. First of all, we derive the trigonometric moments of the SCN distribution as follows:

Theorem 2.1. *Let $\theta^* \sim SCN(0, \varphi^2)$, then the trigonometric moments are as follows:*

$$\alpha_p = \frac{1}{\sqrt{2\pi}\varphi} \sum_{k \in R_c} \binom{p}{k} c^1_{p-k} \Gamma\left(\frac{p-k+1}{2}\right) \Psi\left(\frac{p-k+1}{2}, \frac{3-k}{2}; \frac{1}{2\varphi^2}\right),$$

$$\beta_p = 0,$$

$$\alpha_{-p} = \alpha_p, \quad p \in \mathbb{N}. \text{ Furthermore } \alpha_0 = 1,$$

where $\Psi(\alpha,\gamma;z)$ has an integral representation as $(1/\Gamma(\alpha)) \int_0^\infty e^{-zt} t^{\alpha-1} (1+t)^{\gamma-\alpha-1} dt$ (Gradshteyn and Ryzhik; 2007). $\Psi(\alpha,\gamma;z)$ is related to a confluent hypergeometric function as follows:

$$\Psi(\alpha,\gamma;z) = \frac{\Gamma(1-\gamma)}{\Gamma(\alpha-\gamma+1)} \Phi(\alpha,\gamma;z) + \frac{\Gamma(\gamma-1)}{\Gamma(\alpha)} z^{1-\gamma} \Phi(\alpha-\gamma+1, 2-\gamma; z).$$

A confluent hypergeometric function has a second notation ${}_1F_1(\alpha;\gamma;z)$. R_c denote the set of values such that $\{k|k = 0, 1, \ldots, p$ satisfying $\{p-k = 2m,\, m = 0, 1, 2, \ldots\}\}$.

Proof. For cosine moments, we use the transformation $x = \tan(\theta^*)$. So $\cos(p\theta^*)$ can be expressed as a function of x using Lemma 2.2. The integrand is an even function of x when $p-k$ is even so we use this property. When $p-k$ is odd, the integral is 0 since the integrand is an odd function of x. We change the order of the summation and integration as well as apply a transformation, $y = x^2$, then an intermediate expression is

$$\alpha_p = \frac{1}{\sqrt{2\pi}\varphi} \sum_{k \in R_c} \binom{p}{k} c^1_{p-k} \int_0^\infty y^{\frac{p-k-1}{2}} (1+y)^{-\frac{p}{2}} \exp\left(-\frac{y}{2\varphi^2}\right) dy.$$

□

The result follows immediately by the integral formula 9.211.4 (Gradshteyn and Ryzhik, 2007). Since the SCN distribution is symmetric, $\beta_p = 0$. By the property of sine and cosine functions, the remaining results are obvious.

The trigonometric moments of the SCL distribution are as follows:

Theorem 2.2. *Let $\theta^* \sim SCL(0, \varphi)$, then the trigonometric moments are as follows:*

$$\alpha_p = \frac{1}{2\sqrt{\pi}\Gamma\left(\frac{p}{2}\right)\varphi} \sum_{k \in R_c} \binom{p}{k} c^1_{p-k} G^{31}_{13}\left(\frac{1}{4\varphi^2} \,\middle|\, \begin{matrix} \frac{1-p+k}{2} \\ \frac{k-1}{2}, 0, \frac{1}{2} \end{matrix}\right),$$

$$\beta_p = 0,$$

$$\alpha_{-p} = \alpha_p, \quad p \in \mathbb{N}. \text{ Furthermore } \alpha_0 = 1,$$

where $G_{pq}^{mn}\left(x\,\middle|\,{}^{a_1,\ldots,a_p}_{b_1,\ldots,b_q}\right)$ *is called as Meijer's G-function (Gradshteyn and Ryzhik; 2007).* R_c *denote the set of values such that* $\{k|k=0,1,\ldots,p$ satisfying $\{p-k=2m,\ m=0,1,2,\ldots\}\}$.

Proof. For the cosine moments, we use the transformation $x=\tan(\theta^*)$. So $\cos(p\theta^*)$ can be expressed as a function of x using Lemma 2.2. The integrand is an even function of x when $p-k$ is even. The integral is 0 when $p-k$ is odd since the integrand is an odd function of x. We change the order of the summation and integration, so

$$\alpha_p=\frac{1}{\varphi}\sum_{k\in R_c}\binom{p}{k}c^1_{p-k}\int_0^\infty x^{p-k}(1+x^2)^{-\frac{p}{2}}\exp\left(-\frac{x}{\varphi}\right)dx.$$

□

The result follows immediately by the integral formula 3.389.2 (Gradshteyn and Ryzhik, 2007). Since SCL distribution is symmetric, $\beta_p=0$ like any other circular density. By the property of sine and cosine functions, the remaining results are obvious.

We also derive the trigonometric moments of SCG distribution for known δ w.l.o.g. $\mu=0$. Remark that the trigonometric moments still legitimate whether δ is known or not.

Theorem 2.3. *Let* $\theta^*\sim SCG(0,\varphi)$*, then the trigonometric moments are as follows:*

$$\alpha_p=\frac{1}{2\sqrt{\pi}\Gamma(\gamma)\Gamma(p)\varphi^\gamma}\sum_{k=0}^{2p}\binom{2p}{k}c^1_{2p-k}G^{31}_{13}\left(\frac{1}{4\varphi^2},\,\middle|\,\begin{array}{c}1-\dfrac{2p-k+\gamma}{2}\\ \dfrac{k-\gamma}{2},\,0,\,\dfrac{1}{2}\end{array}\right),$$

$$\beta_p=\frac{1}{2\sqrt{\pi}\Gamma(\gamma)\Gamma(p)\varphi^\gamma}\sum_{k=0}^{2p}\binom{2p}{k}c^2_{2p-k}G^{31}_{13}\left(\frac{1}{4\varphi^2}\,\middle|\,\begin{array}{c}1-\dfrac{2p-k+\gamma}{2}\\ \dfrac{k-\gamma}{2},\,0,\,\dfrac{1}{2}\end{array}\right),$$

$$\alpha_{-p}=\alpha_p \textit{ and } \beta_{-p}=-\beta_p,\quad p\in\mathbb{N}. \textit{ Furthermore } \alpha_0=1 \textit{ and } \beta_0=0.$$

Proof. We use a transformation, $\theta=\theta^*/2$ for the SCG$(0,\varphi)$ random variable, then the distribution of θ becomes 4AG$(0,\varphi)$. For cosine moments, we use the transformation $x=\tan(\theta)$. So $\cos(p\theta^*)=\cos(2p\theta)$ can be expressed as a function of x using Lemma 2.2. We change the order of the summation and integration, then

$$\alpha_p=\frac{1}{\Gamma(\gamma)\varphi^\gamma}\sum_{k=0}^{2p}\binom{2p}{k}c^1_{2p-k}\int_0^\infty x^{2p-k+\gamma-1}(1+x^2)^{-p}\exp\left(-\frac{x}{\varphi}\right)dx.$$

□

The result follows immediately by the integral formula 3.389.2 (Gradshteyn and Ryzhik, 2007). Since the SCG distribution is not symmetric, we need to derive sine moments. For the sine moments, we apply a similar approach; then the result follows immediately by the same integral formula. By the property of sine and cosine functions, the remaining results are obvious.

We derive the trigonometric moments of SCC2 distribution similar to Theorem 2.3.

Corollary 2.1. *Let $\theta^* \sim SCC2(0,\nu)$ with, w.l.o.g., $r=1$, then the trigonometric moments are as follows:*

$$\alpha_p = \frac{1}{\Gamma(\nu/2)\Gamma(p)2^{\frac{\nu}{2}+1}\sqrt{\pi}} \sum_{k=0}^{2p} \binom{2p}{k} c^1_{2p-k} G^{31}_{13}\left(\frac{1}{16} \middle| \begin{array}{c} 1-\frac{2p-k}{2}-\frac{\nu}{4} \\ \frac{k}{2}-\frac{\nu}{4},\ 0,\ \frac{1}{2} \end{array} \right),$$

$$\beta_p = \frac{1}{\Gamma(\nu/2)\Gamma(p)2^{\frac{\nu}{2}+1}\sqrt{\pi}} \sum_{k=0}^{2p} \binom{2p}{k} c^2_{2p-k} G^{31}_{13}\left(\frac{1}{16} \middle| \begin{array}{c} 1-\frac{2p-k}{2}-\frac{\nu}{4} \\ \frac{k}{2}-\frac{\nu}{4},\ 0,\ \frac{1}{2} \end{array} \right),$$

$\alpha_{-p} = \alpha_p$, *and* $\beta_{-p} = -\beta_p$, $p \in \mathbb{N}$. *Furthermore* $\alpha_0 = 1$ *and* $\beta_0 = 0$.

Therefore, for example, the first $\alpha_p = E\cos(p\theta^*), p=1$ of SCC2$(0,\nu)$ with, w.l.o.g., $r=1$ is as follows:

$$\alpha_1 = \frac{1}{\Gamma(\nu/2)2^{\frac{\nu}{2}}\sqrt{\pi}} G^{31}_{13}\left(\frac{1}{16} \middle| \begin{array}{c} \frac{1-\nu}{4} \\ 1-\frac{\nu}{4},\ 0,\ \frac{1}{2} \end{array} \right) - 1.$$

Note that $\cos(2\theta^*) = 2/(1+x^2) - 1 = (1-x^2)/(1+x^2)$, where $x = \tan(\theta^*)$. So the first cosine moment can be checked with this identity and the integral formula 3.389.2 (Gradshteyn and Ryzhik, 2007).

The first sine moment $\beta_p = E\sin(p\theta^*), p=1$ of $SCC2(0,\nu)$ with, w.l.o.g., $r=1$ is as follows:

$$\beta_1 = \frac{1}{\Gamma(\nu/2)2^{\frac{\nu}{2}}\sqrt{\pi}} G^{31}_{13}\left(\frac{1}{16} \middle| \begin{array}{c} \frac{1}{2}-\frac{\nu}{4} \\ \frac{1}{2}-\frac{\nu}{4},\ 0,\ \frac{1}{2} \end{array} \right).$$

3. Statistical Inference

We divide the problems into 2 cases. First assuming μ known and γ unknown, and secondly assuming μ unknown, γ known or both μ and γ unknown. For the first case, we can get very nice theoretical results. In Section 3.3, we examine how to choose an appropriate model among the l-axial exponential family of distributions.

3.1. Assuming μ known and γ unknown

To make the problem in hand, we change the densities (2.3) in the canonical form. If we change $d(\gamma)$ as η and write the densities (2.3) in the canonical form similar to the linear canonical form, then

$$f(\theta^*;\mu,\eta) = a^*\left(r\tan\frac{l(\theta^*-\mu)}{c_p}\right)\exp\left\{c\left(r\tan\frac{l(\theta^*-\mu)}{c_p}\right)\eta - A(\eta)\right\}, \quad l\in\mathbb{N},\ p=1,2, \quad (3.1)$$

where $a^*(r\tan(l(\theta^*-\mu)/c_p)) = r\sec^2(l(\theta^*-\mu)/c_p)a(r\tan(l(\theta^*-\mu)/c_p))$, and $A(\eta) = -\log(b(\gamma))$. When $p=1$, θ^* is in $(-\pi/l+\mu, \pi/l+\mu)$ and for $p=2$, θ^* is in $(\mu, 2\pi/l+\mu)$. The set Ξ of points η for which

$$\int a^*\left(r\tan\frac{l(\theta^*-\mu)}{c_p}\right)\exp\left\{c\left(r\tan\frac{l(\theta^*-\mu)}{c_p}\right)\eta\right\} d_\nu(\theta^*) < \infty$$

is called the natural parameter space of the family and η is called the natural parameter.

Note

$$\int a^* \left(r \tan \frac{l(\theta^* - \mu)}{c_p} \right) \exp \left\{ c \left(r \tan \frac{l(\theta^* - \mu)}{c_p} \right) \eta - A(\eta) \right\} d_\nu(\theta^*) = 1.$$

Differentiate above identity with respect to η to find

$$E_\eta c \left(r \tan \frac{l(\theta^* - \mu)}{c_p} \right) = \frac{d}{d\eta} A(\eta). \tag{3.2}$$

Differentiating above identity with respect to η again to find

$$\mathrm{Var}_\eta c \left(r \tan \frac{l(\theta^* - \mu)}{c_p} \right) = \frac{d^2}{d\eta^2} A(\eta).$$

The log-likelihood for a random sample of size n, $\theta^* = (\theta_1^*, \ldots, \theta_n^*)$, from (3.1) is

$$l(\eta) = \sum_{i=1}^n \log \left\{ a^* \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) \right\} + \eta \sum_{i=1}^n c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) - nA(\eta).$$

The likelihood equation is given by

$$\frac{1}{n} \sum_{i=1}^n c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) = \frac{d}{d\eta} A(\eta).$$

So by (3.2)

$$E_\eta c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) = \frac{d}{d\eta} A(\eta) = \frac{1}{n} \sum_{i=1}^n c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right). \tag{3.3}$$

The left most side of (3.3) is a strictly increasing function of η since

$$\frac{d}{d\eta} E_\eta c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) = \mathrm{Var}_\eta c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right) > 0. \tag{3.4}$$

It follows that equation (3.3) has at most one solution. The conditions of Theorem 3.10 (Lehmann and Casella, 1998) are satisfied. This theorem is related to establishing the existence of a consistent root of the likelihood equation. Furthermore, this theorem asserts that any such sequence is asymptotically normal and efficient.

With probability tending to 1, (3.3) therefore has a solution $\hat{\eta}$. This solution is the maximum likelihood estimator(mle) of η and is unique, consistent, and asymptotically efficient so that

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\mathcal{L}} N \left(0, \frac{1}{I(\eta)} \right), \tag{3.5}$$

where $I(\eta)$ is the Fisher information defined as

$$I(\eta) = E_\eta \left[\frac{d}{d\eta} \log f(\theta^*) \right]^2. \tag{3.6}$$

After direct calculation of (3.6) using (3.3) and (3.4), we obtain it as

$$I(\eta) = E_\eta \left[\frac{d}{d\eta} \log f(\theta^*) \right]^2 = \frac{d^2}{d\eta^2} A(\eta) = \mathrm{Var}_\eta c \left(r \tan \frac{l(\theta_i^* - \mu)}{c_p} \right). \tag{3.7}$$

Furthermore $g(\hat{\eta})$ for any function $g(\cdot)$ is the mle of $g(\eta)$ by the invariance property of mle. Asymptotic distribution of $g(\hat{\eta})$ is as follows:

$$\sqrt{n}(g(\hat{\eta}) - g(\eta)) \xrightarrow{\mathcal{L}} N\left(0, \frac{[g'(\eta)]^2}{I(\eta)}\right),$$

by the delta method provided $g'(\eta)$ exists and is not zero.

Based on this asymptotic distribution of $\hat{\eta}$, (3.5), we can do hypotheses test about η. The test statistic is

$$\mathrm{TS}_0 = \frac{\hat{\eta} - \eta_0}{\sqrt{\dfrac{\left(\dfrac{d^2}{d\eta^2} A(\eta)\Big|_{\eta=\eta_0}\right)^{-1}}{n}}}.$$

For the two sided test of $H_0 : \eta = \eta_0$ *vs.* $H_1 : \eta \neq \eta_0$, we reject H_0 if $|TS_0| > z_{\alpha/2}$ since, under H_0, the distribution of the test statistic is the standard normal distribution. Under $H_0 : \eta \geq \eta_0$, we reject H_0 if $TS_0 < -z_\alpha$. Similarly we reject $H_0 : \eta \leq \eta_0$ if $\mathrm{TS}_0 > z_\alpha$. Since for the one-sided tests, under H_0, the distribution of the test statistic is the standard normal distribution. Approximate $(1-\alpha)100\%$ two-sided confidence interval is immediate by the asymptotic distribution of $\hat{\eta}$ as follow:

$$\hat{\eta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\left(\dfrac{d^2}{d\eta^2} A(\eta)\Big|_{\eta=\hat{\eta}}\right)^{-1}}{n}}$$

since $\hat{\eta}$ is a consistent estimator of η and by Slutsky's theorem.

Furthermore $\sum_{i=1}^n c(r \tan(l(\theta_i^* - \mu)/c_p))$ is a sufficient statistic for η by the Factorization theorem. This family of distributions is a complete so it is a complete sufficient statistic for η. If we can find an unbiased estimator of η, then it is also UMVUE(Uniformly Minimum Variance Unbiased Estimator) for η by the Rao-Blackwell-Lehmann-Scheffé theorem (Lehmann and Casella, 1998). To validate all these theoretical results, an ad hoc approach first gets good estimates of a location parameter μ, for example, circular mean (Jammalamadaka and SenGupta, 2001). Then the data is now free of location so the above theoretical results can be applied which is fruitful. For the following examples, Cramer-Rao inequality is exact and every parameter is UMVUE of the corresponding parameters. Some examples are as follows:

Example 3.1. For Example 2.1 of Section 2.1, let $\eta = 1/\varphi^2$, then $A(\eta) = -\log(\eta)/2$. So mle of η is $n/\sum_{i=1}^n \tan^2(l(\theta_i^* - \mu)/2)$ and mle of φ^2 is $(1/n)\sum_{i=1}^n \tan^2(l(\theta_i^* - \mu)/2)$ by the invariance property of mle. This estimator is UMVUE of φ^2 and Cramer-Rao inequality is exact. An unbiased condition can be checked using the transformation $X_i = \tan(l(\theta_i^* - \mu)/2)$ and normal distribution moments. The asymptotic distribution of $\hat{\eta}$ is

$$\sqrt{n}\,(\hat{\eta} - \eta) \xrightarrow{\mathcal{L}} N\left(0, 2\eta^2\right).$$

Furthermore asymptotic distribution of $\hat{\varphi}^2$ is given by the delta method as follows:

$$\sqrt{n}\left(\hat{\varphi^2} - \varphi^2\right) \xrightarrow{\mathcal{L}} N\left(0, 2\varphi^4\right).$$

The test statistic based on asymptotic distribution is

$$\mathrm{TS}_0 = \frac{\hat{\varphi^2} - \varphi_0^2}{\sqrt{2}\varphi_0^2/\sqrt{n}}.$$

To test $H_0 : \varphi^2 = \varphi_0^2$ *vs.* $H_1 : \varphi^2 \neq \varphi_0^2$, we reject H_0 if $|TS_0| > z_{\alpha/2}$. Under $H_0 : \varphi^2 \geq \varphi_0^2$, we reject H_0 if $TS_0 < -z_\alpha$. Similarly, we reject $H_0 : \varphi^2 \leq \varphi_0^2$ if $TS_0 > z_\alpha$. Approximate $(1-\alpha)100\%$ two-sided confidence interval for φ^2 is given by

$$\hat{\varphi^2} \pm z_{\frac{\alpha}{2}} \frac{\sqrt{2}\hat{\varphi^2}}{\sqrt{n}}$$

by Slutsky's theorem.

An interesting point is that we can also derive the exact distribution of $\hat{\varphi}^2$. For the SCN density, let $X_i = \tan(l(\theta_i^* - \mu)/2)$ then the distribution of X_i/φ are independent and identically distributed as the standard normal distribution. So the distribution of mle of φ^2 is $(\varphi^2/n)\chi^2(n)$ by the property of χ^2 distribution. The test statistic based on the exact distribution is

$$\mathrm{TS}_0 = \frac{n\hat{\varphi^2}}{\varphi_0^2}.$$

The distribution of the test statistic is $\chi^2(n)$ under $H_0 : \varphi^2 = \varphi_0^2$ so we reject H_0 if $TS_0 > \chi^2_{n,\alpha/2}$ or $TS_0 < \chi^2_{n,1-\alpha/2}$. Under $H_0 : \varphi^2 \geq \varphi_0^2$, we reject H_0 if $TS_0 < \chi^2_{n,1-\alpha}$. Similarly we reject $H_0 : \varphi^2 \leq \varphi_0^2$ if $TS_0 > \chi^2_{n,\alpha}$. Since, under the given two one-sided H_0's, the test statistic's distribution is $\chi^2(n)$. Exact $(1-\alpha)100\%$ two-sided confidence interval for φ^2 is given by

$$\left(\frac{n\hat{\varphi^2}}{\chi^2_{n,\frac{\alpha}{2}}}, \frac{n\hat{\varphi^2}}{\chi^2_{n,1-\frac{\alpha}{2}}} \right).$$

Example 3.2. For Example 2.2 of Section 2.1, let $\eta = 1/\varphi$, then $A(\eta) = -\log(\eta)$. Therefore, the mle of η is $n/\sum_{i=1}^n |\tan(l(\theta_i^* - \mu)/2)|$ and mle of φ is $(1/n)\sum_{i=1}^n |\tan(l(\theta_i^* - \mu)/2)|$ by the invariance property of mle. Mle of φ is also UMVUE and Cramer-Rao inequality has exact bound. Asymptotic distribution of $\hat{\eta}$ is

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\mathcal{L}} N(0, \eta^2),$$

and the asymptotic distribution of $\hat{\varphi}$ is given by the delta method as follows:

$$\sqrt{n}(\hat{\varphi} - \varphi) \xrightarrow{\mathcal{L}} N(0, \varphi^2).$$

The test statistic is

$$\mathrm{TS}_0 = \frac{\hat{\varphi} - \varphi_0}{\varphi_0/\sqrt{n}}.$$

Under $H_0 : \varphi = \varphi_0$, we reject H_0 if $|TS_0| > z_{\alpha/2}$. Under $H_0 : \varphi \geq \varphi_0$, we reject H_0 if $TS_0 < -z_\alpha$. Similarly we reject $H_0 : \varphi \leq \varphi_0$ if $TS_0 > z_\alpha$. Approximate $(1-\alpha)100\%$ two-sided confidence interval for φ is given by

$$\hat{\varphi} \pm z_{\frac{\alpha}{2}} \frac{\hat{\varphi}}{\sqrt{n}},$$

by Slutsky's theorem.

Similar to SCN distribution, we can get the exact distribution of mle of φ. Let $X_i = |\tan(l(\theta_i^* - \mu)/2)|$, then the distribution of $\hat{\varphi}$ is $\Gamma(n, \varphi/n)$ by the property of gamma distribution. The test statistic based on the exact distribution is

$$\mathrm{TS}_0 = \frac{2n\hat{\varphi}}{\varphi_0}.$$

The distribution of the test statistic is $\chi^2(2n)$ under $H_0 : \varphi = \varphi_0$ by the property of gamma distribution. So we reject H_0 if $TS_0 > \chi^2_{2n,\alpha/2}$ or $TS_0 < \chi^2_{2n,1-\alpha/2}$. Under $H_0 : \varphi \geq \varphi_0$, we reject H_0 if $TS_0 < \chi^2_{2n,1-\alpha}$. Similarly, we reject $H_0 : \varphi \leq \varphi_0$ if $TS_0 > \chi^2_{2n,\alpha}$. Since, under the given two one-sided H_0's, the test statistic's distribution is $\chi^2(2n)$. Exact $(1-\alpha)100\%$ two-sided confidence interval for φ is given by

$$\left(\frac{2n\hat{\varphi}}{\chi^2_{2n,\frac{\alpha}{2}}}, \frac{2n\hat{\varphi}}{\chi^2_{2n,1-\frac{\alpha}{2}}} \right).$$

Example 3.3. For example 3.3 of Section 2.1, let $\eta = 1/\varphi$, then $A(\eta) = -\delta \log(\eta)$. Therefore, the mle of η is $n\delta / \sum_{i=1}^n \tan(l(\theta_i^* - \mu)/4)$ and mle of φ is $1/n\delta \sum_{i=1}^n \tan(l(\theta_i^* - \mu)/4)$ by the invariance property of mle. Mle of φ is also UMVUE and Cramer-Rao inequality has exact bound. Asymptotic distribution of $\hat{\eta}$ is

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\mathcal{L}} N\left(0, \frac{\eta^2}{\gamma}\right),$$

and the asymptotic distribution of $\hat{\varphi}$ is given by the delta method as follows:

$$\sqrt{n}(\hat{\varphi} - \varphi) \xrightarrow{\mathcal{L}} N\left(0, \frac{\varphi^2}{\gamma}\right).$$

The test statistic is

$$\mathrm{TS}_0 = \frac{\hat{\varphi} - \varphi_0}{\varphi_0/\sqrt{n\gamma}}.$$

Under $H_0 : \varphi = \varphi_0$, we reject H_0 if $|TS_0| > z_{\alpha/2}$. Under $H_0 : \varphi \geq \varphi_0$, we reject H_0 if $\mathrm{TS}_0 < -z_\alpha$. Similarly we reject $H_0 : \varphi \leq \varphi_0$ if $\mathrm{TS}_0 > z_\alpha$. Approximate $(1-\alpha)100\%$ two-sided confidence interval for φ is given by

$$\hat{\varphi} \pm z_{\frac{\alpha}{2}} \frac{\hat{\varphi}}{\sqrt{n\gamma}},$$

by Slutsky's theorem.

Let $X_i = \tan(l(\theta_i^* - \mu)/4)$, then the distribution of $\hat{\varphi}$ is $\Gamma(n\gamma, \varphi/(n\gamma))$ by the property of gamma distribution. We can do hypotheses test based on the exact distribution. The test statistic based on the exact distribution is now

$$\mathrm{TS}_0 = \frac{2n\gamma\hat{\varphi}}{\varphi_0}.$$

The distribution of the test statistic is $\chi^2(2n\gamma)$ under $H_0 : \varphi = \varphi_0$ by the property of gamma distribution. So we reject H_0 if $TS_0 > \chi^2_{2n\gamma,\alpha/2}$ or $TS_0 < \chi^2_{2n\gamma,1-\alpha/2}$. Under $H_0 : \varphi \geq \varphi_0$, we reject

H_0 if $TS_0 < \chi^2_{2n\gamma,1-\alpha}$. Similarly we reject $H_0 : \varphi \le \varphi_0$ if $TS_0 > \chi^2_{2n\gamma,\alpha}$. Since, under the given two one-sided H_0's, the test statistic's distribution is $\chi^2(2n\gamma)$. Exact $(1-\alpha)100\%$ two-sided confidence interval for φ is given by

$$\left(\frac{2n\gamma\hat{\varphi}}{\chi^2_{2n\gamma,\frac{\alpha}{2}}}, \frac{2n\gamma\hat{\varphi}}{\chi^2_{2n\gamma,1-\frac{\alpha}{2}}} \right).$$

Example 3.4. For Example 3.4 of Section 2.1, let $\eta = \nu$, then $A(\eta) = \log(\Gamma(\eta/2)2^{\eta/2})$. Therefore the mle of ν is the solution of the likelihood equation,

$$\frac{1}{2n}\sum_{i=1}^{n} \log\left(r\tan\left(\frac{l(\theta_i^* - \mu)}{4} \right) \right) = \frac{d}{d\nu}\log\Gamma\left(\frac{\nu}{2}\right) - \frac{\log(2)}{2}$$

after simple algebra. Unfortunately in this case we cannot get the mle of ν as in closed form, but still we can get it using some numerical methods. Asymptotic distribution of $\hat{\nu}$ is

$$\sqrt{n}(\hat{\nu} - \nu) \xrightarrow{\mathcal{L}} N\left(0, \frac{4}{\zeta}\left(2, \frac{\nu}{2}\right)\right),$$

where $\zeta(z,q)$ is Riemann's zeta function (Gradshteyn and Ryzhik, 2007) which has an integral representation

$$\zeta(z,q) = \frac{1}{\Gamma(z)}\int_0^\infty \frac{t^{z-1}e^{-qt}}{1-e^{-t}}\, dt.$$

Asymptotic variance of $\sqrt{n}(\hat{\nu} - \nu)$, $1/\mathrm{Var}_\nu\left\{(1/2)\log(r\tan(l(\theta_i^* - \mu)/4))\right\}$, is derived using a transformation, $X_i = r\tan(l(\theta_i^* - \mu)/4)$ which follows $\chi^2(\nu)$ distribution.
$E\log X_i$ is derived using the integral formula 4.352.1 (Gradshteyn and Ryzhik, 2007) as follows:

$$E\log X_i = \psi\left(\frac{\nu}{2}\right) + \log 2, \tag{3.8}$$

where $\psi(z)$ is Euler psi function defined by $\psi(z) = d/dz\log\Gamma(z)$. It has an integral representation as follows:

$$\psi(z) = \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-zt}}{1-e^{-t}} \right) dt.$$

$E(\log X_i)^2$ is derived using the integral formula 4.358.2 (Gradshteyn and Ryzhik, 2007) as follows:

$$E(\log X_i)^2 = \left[\psi\left(\frac{\nu}{2}\right) + \log 2\right]^2 + \zeta\left(2, \frac{\nu}{2}\right). \tag{3.9}$$

Hence asymptotic variance of $\sqrt{n}(\hat{\nu} - \nu)$ is immediate using (3.8) and (3.9). We may also derive $I(\eta)$ using (3.7) as follows:

$$I(\nu) = A''(\nu) = \frac{\Gamma''\left(\frac{\nu}{2}\right)\Gamma\left(\frac{\nu}{2}\right) - \left(\Gamma'\left(\frac{\nu}{2}\right)\right)^2}{4\left(\Gamma\left(\frac{\nu}{2}\right)\right)^2}.$$

We use the asymptotic variance since this closed form is little bit more complicated than the asymptotic variance.

The test statistic is

$$\mathrm{TS}_0 = \frac{\hat{\nu} - \nu_0}{\sqrt{4/\{\zeta\,(2, \nu_0/2) \cdot n\}}}.$$

Under $H_0 : \nu = \nu_0$, we reject H_0 if $|TS_0| > z_{\alpha/2}$. Under $H_0 : \nu \geq \nu_0$, we reject H_0 if $TS_0 < -z_\alpha$. Similarly we reject $H_0 : \nu \leq \nu_0$ if $TS_0 > z_\alpha$. Approximate $(1-\alpha)100\%$ two-sided confidence interval for ν is given by

$$\hat{\nu} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{4}{\{\zeta\,(2, \hat{\nu}/2) \cdot n\}}},$$

by Slutsky's theorem.

3.2. Assuming μ unknown, γ known or both μ and γ unknown

In these situations, nice theoretical results are not easy. However, we can still get the mles of unknown parameters using any numerical routines. The log-likelihood for a random sample of size n, $\theta^* = (\theta_1^*, \ldots, \theta_n^*)$, from (2.3) is

$$l(\mu, \gamma) = \sum_{i=1}^n \log\left\{a^*\left(r \tan\frac{l(\theta_i^* - \mu)}{c_p}\right)\right\} + n \log b(\gamma) + d(\gamma) \sum_{i=1}^n c\left(r \tan\frac{l(\theta_i^* - \mu)}{c_p}\right).$$

The corresponding estimates can be computed by direct minimization (Byrd *et al.*, 1995) of the minus log-likelihood itself. Byrd's method allows box constraints, that is, each variable can be given a lower and/or upper bound. For l-axial exponential family, to improve estimation process we can use ranges of μ and γ of the likelihood as box constraints. These constraints depend on a specific distribution.

Another possible approach is the profile log-likelihood method. We have a random sample of size n, $\theta^* = (\theta_1^*, \ldots, \theta_n^*)$, from (3.1). For fixed μ, the mle of η is a solution to the likelihood equation of canonical form, *i.e.*

$$\frac{1}{n} \sum_{i=1}^n c\left(r \tan\frac{l(\theta_i^* - \mu)}{c_p}\right) = \frac{d}{d\eta} A(\eta). \tag{3.10}$$

Denote it as $\hat{\eta}$. Thus the profile log-likelihood is

$$l(\mu) = l(\mu, \hat{\eta}) = \sum_{i=1}^n \log\left\{a^*\left(r \tan\frac{l(\theta_i^* - \mu)}{c_p}\right)\right\} + \hat{\eta} \sum_{i=1}^n c\left(r \tan\frac{l(\theta_i^* - \mu)}{c_p}\right) - nA(\hat{\eta}). \tag{3.11}$$

The solution of the profile log-likelihood equation for μ is hard to get in closed form, but it is relatively easy to use a numerical maximization subroutine to obtain it. For an initial value of μ, we can use any measure of center, for example, circular mean (Jammalamadaka and SenGupta, 2001) defined by

$$\overline{\mu}_0 = \begin{cases} \tan^{-1} \dfrac{S}{C}, & \text{if } C > 0, S \geq 0, \\ \dfrac{\pi}{2}, & \text{if } C = 0, S > 0, \\ \tan^{-1} \dfrac{S}{C} + \pi, & \text{if } C < 0, \\ \tan^{-1} \dfrac{S}{C} + 2\pi, & \text{if } C \geq 0, S < 0, \\ \text{undefined}, & \text{if } C = 0, S = 0, \end{cases}$$

where $(C,S) = (\sum_{i=1}^n \cos(\theta_i^*), \sum_{i=1}^n \sin(\theta_i^*))$. We iterate the above profile approach until convergence, *i.e.*,

An algorithm.

Step 1: Use $\overline{\mu}_0$ as an initial value of μ,

Step 2: $\hat{\eta}$ is a solution to the likelihood equation of canonical form (3.10),

Step 3: $\hat{\mu}$ is the maximizer of the profile log-likelihood (3.11),

Step 4: Check convergence. If not, go to step 2.

3.3. Model checking

In this section, we concern about how to choose an appropriate model among l-axial exponential family of distributions. The first approach is a graphical one and the second method is a theoretical one.

One rough graphical approach is to use a circular data plot with a *pdf* plot of l-axial exponential family of distributions. A little bit enhanced method uses Healy's plot (Healy, 1968). Healy's plot is based on

$$d_i = r \tan \frac{l(\theta_i^* - \mu)}{c_p}, \quad (i = 1, \ldots, n). \tag{3.12}$$

In addition, d_i is sampled from exponential family of distributions if the fitted model is appropriate. Practically the exact parameter values in equation (3.12) need to be replaced by estimates. Above d_i then sorted and plotted against exponential family of distributions percentage points. Similarly, the cumulative probabilities of an exponential family of distributions can be plotted against their nominal values $1/n, 2/n, \ldots, 1$; the points should lie on the bisection line of the quadrant.

The above approaches are graphical methods, whereas the following one is a theoretical one. Suppose θ^* follows a member of the l-axial exponential family of distributions and we wish to test the hypothesis,

$$H_0 : F_{\theta^*}(\theta^*) = F_0(\theta^*) \ \forall\, x \quad vs. \quad H_1 : \exists\, x \text{ such that } F_{\theta^*}(\theta^*) \neq F_0(\theta^*),$$

where $F_0(\theta^*)$ is given by (2.4). Then the Kolmogorov-Smirnov test (Lehmann and Romano, 2005)) can be adopted. Given a random sample of size n, $\theta^* = (\theta_1^*, \ldots, \theta_n^*)$, from the l-axial exponential family of distributions, we first arrange those in increasing order of magnitude. The empirical distribution function is defined by

$$\hat{F}_n(\theta^*) = \begin{cases} 0, & \text{if } \ \theta^* < \theta_{(1)}^*, \\ \dfrac{i}{n}, & \text{if } \ \theta_{(i)}^* \leq \theta^* < \theta_{(i+1)}^*, \\ 1, & \text{if } \ \theta_{(n)}^* \leq \theta^*. \end{cases}$$

The value of Kolmogorov-Smirnov statistic is defined by

$$D_n = \sup_{\theta^*} \sqrt{n} \left| \hat{F}_n(\theta^*) - F_0(\theta^*) \right|.$$

The Kolmogorov-Smirnov test rejects the null hypothesis if $D_n > s_{n,1-\alpha}$, where $s_{n,1-\alpha}$ is the $1-\alpha$ quantile of the null distribution of D_n when F_0 is the uniform $U(0,1)$ distribution (Smirnov, 1948).

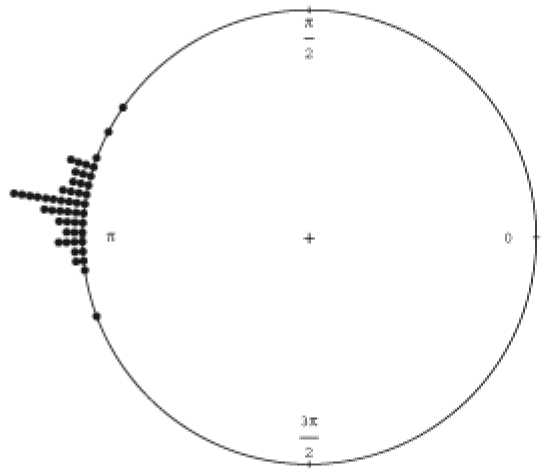

Figure 4.1. Circular data plot of samples of orientations of termite mounds

The finite sampling distribution of D_n under F_0 is the same for all continuous F_0, but its exact form is difficult to express. By the duality of tests and confidence regions, the Kolmogorov-Smirnov test can be inverted to yield uniform confidence bands for F, given by

$$R_{n,1-\alpha} = \left\{ F : \sup_{\theta^*} \sqrt{n} \left| \hat{F}_n(\theta^*) - F(\theta^*) \right| \leq s_{n,1-\alpha} \right\}.$$

By construction, $P_F\{F \in R_{n,1-\alpha}\} = 1 - \alpha$ if F is continuous. Therefore, the confidence band is then

$$\max\left\{0, \hat{F}_n(\theta^*) - s_{n,1-\alpha}\right\} \leq F(\theta^*) \leq \min\left\{1, \hat{F}_n(\theta^*) + s_{n,1-\alpha}\right\}.$$

4. Real Data Analysis

Figure 4.1 shows samples of orientations of termite mounds of Amitermes laurensis and their mean orientations, at 10^{th} site in Cape York Peninsula, North Queensland (Fisher, 1993). It is of interest to determine whether the mean orientations are consistent. However we consider only fitting our suggested model to this data. To fit the data using well-known von Mises model, we need to convert them to vectors first and then fit them using von Mises model since the data are axial. Because of this cumbersome data transformation, we suggested semicircular models to fit any axial data directly. Among them, we use the semicircular normal distribution to fit the data. For this data set, the corresponding maximum likelihood estimates are given by $\hat{\mu} = 3.01$, and $\hat{\varphi} = 0.167$. Histogram with the estimated *pdf* and Healy's plot (Healy, 1968) are shown in Figure 4.2. A visual inspection of Figure 4.2 indicates a satisfactory fit of the density to the data.

5. Conclusion

We derived the 4-axial (or semicircular) exponential family of distributions via the projection of the exponential family of distributions over a quarter-circular (or semicircular) segment. Then we extended it to the l-axial exponential family of distributions using the simple transformations for

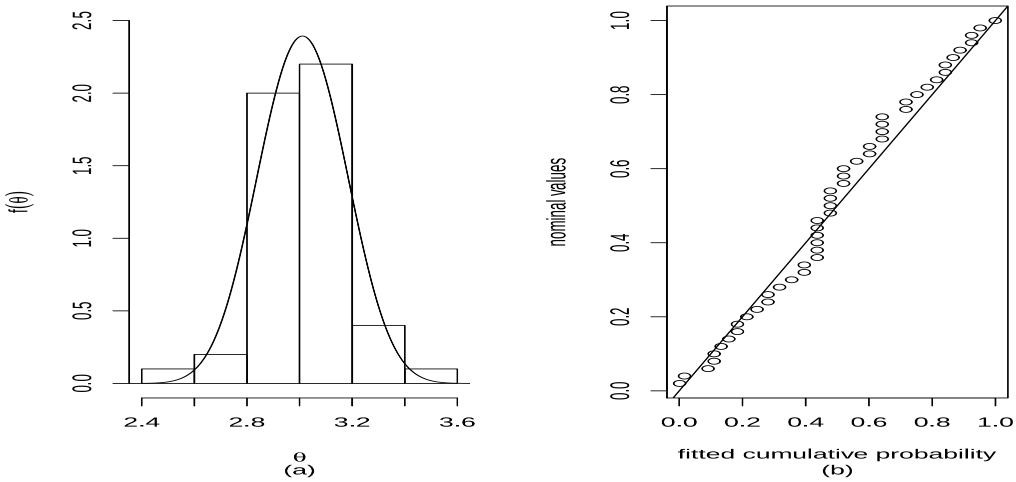

Figure 4.2. (a) Histogram with the estimated *pdf* and (b) Healy's plot

modeling any arc of arbitrary length say $2\pi/l$ for $l = 1, 2, \ldots$. Occasionally, measurements result in any arc of arbitrary length. A derived new family of the distributions can be used to model symmetric or skewed angular data. Asymptotic results reveal that linear exponential family of distributions can be used to approximate the l-axial exponential family of distributions. Some basic properties of the l-axial exponential family of distributions are introduced.

Trigonometric moments are derived for members of l-axial exponential family of distributions. We find that, by a simple transformation, the k^{th} sine(cosine) moment of the semicircular exponential family of distributions is the same as the $2k^{th}$ sine(cosine) moment of the 4-axial exponential family of distributions because of the transformation we are using. Furthermore the k^{th} sine(cosine) moment of the circular exponential family of distributions is the same as the $2k^{th}$ sine(cosine) moment of the semicircular exponential family of distributions.

When a location parameter is known, we suggested how to develop a general maximum likelihood estimator and how to do hypothesis tests. Confidence intervals are also developed by the duality of tests and confidence intervals. To validate all these theoretical results, an ad hoc approach is suggested using a good estimates of a location parameter μ. When both parameters are unknown, we derived the mles of unknown parameters using by direct minimization of the minus log-likelihood itself. A profile likelihood approach is also introduced. For model checking, a graphical approach and a theoretical approach are adopted. To demonstrate our proposed model we applied semicircular normal distribution to the samples of orientations of termite mounds of Amitermes laurensis, and their mean orientations, at 10^{th} site in Cape York Peninsula, North Queensland (Fisher, 1993). A visual inspection of Healy's plot indicates a satisfactory fit of the density to the data.

As a future study, we mention some general comments. We only considered the exponential family of distributions, but eventually all continuous distributions can be transformed to l-axial distributions. For example, let $X \sim U(0,1)$, then take the same projection method so far, i.e., $X = r\tan(\theta)$. The *pdf* becomes

$$f(\theta) = r\sec^2(\theta), \quad 0 < \theta < \tan^{-1}\left(\frac{1}{r}\right).$$

As mentioned earlier, we assume without loss of generality $r = 1$, then the support of θ becomes $(0, \pi/4)$. Similar to previous approach, let $\theta^* = 8\theta/l$, then the *pdf* of θ^* becomes

$$f(\theta^*) = \frac{l}{8} \sec^2\left(\frac{l\theta^*}{8}\right), \quad 0 < \theta^* < \frac{2\pi}{l},$$

which is the *pdf* of l-axial uniform distribution. Furthermore, this type of linear bounded support also happen at the exponential family. For example the beta distribution has the same support as the uniform distribution on $(0, 1)$. We can develop l-axial beta distribution even though support is not $\Re$ or $\Re^+$. Let $X = r\tan(\theta)$, then the *pdf* becomes

$$f(\theta; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (r\tan(\theta))^{\alpha-1}(1 - r\tan(\theta))^{\beta-1} r\sec^2(\theta), \quad 0 < \theta < \tan^{-1}\left(\frac{1}{r}\right).$$

As mentioned earlier, we can assume without loss of generality $r = 1$, then the support of θ becomes $(0, \pi/4)$. Similar to previous approach, let $\theta^* = 8\theta/l$, then the *pdf* of θ^* becomes

$$f(\theta^*; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\tan\frac{l\theta^*}{8}\right)^{\alpha-1} \left(1 - \tan\frac{l\theta^*}{8}\right)^{\beta-1} \sec^2\left(\frac{l\theta^*}{8}\right), \quad 0 < \theta^* < \frac{2\pi}{l},$$

which is the *pdf* of l-axial beta distribution.

References

Ahn, B. J. and Kim, H. M. (2008). A new family of semicircular models: The semicircular Laplace distributions, *Communications of the Korean Statistical Society*, **15**, 775–781.

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

Casella, G. and Berger, R. L. (2002). *Statistical Inference*, 2nd Edition, Duxbury Press.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*, Cambridge University Press.

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products(7^{th} edition)*, Academic Press.

Guardiola, J. H. (2004). *The Semicircular Normal Distribution, Ph. D. Dissertation*, Baylor University, Institute of Statistics.

Healy, M. J. R. (1968). Multivariate normal plotting, *Applied Statistics*, **17**, 157–161.

Jammalamadaka, S. R. and Kozubowski, T. J. (2003). A New family of circular models: The wrapped Laplace distributions, *Advances and Applications in Statistics*, **3**, 77–103.

Jammalamadaka, S. R. and Kozubowski, T. J. (2004). New families of wrapped distributions for modeling skew circular data, *Communications in Statistics-Theory and Method*, **33**, 2059–2074.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*, World Scientific Publishing, Singapore.

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

Jones, T. A. (1968). Statistical analysis of orientation data, *Journal of Sedimentary Petrology*, **38**, 61–67.

Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, 2^{nd}edition, Springer Science, New York.

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*, 3^{rd}edition, Springer Science, New York.

Mardia, K. V. and Jupp, P. (2000). *Directional Statistics*, John Wiley and Sons, Chichester.

Pewsey, A. (2000). The wrapped skew-normal distribution on the circle, *Communications in Statistics-Theory and Method*, **29**, 2459–2472.

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

Pewsey, A. (2004). Testing for circular reflective symmetry about a known median axis, *Journal of Applied Statistics*, **31**, 575–585

Pewsey, A. (2006). Modelling asymmetrically distributed circular data using the wrapped skew-normal distribution, *Environmental and Ecological Statistics*, **13**, 257–269.

Pewsey, A. (2008). The wrapped stable family of distributions as a flexible model for circular data, *Computational Statistics & Data Analysis*, **52**, 1516–1523.

Pewsey, A., Lewis, T. and Jones, M. C. (2007). The wrapped t family of circular distributions, *Australian & New Zealand Journal of Statistics*, **49**, 79–91.

Smirnov, N. V. (1948). Tables for estimating the goodness of fit of empirical distributions, *Annals of Mathematical Statistics*, **19**, 279–281.