Far East J. Theo. Stat. **20**(2) (2006), 207-216

THE SEMI-CIRCULAR NORMAL DISTRIBUTION

JOSÉ H. GUARDIOLA[*], JAMES D. STAMEY[**], JOHN W. SEAMAN[**] and HASSAN ELSALLOUKH[***]

[*]Department of Computing and Mathematical Sciences
Texas A&M University at Corpus Christi
Corpus Christi Texas, 78412, U. S. A.
e-mail: jose.guardiola@tamucc.edu

[**]Department of Statistical Science
Baylor University, Waco TX 76798, U. S. A.
e-mail: james_stamey@baylor.edu
john_seaman@baylor.edu

[***]Department of Mathematics and Statistics
University of Arkansas at Little Rock
Little Rock, Arkansas, 72204, U. S. A.
e-mail: hxelsalloukh@ualr.edu

Abstract

We introduce the semicircular normal (SCN) family of distributions for observations on the half-circle and for axial data. Suitably transformed, it can also be used for fully circular data. For small angles, the SCN and the normal distribution are equivalent and there is a limiting relationship between the SCN and the von Mises distribution. We also develop a bivariate version of the SCN. Inference for the SCN is straightforward, using method of moments, maximum likelihood, or Bayesian estimators. We illustrate the applicability of the SCN with some examples.

2000 Mathematics Subject Classification: 62E15.

Keywords and phrases: axial data, semicircular distribution, von Mises distribution for half-circle, angular distribution.

Received June 24, 2006

1. Introduction

Suppose a herpetologist observes sea turtles emerging from the surf to locate nesting sites, where the directions taken by the turtles are confined to a 180° arc. Similarly, an ornithologist observing the departure of sea birds form the shear face of a cliff need only record angles from within $[0, \pi]$ radians. A full-circular distribution, such as the von Mises, see for example, Jammalamadaka and SenGupta [3], is not needed in situations such as these. We refer to observations in these cases as *semicircular data*. In this paper, we introduce a new distribution for semicircular data. The new distribution is constructed by projecting the support of a normal density onto a semicircle. We shall call this the *semicircular normal* (SCN) family of distributions. This new family of distributions is simpler than other families for angular data, but no less flexible in application. Indeed, the SCN family can be extended to bimodal forms and to the full circle. Furthermore, it is particularly adept as a model for axial data.

2. The Semicircular Normal Distribution

The *semicircular normal* (SCN) *distribution* is obtained by projecting a normal distribution over a semicircular segment. In this section, we derive the density and consider its properties.

Let x have a normal distribution with mean zero and variance σ^2. For a positive real number r, define the angle θ by

$$\theta = g(x) = \arctan\left(\frac{x}{r}\right).$$

Hence $x = g^{-1}(\theta) = r\tan(\theta)$. It follows that $\frac{dx}{d\theta} r\sec^2(\theta)$ and the density of θ is

$$\begin{aligned} f(\theta) &= f[g^{-1}(\theta)]\frac{dx}{d\theta} \\ &= \frac{r}{\sqrt{2\pi}\sigma}\sec^2(\theta)\exp\left[-\frac{r^2\tan^2(\theta)}{2\sigma^2}\right], \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}, \quad \sigma, r \in \mathbb{R}^+. \end{aligned}$$

Let $\varphi = \sigma/r$. Then, allowing for negative angles, we have

$$f(\theta) = \frac{1}{\sqrt{2\pi}\varphi} \sec^2(\theta) \exp\left[-\frac{\tan^2(\theta)}{2\varphi^2}\right], \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}, \quad \varphi \in \mathbb{R}^+. \qquad (1)$$

We refer to (1) as the *semicircular normal* (SCN) density and write $\theta \sim \text{SCN}(0, \varphi)$. Notice that (1) is not defined at $\theta = \pm\pi/2$ as $\cos(\pi/2) = 0$ and therefore $\sec^2(\pi/2)$ is not defined. However, $\lim_{\theta \to \pi/2} f(\theta) = 0$. More generally, we introduce the parameter μ as the location parameter for the SCN relative to the horizontal axis and write the density as

$$f(\theta | \mu) = \frac{1}{\sqrt{2\pi}\varphi} [\sec(\theta - \mu)]^2 \exp\left\{-\frac{[\tan(\theta - \mu)]^2}{2\varphi^2}\right\}, -\frac{\pi}{2} + \mu < \theta < \frac{\pi}{2} + \mu, \quad \varphi \in \mathbb{R}^+.$$

It is straightforward to show that the cumulative distribution for θ is

$$F(\alpha) = \Phi\left[\frac{1}{\varphi} \tan(\alpha - \mu)\right],$$

where Φ is the standard normal CDF. Equivalently, $F(\alpha) = (1/2)erf \times [\tan(\alpha - \mu)/\sqrt{2}\varphi]$, where *erf* is the error function defined by $erf(x) = (2/\sqrt{\pi})\int_0^x e^{-u^2} du$. The function $erf(x)$ differs from the CDF of a normal distribution only by a multiplicative constant. See for example, Spiegel [5, p. 183].

3. Inference for the SCN

We now consider estimation of the parameters of the SCN. Our main interest is Bayesian methodology, but we first derive maximum likelihood estimators for μ and φ since the MLE. Suppose that we have a random sample, $\theta_1, \theta_2, \ldots, \theta_n$, from a SCN with pdf (1). Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$. Then the log-likelihood is

$$\ln l(\mu, \varphi | \boldsymbol{\theta}) = -n \ln(\varphi) - 2\sum_{i=1}^{n} \ln[\cos(\theta_i - \mu)] - \frac{1}{2\varphi^2} \sum_{i=1}^{n} \tan^2(\theta_i - \mu). \qquad (2)$$

We find the MLE for μ assuming φ is known. From (2), the log-likelihood equation for μ is

$$-2\sum_{i=1}^{n}\tan(\theta_i - \mu) + \frac{1}{\varphi^2}\sum_{i=1}^{n}\tan(\theta_i - \mu)\sec^2(\theta_i - \mu) = 0. \tag{3}$$

There is no closed form solution for μ in (3) so we must solve numerically to obtain the maximum likelihood estimator. However, in the special case of small $\theta_i - \mu$, simplifying assumptions can be made using small angle approximations. In that case

$$\tan(\theta_i - \mu) \approx \theta_i - \mu \tag{4}$$

and

$$\sec^2(\theta_i - \mu) \approx 1. \tag{5}$$

Substituting (4) and (5) into (3), we have

$$-2\sum_{i=1}^{n}(\theta_i - \mu) + \frac{1}{\varphi^2}\sum_{i=1}^{n}(\theta_i - \mu)(1)^2 = 0,$$

or equivalently, we have for small $\theta_i - \mu$, $\left(\frac{1}{\varphi^2} - 2\right)\sum_{i=1}^{n}(\theta_i - \mu) = 0$, for $\varphi^2 \neq 1/2$. This yields the small angle maximum likelihood estimator

$$\hat{\mu} \approx \frac{1}{n}\sum_{i=1}^{n}\theta_i.$$

Now suppose μ is known and φ is unknown. We have

$$\frac{\partial \ln[l(\mu, \varphi \mid \boldsymbol{\theta})]}{\partial\varphi} = -\frac{n}{\varphi} + \frac{1}{\varphi^3}\sum_{i=1}^{n}\tan^2(\theta_i - \mu).$$

Setting this equation to 0 in order to find the maximum we obtain

$$-\frac{n}{\varphi} + \frac{1}{\varphi^3}\sum_{i=1}^{n}\tan^2(\theta_i - \mu) = 0.$$

Obtaining the positive solution for φ we have the estimator

$$\hat{\varphi} = \left[\frac{1}{n}\sum_{i=1}^{n}\tan^2(\theta_i - \mu)\right]^{\frac{1}{2}}, \tag{6}$$

assuming μ is known.

When both parameters of the SCN distribution, φ and μ are unknown, the equations for the MLE have to be solved iteratively. Although it does not guarantee a global maximum in general, alternating between equations (3) and (6) in order to maximize the likelihood equation has provided reasonable results in our experience. A good initial value for μ can be obtained using $\hat{\mu} \approx \frac{1}{n}\sum_{i=1}^{n}\theta_i$, even when the angles are not small, and convergence is typically fast.

The large sample variance of the maximum likelihood estimators can be approximated by inverting the negative of the Hessian matrix. The required second derivatives are

$$\begin{aligned}
\frac{\partial^2 \ln[l(\mu, \varphi \mid \boldsymbol{\theta})]}{\partial \mu^2} &= 2\sum_{i=1}^{n}\sec^2(\theta_i - \mu) \\
&\quad - \frac{1}{2\varphi^2}\sum_{i=1}^{n}[2\sec^4(\theta_i - \mu) + 4\sec^2(\theta_i - \mu)\tan^2(\theta_i - \mu)], \\
\frac{\partial^2 \ln[l(\mu, \varphi \mid \boldsymbol{\theta})]}{\partial \varphi^2} &= \frac{n}{2\varphi^4}, \\
\frac{\partial^2 \ln[l(\mu, \varphi \mid \boldsymbol{\theta})]}{\partial \varphi \partial \mu} &= \sum_{i=1}^{n}\sec^2(\theta_i - \mu)\tan(\theta_i - \mu).
\end{aligned}$$

Once the variance is approximated, large sample confidence intervals can be constructed.

We next consider Bayesian inferential methods for the SCN distribution. We begin with some simple examples using single observations and one parameter known. Ultimately we provide a Bayesian approach to inference or the full model. Jammalamadaka and SenGupta [3, p. 278] have noted that "Attempts at Bayesian inference for circular data have not been as successful on the analytical front as they have been in the linear case, partly for lack of nice conjugate priors in the general case". As we shall see, the use of Markov chain Monte Carlo methods alleviates this difficulty.

For our Bayesian analysis it is convenient to reparameterize the likelihood. In place of φ we substitute $\kappa = 1/\varphi^2$. The SCN density is then

$$f(\theta \mid \mu, \kappa) = \frac{\kappa^{1/2}}{\sqrt{2\pi}}[\sec(\theta - \mu)]^2 \exp\left\{-\frac{\kappa}{2}[\tan(\theta - \mu)]^2\right\},$$

$$-\frac{\pi}{2} + \mu < \theta < \frac{\pi}{2} + \mu,\ \varphi \in \mathbb{R}^+.$$

We write $\theta \sim \mathrm{SCN}(\mu, \kappa)$. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ be n iid observations from a $\mathrm{SCN}(\mu, \kappa)$ distribution.

It is often desirable to use a prior structure that it is dominated by the likelihood in the posterior. Such "non-informative" priors are convenient when little or no prior information is available. For example, we may use $p_1(\mu) = p_2(\kappa) = 1/2\pi$, assuming μ and κ are independent. Then we have the posterior

$$p(\mu, \kappa \mid \boldsymbol{\theta}) \propto l(\mu, \kappa \mid \boldsymbol{\theta}) p_1(\mu) p_2(\kappa)$$

$$= \frac{1}{(2\pi)^2}\kappa^{n/2}\prod_{i=1}^{n}\sec^2(\theta_i - \mu)\exp\left[-\frac{\kappa}{2}\sum_{i=1}^{n}\tan^2(\theta_i - \mu)\right], \qquad (7)$$

which is of course proportional to the likelihood. We can maximize either (7) or the likelihood to obtain the posterior mode using the same approach as finding maximum likelihood estimators in the previous section as a point estimator for μ and κ.

An informative prior structure assuming prior independence of μ and κ can be constructed as follows. For μ consider the $\mathrm{SCN}(\delta, \nu)$ with density

$$p_1(\mu) = \frac{\nu^{1/2}}{\sqrt{2\pi}}[\sec(\mu - \delta)]^2 \exp\left\{-\frac{\nu}{2}[\tan(\mu - \delta)]^2\right\}.$$

This choice affords a fairly flexible family of priors for μ. For κ we use a gamma with parameters α, β:

$$p_2(\kappa) = \frac{\beta^\alpha}{\Gamma(\alpha)}\kappa^{\alpha-1}e^{-\beta\kappa}.$$

We choose the gamma prior because it has the advantage of producing a gamma posterior for κ as well. More details about gamma prior elicitation is discussed by Martz and Waller [4, pp. 318-324]. Thus the posterior is

$$p(\mu, \varphi \,|\, \boldsymbol{\theta}) \propto \kappa^{n/2}\prod_{i=1}^{n}\sec^2(\theta_i - \mu)\exp\left[-\frac{\kappa}{2}\sum_{i=1}^{n}\tan^2(\theta_i - \mu)\right]$$

$$\times \frac{\nu^{1/2}}{\sqrt{2\pi}}\sec^2(\mu - \delta)\exp\left[-\frac{\nu}{2}\tan^2(\mu - \delta)\right]\frac{\beta^\alpha}{\Gamma(\alpha)}\kappa^{\alpha-1}e^{-\beta\kappa}$$

$$= \left[\prod_{i=1}^{n}\sec^2(\theta_i - \mu)\right]\sec^2(\mu - \delta)\exp\left[-\frac{\nu}{2}\tan^2(\mu - \delta)\right]$$

$$\times \kappa^{n/2+\alpha-1}\exp\left\{-\kappa\left[\beta + \frac{1}{2}\sum_{i=1}^{n}\tan^2(\theta_i - \mu)\right]\right\}.$$

Again, posterior modes are straightforward to find.

To obtain the full posterior we employ MCMC methods using the software package Winbugs 1.4 (Imperial College and Medical Research Council, "Winbugs 1.4", 1996-2003). We shall illustrate the use of these methods in the examples.

4. Example

For this example we consider an axial statistical analysis using the von Mises distribution similar to that in Adams [1]. His analysis supports the conclusion that spindles do not rotate randomly, rather, they spend the most of their time aligned parallel or antiparallel to the direction in which they will later enter anaphase and undergo cell division. For more details see Adams [1]. Here we apply our methods to data generated to resemble that described by Adams. We compare the results obtained by using the SCN with those using the von Mises distribution.

Twenty four observations were generated with S-Plus and the CircStats package provided by Jammalamadaka and Sen Gupta [3]. The S-Plus code used to generate the data is available from the first author. For the Bayesian analysis, non-informative priors for μ and κ were used. For the MCMC algorithm, inferences were based on five chains each with 5,000 burn in iterations followed by 50,000 sample iterations. The Gelman-Rubin statistic and trace plots were also used to assess convergence. No problems with convergence were apparent. Posterior summaries of this Bayesian analysis using the SCN are provided in Table 1. The posterior modes of μ and κ are found iteratively using the closed form derivatives of the full posterior.

A maximum likelihood analysis of the SCN model parameters was performed as well. The results are summarized in Table 2. Note the close agreement between the MLEs and the point estimates from the Bayesian analysis.

Table 1. Posterior summaries using a Gibbs sampler for the SCN model

Quantity	Mean	Standard Deviation	Lower limit CI 2.5%	Upper Limit CI 97.5%
μ	179.3°	3.33°	172.8°	185.9°
κ	10.54	3.0	5.49	17.19
θ_{n+1}	179.3°	17.2°	146.3°	212.4°

Table 2. MLE summaries for the SCN model. The corresponding confidence intervals are indicated below

Quantity	Mode	Standard Error	2.5%	97.5%
μ (degrees)	179.4°	3.32°	172.7°	185.9°
κ	10.2	3.0	5.5	17.2

5. Conclusion

The SCN distribution is a good alternative to the von Mises distribution as it is able to accommodate axial data without the need of any transformation, i.e., there is no need to double the angles module 360. Furthermore, computations involving the SCN do not require use of Bessel functions. In our experience, inference based on the SCN differs little from that based on the von Mises when the data are highly concentrated around small angles. However, as variability in the data increases, or the data do not concentrate around small angles, the two models can differ markedly.

In this paper the semicircular normal (SCN) family of distributions is introduced. Ideally suited for observations on the half-circle and for axial data, it can also be used to analyze fully circular data. A bivariate version is also derived. Moment, maximum likelihood and Bayesian estimators are derived and all can be evaluated using standard software. Potential areas of future work are goodness of fit tests, allowing for covariates, and multivariate versions beyond the bivariate model derived in this paper.

Acknowledgement

The authors would like to thank Dr. R. Adams for kindly providing part of the data used in his 1996 Journal of Neuroscience paper.

References

[1] R. Adams, Metaphase spindles rotate in the Neuroepithelium of rat cerebral cortex, J. Neuroscience 16 (1996), 7610-7618.

[2] N. I. Fisher, Statistical Analysis of Circular Data, Cambridge University Press, Cambridge, 1993.

[3] S. R. Jammalamadaka and A. SenGupta, Topics in Circular Statistics, World Scientific Publishing Co., Singapore, 2001.

[4] H. F. Martz and R. A. Waller, Bayesian Reliability Analysis, John Wiley & Sons, New York, 1982.

[5] R. M. Spiegel, Mathematical Handbook of Formulas and Tables, McGraw-Hill, New York, 1968.