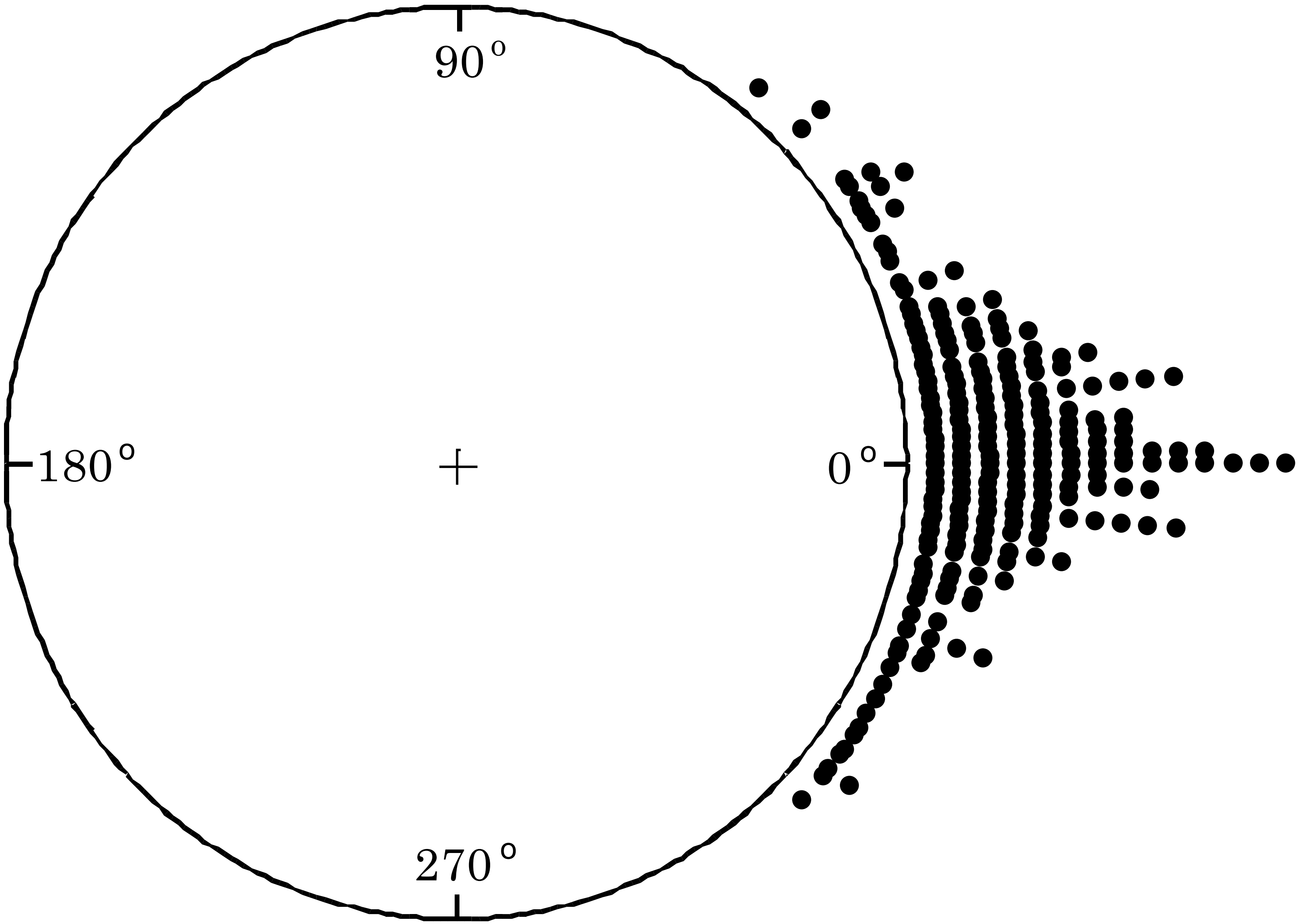
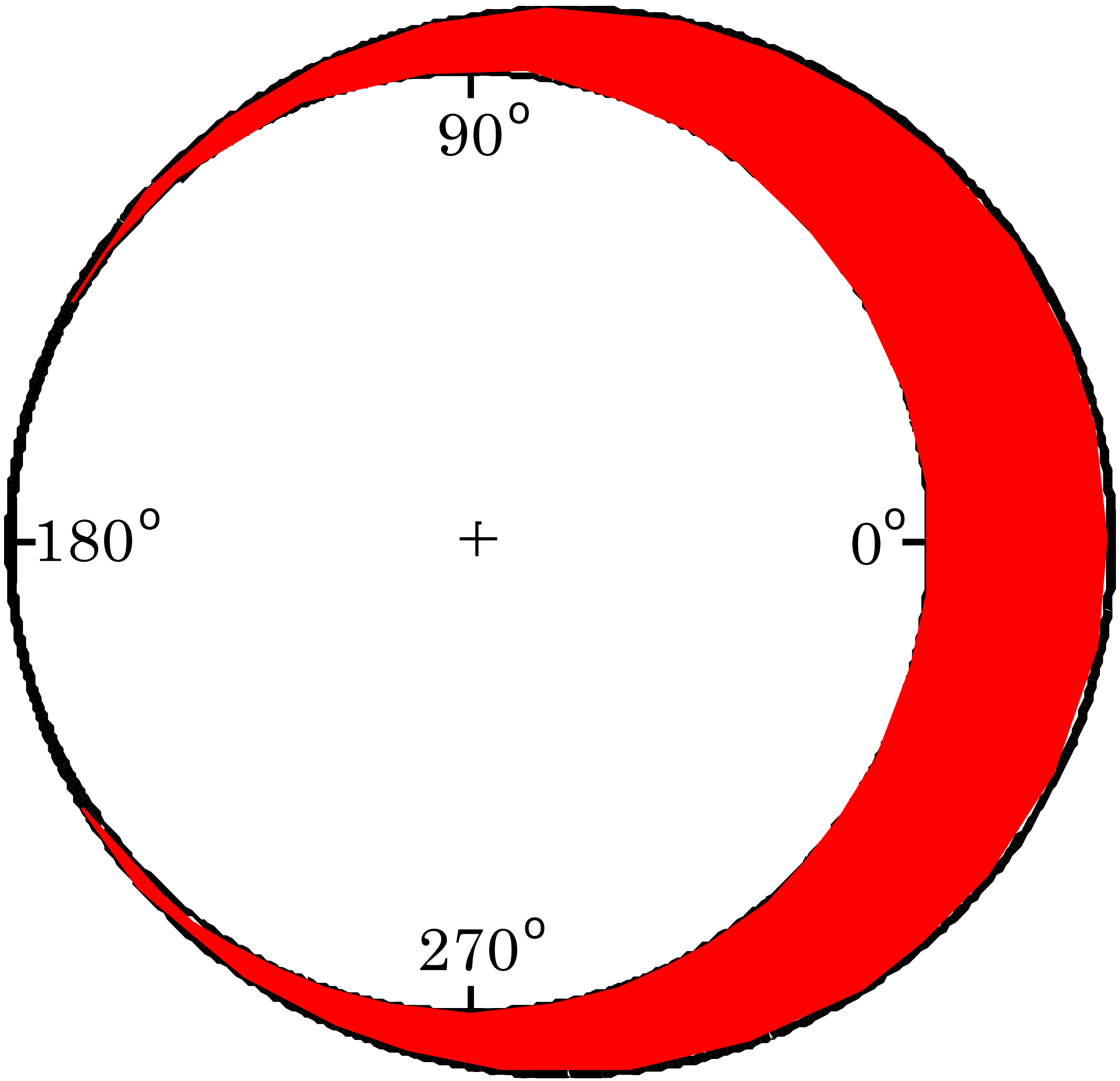
Modelling Circular Data Using the Highly Flexible Wrapped α-Stable Family of Distributions

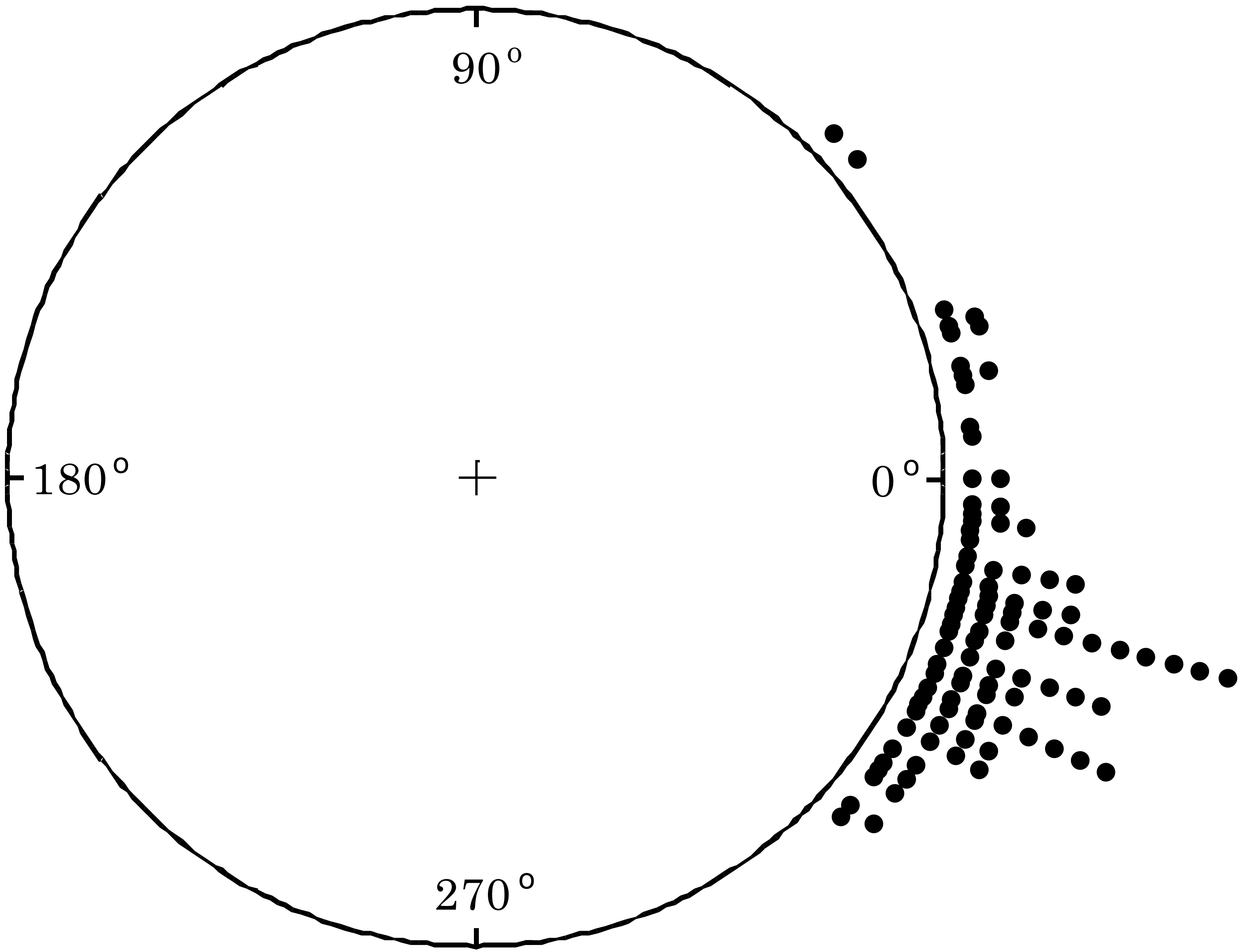
ARTHUR PEWSEY

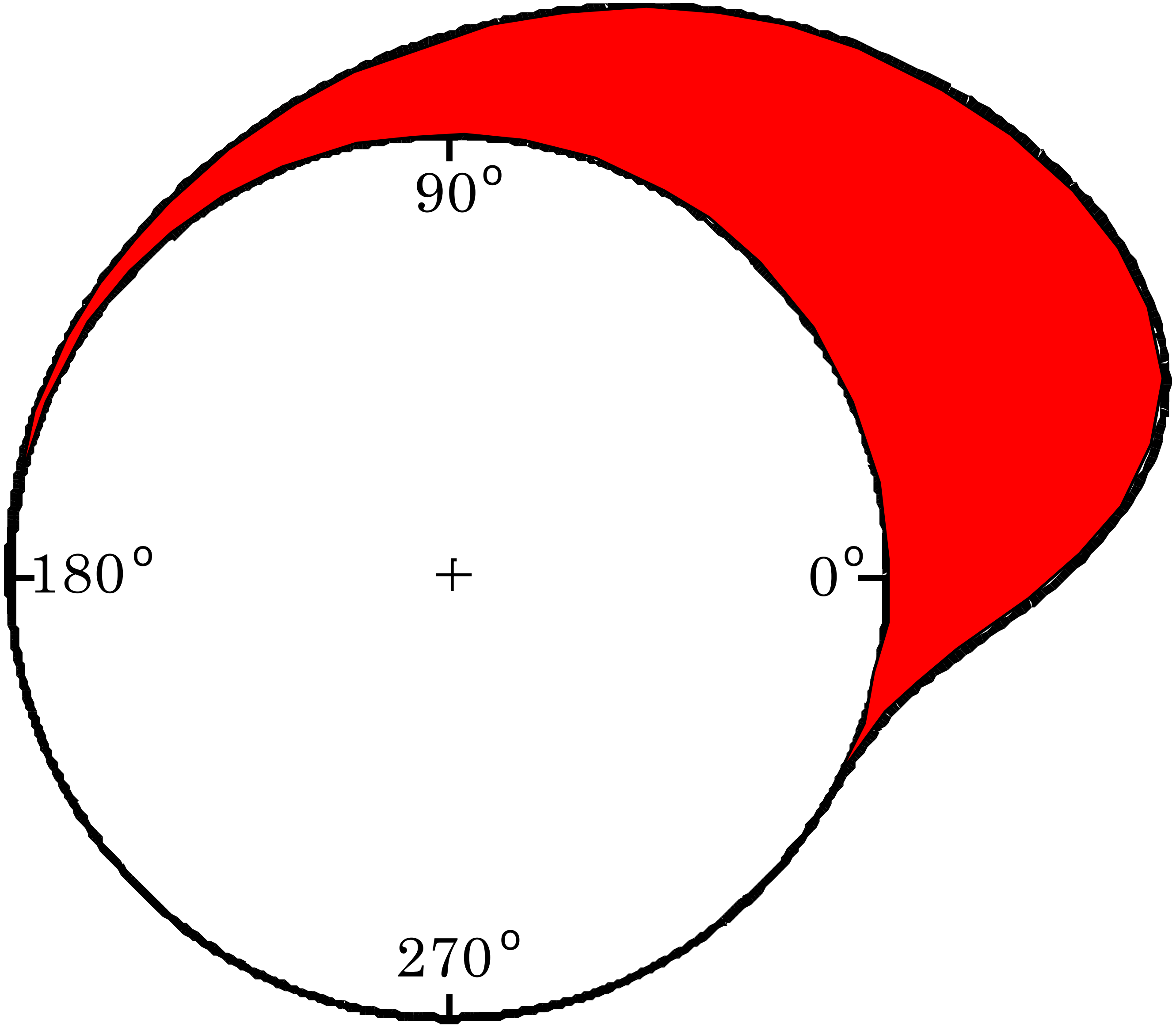
Universidad de Extremadura, Cáceres, Spain

apewsey@guadiana.unex.es


90°
180°
0°
270°


90°
180°
0°
270°


90°
180°
0°
270°


90°
180°
0°
270°

Overview

1. Introduction
    - Linear Stable Distributions
    - Circular Models
2. Definition and Fundamental Properties
    - Characteristic Function
    - Trigonometric Moments
    - Density
3. Estimation
    - Method of Moments
    - Maximum Likelihood
4. Example
    - Headings of Migrating Birds
5. Extension to Finite Mixture Modelling

1. Introduction

1.1 Linear Stable Distributions

- Four-parameter stable family of distributions on $\Re$ - Lévy (1924).

  Feller (1971), Zolotarev (1986), Samorodnitsky & Taqqu (1994), Nolan (2005) - highly extensive bibliography.

- Densities are unimodal (Yamazato 1978) and infinitely differentiable.

Further Background

- Only known cases with closed form expressions: normal, Cauchy and Lévy distributions.
- Important theoretically because of the Generalised Central Limit Theorem.
- Useful because of its great modelling flexibility: symmetric and asymmetric cases; can model peakedness and heavy tailedness.
- Characteristic function. Nolan (2004) refers to 11 different parametrisations.

Variant of (M) parametrisation of Zolotarev (1986)

- Recommended for conducting numerical work and modelling.
- Simplest location-scale parametrisation that is jointly continuous in all four parameters.
- $X \sim S(\alpha, \beta, \gamma, \delta_0; 0)$ if $\phi_X(t) = E\left(e^{itX}\right)$ is given by

$$\phi_X(t) = \begin{cases} \exp\left(-\gamma^{\alpha}|t|^{\alpha}\left[1 + i\beta\,\mathrm{sign}(t)\tan\left(\frac{\pi\alpha}{2}\right)\left\{(\gamma|t|)^{1-\alpha} - 1\right\}\right] + i\delta_0 t\right), & \alpha \neq 1, \\ \exp\left(-\gamma|t|\left[1 + i\beta\,\mathrm{sign}(t)\frac{2}{\pi}\log(\gamma|t|)\right] + i\delta_0 t\right), & \alpha = 1. \end{cases}$$

$\alpha \in (0,2]$ - index of stability; $\beta \in [-1,1]$ - skewness parameter;

$\gamma > 0$ - scale parameter; $\delta_0 \in \Re$ - location parameter.

Variant of (M) parametrisation of Zolotarev (1986)

$$\phi_X(t) = \begin{cases} \exp\left(-\gamma^{\alpha}|t|^{\alpha}\left[1 + i\beta \,\mathrm{sign}(t)\tan\left(\frac{\pi\alpha}{2}\right)\left\{(\gamma|t|)^{1-\alpha} - 1\right\}\right] + i\delta_0 t \right), & \alpha \neq 1, \\ \exp\left(-\gamma|t|\left[1 + i\beta \,\mathrm{sign}(t)\frac{2}{\pi}\log(\gamma|t|)\right] + i\delta_0 t \right), & \alpha = 1. \end{cases}$$

- α and β determine distribution shape; α controls the peakedness and tails, β controls skewness (in general).
- Symmetric about δ_0 when $\beta = 0$.
- Generally, "totally skewed to the right (left)" if $\beta = 1$ ($\beta = -1$).

Variant of (M) parametrisation of Zolotarev (1986)

$$\phi_X(t) = \begin{cases} \exp\left(-\gamma^\alpha |t|^\alpha \left[1 + i\beta \,\mathrm{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right)\left\{(\gamma|t|)^{1-\alpha} - 1\right\}\right] + i\delta_0 t\right), & \alpha \neq 1, \\ \exp\left(-\gamma|t|\left[1 + i\beta \,\mathrm{sign}(t) \frac{2}{\pi} \log(\gamma|t|)\right] + i\delta_0 t\right), & \alpha = 1. \end{cases}$$

- Normal: α=2, β=0. Cauchy: α=1, β=0. Lévy: α=1/2, β=1.
- As $\alpha \to 2$, β has progressively less effect.
- As $\alpha \to 0$, density becomes increasingly spiked.
- Normal distribution - only case with finite variance.
  $S(2, 0, \gamma, \delta_0;0) = N(\delta_0, 2\gamma^2)$.
- For $0<\alpha<2$, $E\left(|X|^p\right)$ only finite for $0<p<\alpha$.

1.2 Circular Models

- Models most commonly referred to (von Mises, wrapped normal, wrapped Cauchy, cardioid,...) all symmetric.

- Circular data seldom symmetrically distributed (Mardia 1972, p. 10).

- (Few) parametric distributions capable of modelling asymmetry:
    1. Projected normal distribution - Mardia (1972, p.52)
    2. Generalisation of the von Mises distribution - Cox (1975)
    3. Densities defined as non-negative finite trigonometric sums - Fernández-Durán (2004)
    4. Wrapped skew-normal distribution - Pewsey (2000)

Wrapped Stable Family

- First appearance - Mardia (1972, p. 57): reference to family’s density, based on the standard parametrisation used in analytical studies of linear stable distributions.

- Family's three-parameter wrapped symmetric stable sub-class considered by: Jammalamadaka & SenGupta (2001, sec. 2.2.8), SenGupta & Pal (2001), Gatto & Jammalamadaka (2003).

2. Definition and Fundamental Properties

2.1 Characteristic Function

- Consider the linear random variable $X \sim S(\alpha, \beta, \gamma, \delta_0; 0)$.
- Define the circular random variable $\Theta = X \pmod{2\pi} \in [0, 2\pi)$.
- Then $\Theta \sim WS(\alpha, \beta, \gamma, \delta_0^*; 0)$ where $\delta_0^* = \delta_0 \pmod{2\pi} \in [0, 2\pi)$.
- Elements of the characteristic function of Θ are given by $\phi_\Theta(p) = \phi_X(p)$. For $p = 1, 2, \ldots$

$$\phi_\Theta(p) = \begin{cases} \exp\left(-\gamma^\alpha p^\alpha \left[1 + i\beta \tan\left(\frac{\pi\alpha}{2}\right)\left\{(\gamma p)^{1-\alpha} - 1\right\}\right] + i\delta_0^* p\right), & \alpha \neq 1, \\ \exp\left(-\gamma p\left[1 + i\beta \frac{2}{\pi}\log(\gamma p)\right] + i\delta_0^* p\right), & \alpha = 1. \end{cases}$$

2.2 Trigonometric Moments

- $\phi_\Theta(p) = E\left(e^{ip\Theta}\right) = \alpha_p + i\beta_p$, where α_p and β_p are the *p*th cosine and sine moments.

- $\alpha_p = \rho_p \cos(\mu_p)$ and $\beta_p = \rho_p \sin(\mu_p)$, where $\rho_p \in [0,1]$ and $\mu_p \in [0,2\pi)$ are the *p*th mean resultant length and *p*th mean direction.

- For $p = 1,2,...$

$$\rho_p = \exp\left\{-(\gamma p)^\alpha\right\} = \rho^{p^\alpha},$$

where $\rho = \rho_1$ is the mean resultant length, and

$$\mu_p = \begin{cases} \delta_0^* p + \beta \tan\left(\dfrac{\pi\alpha}{2}\right)\left\{(\gamma p)^\alpha - \gamma p\right\} \pmod{2\pi}, & \alpha \neq 1, \\ \delta_0^* p - \beta \dfrac{2}{\pi} \gamma p \log(\gamma p) \pmod{2\pi}, & \alpha = 1. \end{cases}$$

Trigonometric Moments About Mean Direction

- *p*th cosine and sine moments about mean direction, $\boxed{\mu = \mu_1}$:

$$\bar{\alpha}_p = E[\cos\{p(\Theta-\mu)\}] = \begin{cases} \rho_p \cos\left\{\beta \tan\left(\dfrac{\pi\alpha}{2}\right)\gamma^{\alpha}\left(p^{\alpha}-p\right)\right\}, & \alpha \neq 1, \\ \rho_p \cos\left\{-\beta \dfrac{2}{\pi}\gamma p \log(p)\right\}, & \alpha = 1, \end{cases}$$

and

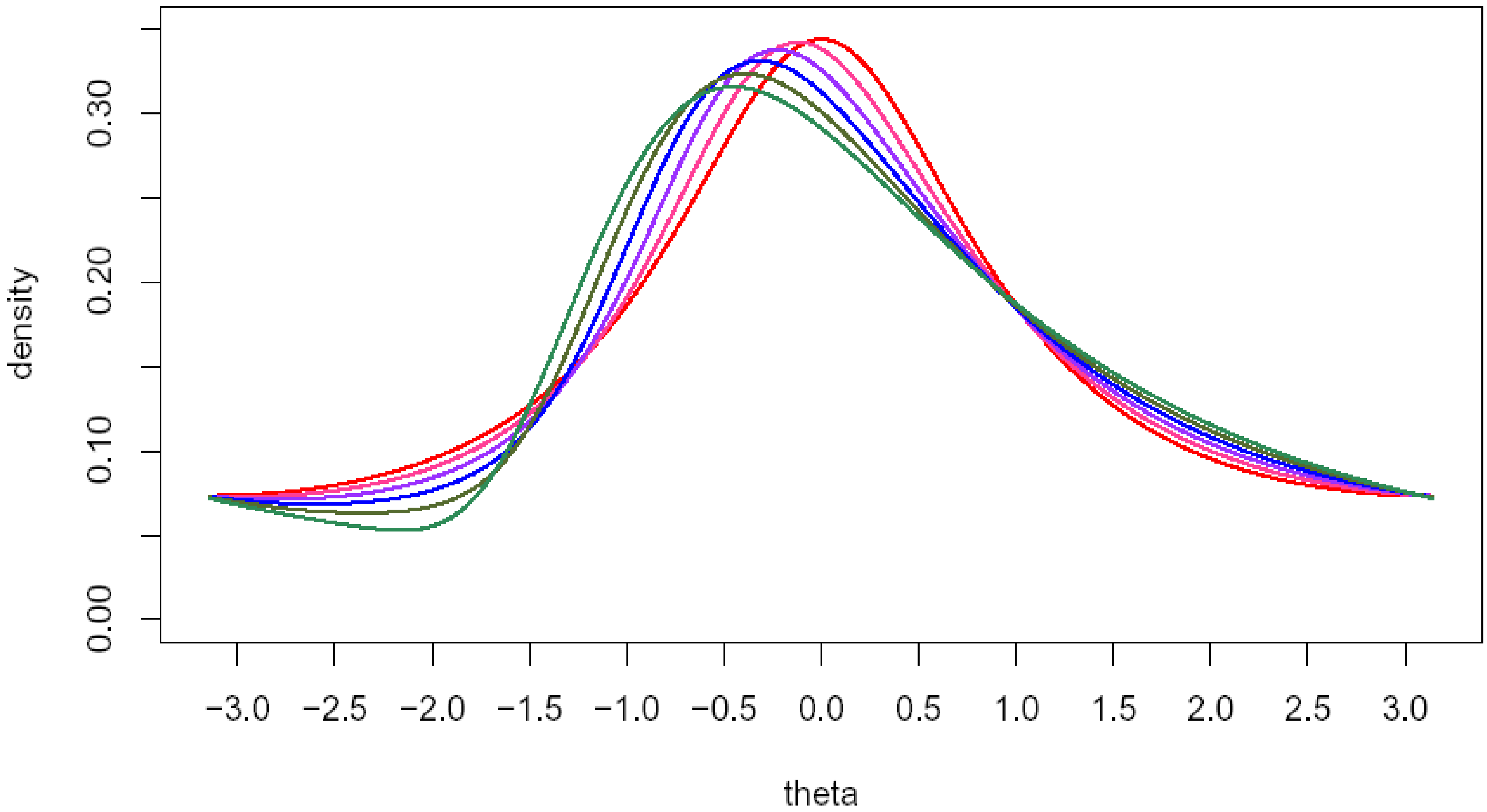
$$\bar{\beta}_p = E[\sin\{p(\Theta-\mu)\}] = \begin{cases} \rho_p \sin\left\{\beta \tan\left(\dfrac{\pi\alpha}{2}\right)\gamma^{\alpha}\left(p^{\alpha}-p\right)\right\}, & \alpha \neq 1, \\ \rho_p \sin\left\{-\beta \dfrac{2}{\pi}\gamma p \log(p)\right\}, & \alpha = 1. \end{cases}$$

2.4 Density

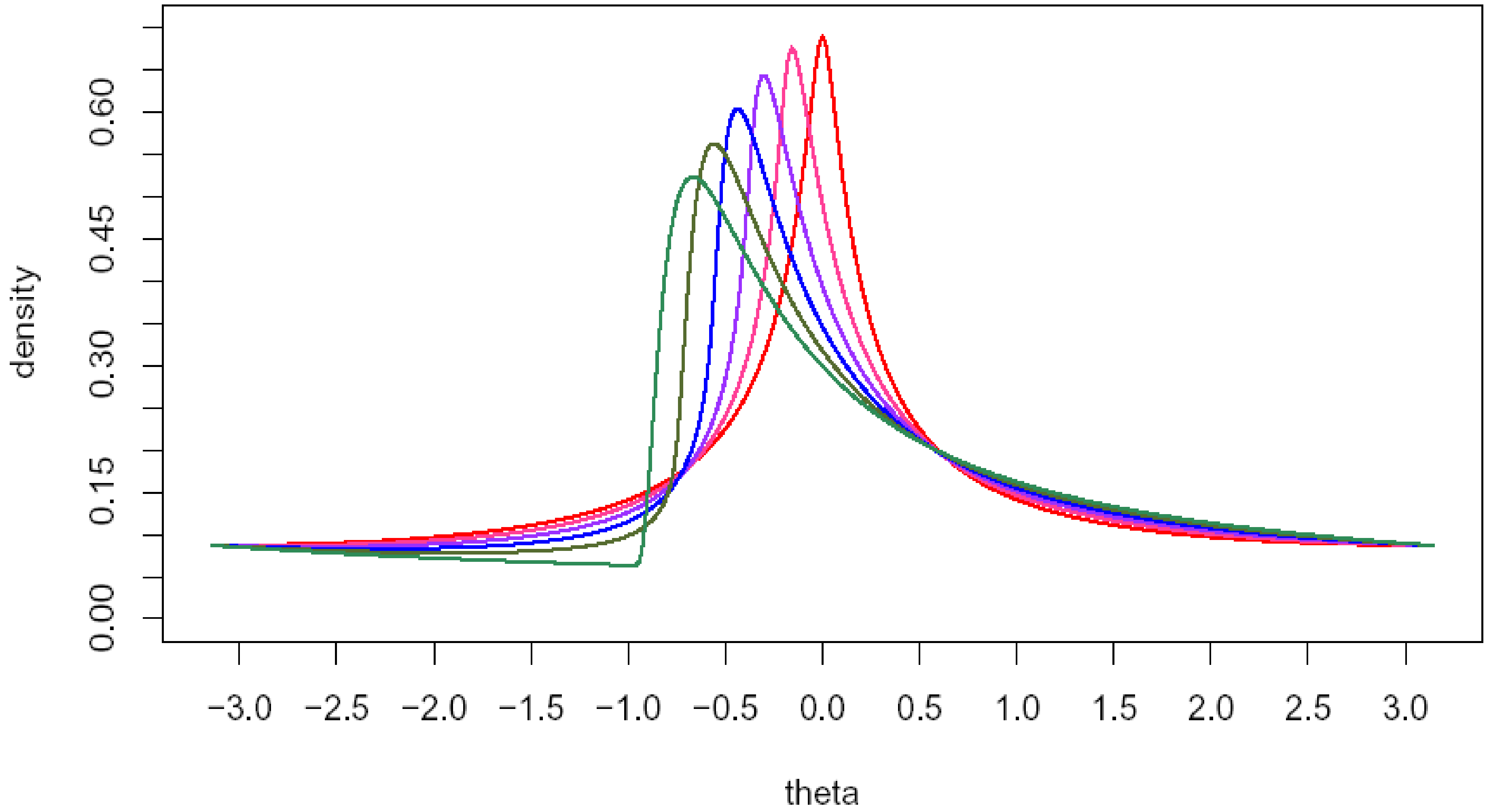
- Denoting $(\alpha, \beta, \gamma, \delta_0^*)$ by $\underline{\xi}$, the density of Θ can be represented as

$$f(\theta;\underline{\xi}) = \frac{1}{2\pi}\left\{1 + 2\sum_{p=1}^{\infty}\rho_p \cos(p\theta - \mu_p)\right\}.$$

- Wrapped normal distribution obtained when $\alpha = 2$ (irrespective of value of β).
- No skew extensions of the wrapped normal distribution; wrapped skew-normal class is complementary to wrapped stable family.
- As $\alpha \to 0$, density becomes increasingly more spiked.
- Must approximate density.



Examples of wrapped stable densities with $\boxed{\alpha = 1}$, $\gamma = 1$, $\delta_0^* = 0$ and, starting from the mode of the symmetric density (in red) and working left, $\beta = 0\ (0.2)\ 1$. Support is $[-\pi, \pi]$.



Examples of wrapped stable densities with $\boxed{\alpha = 1/2}$, $\gamma = 1$, $\delta_0^* = 0$ and, starting from the mode of the symmetric density (in red) and working left, β = 0 (0.2) 1. Support is [-π, π].

3. Estimation

3.1 Method of Moments

- $\underline{\theta} = (\theta_1, \ldots, \theta_n)$ - random sample of n observations from a WS(α, β, γ, δ_0^*;0) distribution.
- pth sample cosine and sine moments: $a_p = \frac{1}{n}\sum_{i=1}^{n}\cos(p\theta_i)$ and $b_p = \frac{1}{n}\sum_{i=1}^{n}\sin(p\theta_i)$.
- pth sample mean resultant length: $\bar{R}_p = (a_p^2 + b_p^2)^{1/2}$.

Mean Direction and Associated Trigonometric Moments

- If $\bar{R}_p = 0$, pth sample mean direction, $\bar{\theta}_p$, undefined. Otherwise,

$$\bar{\theta}_p = \begin{cases} \tan^{-1}\left(b_p / a_p\right), & a_p > 0, b_p \geq 0, \\ \pi/2, & a_p = 0, b_p > 0, \\ \tan^{-1}\left(b_p / a_p\right) + \pi, & a_p < 0, \\ \tan^{-1}\left(b_p / a_p\right) + 2\pi, & a_p \geq 0, b_p < 0, \end{cases}$$

where $\tan^{-1}$ takes values in [-π/2, π/2].

- pth sample cosine and sine moments about mean direction,

$\boxed{\bar{\theta} = \bar{\theta}_1}$: $\bar{a}_p = \frac{1}{n}\sum_{i=1}^{n}\cos\left(p\left(\theta_i - \bar{\theta}\right)\right)$ and $\bar{b}_p = \frac{1}{n}\sum_{i=1}^{n}\sin\left(p\left(\theta_i - \bar{\theta}\right)\right)$.

- Population analogues of $\bar{R}_p$, $\bar{\theta}_p$, $\bar{a}_p$ and $\bar{b}_p$ are ρ_p, μ_p, $\bar{\alpha}_p$ and $\bar{\beta}_p$.

Moment Estimates

- Equating $\bar{R}_1$, $\bar{R}_2$, $\bar{b}_2$ and $\bar{\theta}$ with ρ, ρ_2, $\bar{\beta}_2$ and μ, respectively:

$$\tilde{\alpha} = \frac{\log\{\log(\bar{R}_2)/\log(\bar{R}_1)\}}{\log(2)},$$

$$\tilde{\gamma} = \begin{cases} \exp[\log\{-\log(\bar{R}_1)\}/\tilde{\alpha}], & \tilde{\alpha} \neq 1, \\ -\log(\bar{R}_1), & \tilde{\alpha} = 1, \end{cases}$$

$$\tilde{\beta} = \begin{cases} \sin^{-1}(\bar{b}_2/\bar{R}_2)/\{\tan(\tilde{\alpha}\pi/2)\tilde{\gamma}^{\tilde{\alpha}}(2^{\tilde{\alpha}}-2)\}, & \tilde{\alpha} \neq 1, \\ \sin^{-1}(\bar{b}_2/\bar{R}_2)/\{-(4/\pi)\tilde{\gamma}^{\tilde{\alpha}}\log(2)\}, & \tilde{\alpha} = 1, \end{cases}$$

$$\tilde{\delta}_0^* = \begin{cases} \bar{\theta} - \tilde{\beta}\tan(\tilde{\alpha}\pi/2)(\tilde{\gamma}^{\tilde{\alpha}} - \tilde{\gamma}) \pmod{2\pi}, & \tilde{\alpha} \neq 1, \\ \bar{\theta} + (2/\pi)\tilde{\beta}\tilde{\gamma}\log(\tilde{\gamma}) \pmod{2\pi}, & \tilde{\alpha} = 1. \end{cases}$$

- Problems: $\boxed{\tilde{\alpha}}$ and $\boxed{\tilde{\beta}}$ do not always lie in (0,2] and [-1,1].

3.2 Maximum Likelihood

- Constrained numerical maximisation of log-likelihood function

$$\ell(\underline{\xi};\underline{\theta}) = -n\log(2\pi) + \sum_{i=1}^{n}\log\left\{1 + 2\sum_{p=1}^{\infty} A_{pi}\right\},$$

where $A_{pi} = \rho_p \cos(p\theta_i - \mu_p)$.

- FORTRAN program:

  a) Direct-search simplex algorithm of Nelder & Mead (1965). Only requires a reliable approximation to objective function.

  b) Designed for $\alpha > 0.4$.

  c) Execution time depends on α, n, and tolerance of approximation.

Practicalities

- Use method of moments estimates as starting values unless data distribution is close to circular uniform.

- If $\tilde{\alpha} > 2$, set starting value for α marginally below 2.

- If $|\tilde{\beta}| > 1$, can use default starting values of β=0 and $\delta_0^* = \bar{\theta}$.

- For data with a single mode, have never found multiple maxima on the log-likelihood surface.

- For data close to being uniformly distributed, use a grid based search covering [0,2π) for δ_0^*.

Other Forms of Inference

- For interior points of the parameter space: use standard asymptotic theory for confidence regions (information matrix / profile likelihood methods) and likelihood ratio testing.

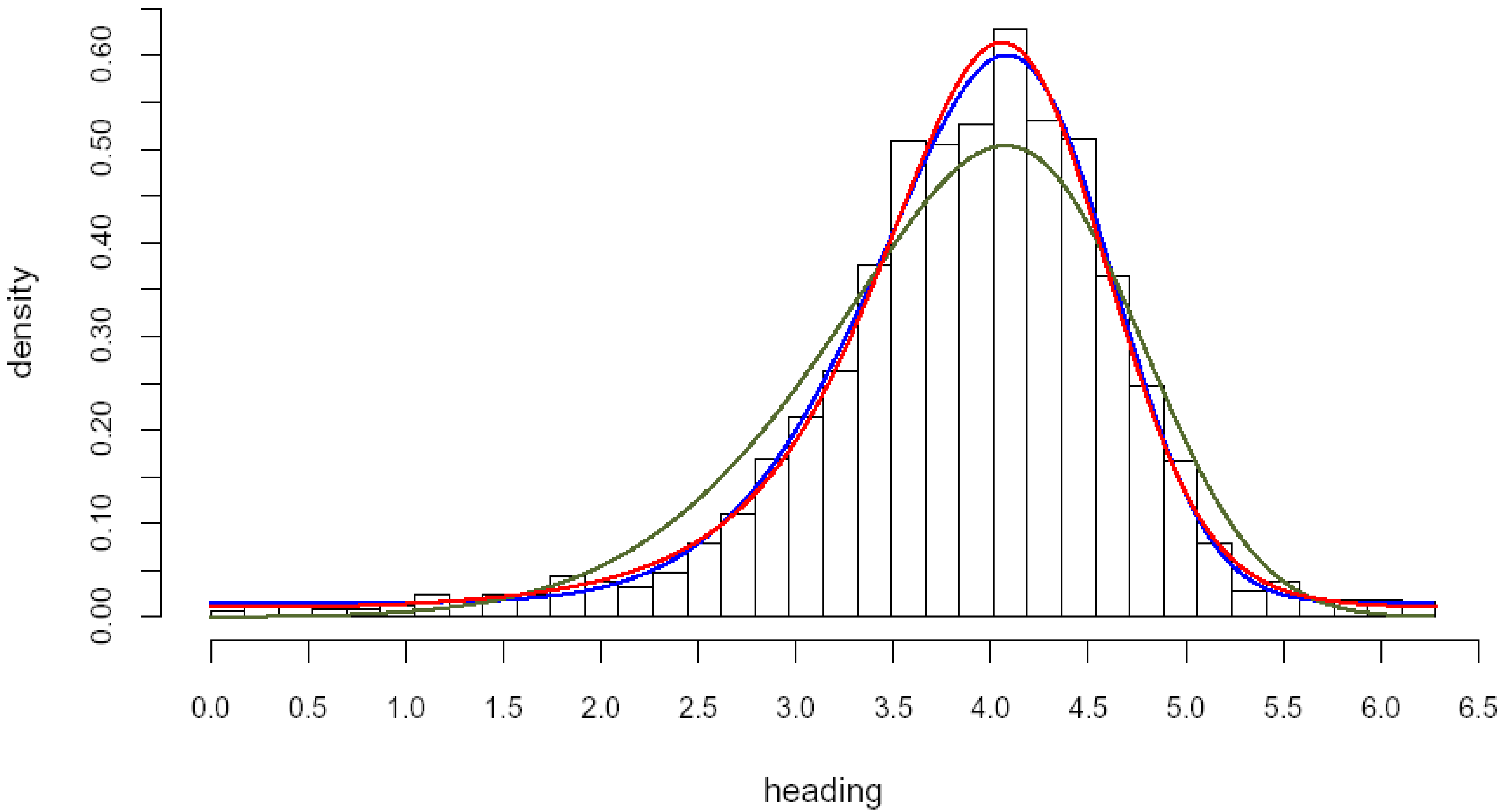
- For points on the boundary of the parameter space, usual regularity conditions do not hold.

  Example: Cannot frame null hypothesis of underlying wrapped normal distribution versus some skew alternative in terms of "skewness" parameter of the wrapped stable family. However, wrapped normal distributions correspond to interior points of the parameter space of the wrapped skew-normal class.

4. Example

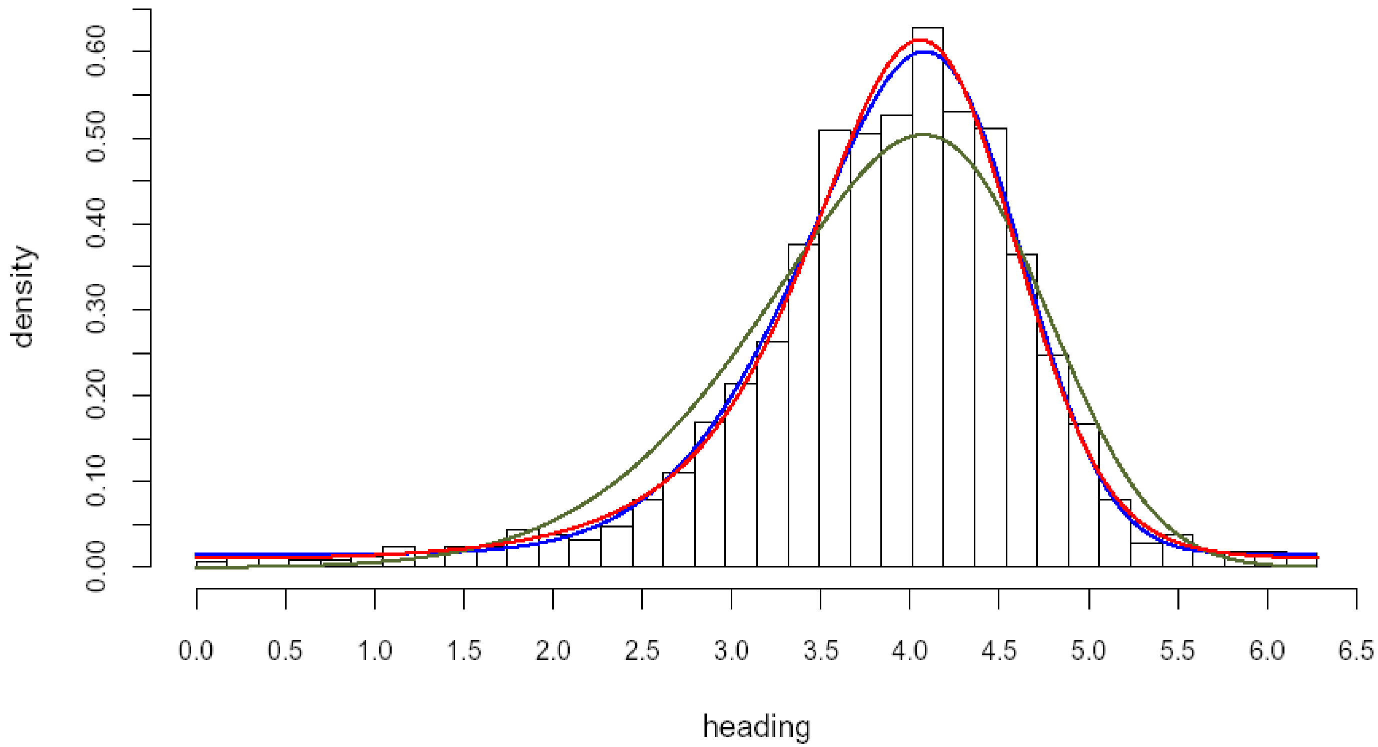
Headings of Migrating Birds

- Data set consisting of n = 1827 ‘headings’ of birds - Bruderer & Jenni (1990).
- Recorded near Stuttgart in Germany during the autumnal migration period of 1987.
- A ‘heading’ is the direction, measured clockwise from north, of a bird's body during flight.



Histogram of the headings and densities of maximum likelihood fits; wrapped stable, wrapped skew-normal and circular uniform mixture, and wrapped skew-normal. The support is $[0,2\pi)$.

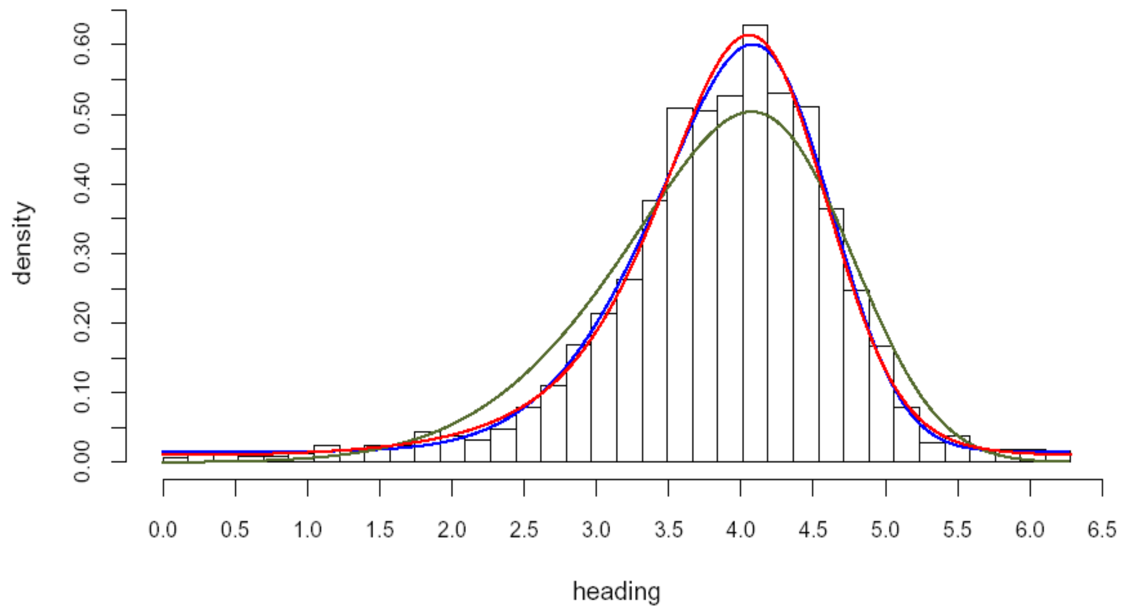
Results



Distribution	ℓ	χ^2 g-o-f
wrapped skew-normal	-2202.06	0
wrapped skew-normal + circular uniform mixture	-2128.03	0.2
wrapped stable	-2127.73	0.2

Results

- Wrapped stable fit: WS(1.54, -0.64, 0.46, 4.01;0).
- Approx. 95% CI's (standard asymptotic theory + profile likelihood methods): (1.46, 1.62), (-0.81,-0.45), (0.44, 0.48), (3.97, 4.05).
- Would appear that underlying distribution is negatively skew, with heavier than wrapped normal (α = 2) tails but lighter than wrapped skew-Cauchy (α = 1) tails.



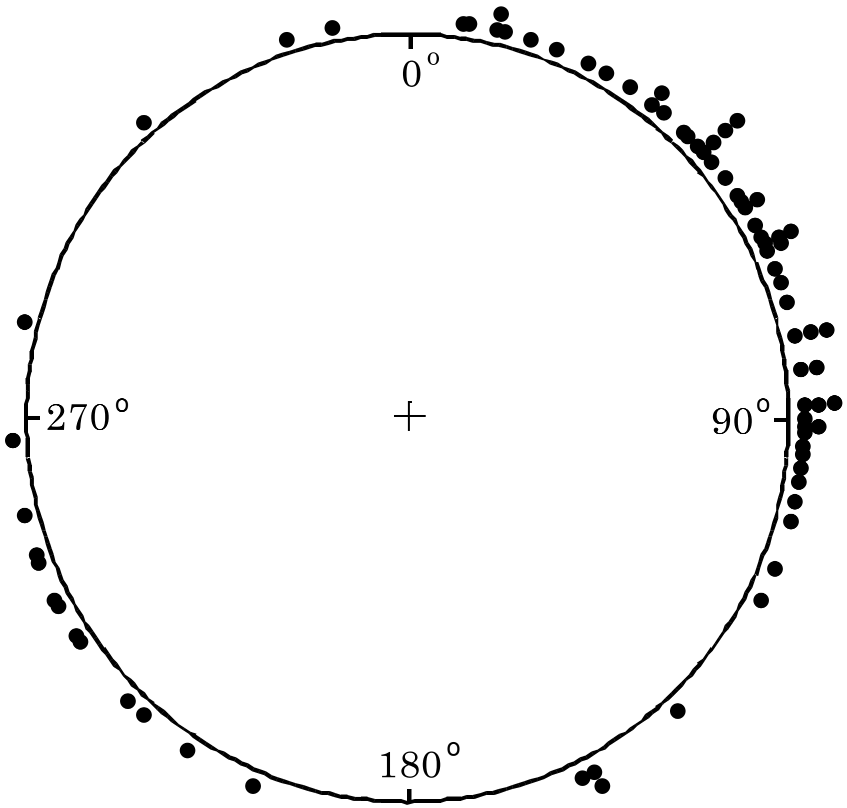
Interpretation

- Wrapped stable fit: single negatively skewed population with heavy tails.
- Mixture distribution fit: $\boxed{g(\theta) = pf(\theta) + (1-p)\frac{1}{2\pi}}$

  10%, circular uniform; 90% negatively skewed, but light-tailed, distribution.

5. Extension to Finite Mixture Modelling

- Here we have considered use of wrapped stable family as a model for unimodal circular data.
- Examples abound of circular data sets with $m \geq 2$ modes.



- Finite mixtures of wrapped stable distributions provide highly flexible models for multimodal circular data.

- Potentially rather parameter heavy (no. of parameters is $5m$-1).

- Experience of finite mixture modelling with wrapped stable component distributions most encouraging - using obvious extensions of the likelihood based methodology described here.

Brief Summary

We have:

1. Considered a parametrisation of linear stable distributions and wrapped circular analogue.

2. Presented family's basic properties.

3. Studied parameter estimation using the method of moments and maximum likelihood.

4. Illustrated the application of the family in the modelling of a large skew circular data set.

5. Referred to the use of the family in finite mixture modelling.

References

Bruderer, B., & Jenni, L. (1990), "Migration Across the Alps," in *Bird Migration: Physiology and Ecophysiology*, ed. E. Gwinner, Berlin: Springer, pp. 60-77.

Cox, D.R. (1975), Discussion of "Statistics of Directional Data," by K.V. Mardia, *Journal of the Royal Statistical Society*, Ser. B, 37, 380-381.

Feller, W. (1971), *An Introduction to Probability Theory and its Applications* (Vol. 2; 2nd ed.), New York: Wiley.

Fernández-Durán, J.J. (2004), "Circular Distributions Based on Nonnegative Trigonometric Sums," *Biometrics*, 60, 499-503.

Gatto, R., & Jammalamadaka, S.R. (2003), "Inference for Wrapped Symmetric Alpha-Stable Circular Models," *Sanhkya*, 65, 333-355.

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

Lévy, P. (1924), "Théorie des Erreurs la Loi de Gauss et les Lois Exceptionnelles," *Bulletin de la Société Mathématique de France*, 52, 49-85.

Mardia, K.V. (1972), *Statistics of Directional Data*, London: Academic Press.

Nelder, J.A., & Mead, R. (1965), "A Simplex Method for Function Minimization," *Computer Journal*, 7, 308-313.

Nolan, J.P. (2004), "Stable Distributions: Models for Heavy Tailed Data," American University, Department of Mathematics and Statistics, available at *http://academic2.american.edu/~jpnolan/stable/chap1.pdf*.

Nolan, J.P. (2005), "Bibliography on Stable Distributions, Processes and Related Topics," American University, Department of Mathematics and Statistics, available at *http://academic2.american.edu/~jpnolan/stable/StableBibliography.pdf*.

Pewsey, A. (2000), "The Wrapped Skew-Normal Distribution on the Circle," *Communications in Statistics - Theory and Methods*, 29, 2459-2472.

Pewsey, A. (2006), "Modelling Asymmetrically Distributed Circular Data Using the Wrapped Skew-Normal Distribution," *Environmental and Ecological Statistics*. In press.

Samorodnitsky, G., & Taqqu, M.S. (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, New York: Chapman & Hall/CRC.

SenGupta, A., & Pal, C. (2001), "On Optimal Tests for Isotropy Against the Symmetric Wrapped Stable-Circular Uniform Mixture Family," *Journal of Applied Statistics*, 28, 129-143.

Yamazato, M. (1978), "Unimodality of Infinitely Divisible Distribution Functions of Class L," *The Annals of Probability*, 6, 523-531.

Zolotarev, V.M. (1986), *One-dimensional Stable Distributions*, Vol. 65 of "Translations of Mathematical Monographs", Providence, RI: American Mathematical Society. Translation from the original 1983 Russian edition.