Modelling Asymmetrically Distributed Circular Data

Arthur Pewsey
Departmento de Matemáticas
Universidad de Extremadura
1. Statistical Context

Circular Data \{ Spherical Data \} \quad \text{Directional Data}

Modelling data on different manifolds
(Kanti Mardia, John Kent)
2. Some Data

Data Set 1. Initial headings of 230 Chinese painted quail on exit from a straight, 1m long, corridor. The zero direction corresponds to the orientation of the corridor.
Data Set 2. Orientations of 730 red wood ants in relation to a black target placed at an angle of $180^\circ$ from the zero direction. Each dot represents the direction followed by five ants.
Data Set 3. Orientations, measured in a clockwise direction from North (in degrees), of 76 turtles after egg laying.
3. Data Sources

Animal orientation experiments

**Earth Sciences:** Paleocurrents, Paleomagnetic directions

**Medicine:** Posture, Flexion of limbs

**Meteorology:** Wind directions, Times of day of thunderstorms

**Ecology:** Direction of wind (transportation of pollutants)

**Psychology:** Mental maps (representing surroundings)

**Image Analysis:** Machine vision, Orientation of textures, Orientation of ridges on fingerprints
4. Linear/Circular Geometry
5. Vector Resolution
6. Circular Statistics

\( \theta \) - random angle; \( \theta_1, \ldots, \theta_n \) - random sample of \( n \) observations

Sample mean resultant length
\[
\bar{R} = \sqrt{a_1^2 + b_1^2}
\]

\[
a_p = \frac{1}{n} \sum_{i=1}^{n} \cos p \theta_i, \quad b_p = \frac{1}{n} \sum_{i=1}^{n} \sin p \theta_i
\]

\( p \)th order trigonometric moments about the zero direction

If \( \bar{R} = 0 \), sample mean direction is undefined

If \( \bar{R} > 0 \), \( \bar{\theta} = \begin{cases} 
\tan^{-1} \left( \frac{b_1}{a_1} \right) & \text{if } a_1 \geq 0 \\
\pi + \tan^{-1} \left( \frac{b_1}{a_1} \right) & \text{if } a_1 < 0 
\end{cases} \)

where \( \tan^{-1}(x) \in [-\pi/2, \pi/2] \)
Sample trigonometric moments about $\bar{\theta}$

$$\bar{a}_p = \frac{1}{n} \sum_{i=1}^{n} \cos p(\theta_i - \bar{\theta}) \quad \text{and} \quad \bar{b}_p = \frac{1}{n} \sum_{i=1}^{n} \sin p(\theta_i - \bar{\theta})$$

$$\bar{b}_1 = \frac{1}{n} \sum_{i=1}^{n} \sin (\theta_i - \bar{\theta}) = 0$$

$$\bar{b}_2 = \frac{1}{n} \sum_{i=1}^{n} \sin 2(\theta_i - \bar{\theta}) \quad \text{Measure of circular skewness}$$

Batschelet (1965)
7. Population Characteristics

Characteristic function

\[ \{ \psi_p : p = 0, \pm 1, \ldots \} \text{ for } \psi_p = \alpha_p + i\beta_p. \]

\[ \alpha_p = E(\cos p\theta), \quad \beta_p = E(\sin p\theta) \]

**Trigonometric moments of \( \theta \) about the zero direction**

\[ p = 1 : \quad \psi_1 = \alpha_1 + i\beta_1 = \rho e^{i\mu} \]

\( \mu, \) mean direction; \( \rho, \) mean resultant length

**Trigonometric moments about \( \mu \)**

\[ \overline{\alpha}_p = E\{ \cos p(\theta - \mu) \} \quad \text{and} \quad \overline{\beta}_p = E\{ \sin p(\theta - \mu) \} \]
8. Symmetric Models

8.1 Non-wrapped

Uniform

No preferred direction; Limiting distribution of circular analogue of CLT

$$f(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta < 2\pi,$$
von Mises $\text{VM}(\mu, \kappa)$

Mardia & Jupp (1999), five constructions; Assumed advanced parametric methods

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\},$$

$0 \leq \mu < 2\pi$, mean direction; $\kappa > 0$, concentration parameter; $I_p(.)$, modified Bessel function - first kind, order $p$; $\rho = A(\kappa) = I_1(\kappa)/I_0(\kappa) \in [0,1]$, mean resultant length.
8.2 Wrapped Schmidt(1917)

General Result

If the linear random variable $X$ has c.f. $\psi(t)$ then the c.f. of $\Theta = X(\text{mod } 2\pi)$ is

$$\{\psi(p) : p = 0, \pm 1, \ldots\}.$$

Wrapped Normal $WN(\mu, \sigma)$

Position of particle under Brownian motion on the circle. $VM(\mu, \kappa) \approx WN(\mu, \sigma)$, $\sigma = \sqrt{-2 \log[A(\kappa)]}$. Discrimination Pewsey & Jones (2003).

$$f(\theta; \mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \sum_{k=-\infty}^{\infty} \exp \left\{ -\frac{(\theta - \mu - 2\pi k)^2}{2\sigma^2} \right\}.$$
Wrapped Cauchy

Heavier “tails” than wrapped normal; density can be expressed in closed form.

\[ f(\theta; \mu, \rho) = \frac{1}{2\pi} \left( \frac{1 - \rho^2}{1 + \rho^2 - \rho \cos(\theta - \mu)} \right), \quad \rho \in [0, 1]. \]

Symmetric wrapped \( \alpha \)-stable \text{ Mardia (1972)}

Includes wrapped normal, wrapped Cauchy, ... (but not von Mises)

\[ f(\theta; \mu, \alpha, \varsigma) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \exp(-\varsigma^k \alpha) \cos\{k(\theta - \mu)\}, \]

where \( \alpha \in (0, 1) \cup (1, 2] \) and \( \varsigma \geq 0. \)

Closed form; Includes von Mises, wrapped Cauchy, ... (but not wrapped normal)

\[
f(\theta; \mu, \kappa, \psi) = \frac{\{1 + \tanh(\kappa \psi) \cos(\theta - \mu)\}^{1/\psi}}{2\pi \, P_{\nu}(\cosh(\kappa \psi))},
\]

\(\kappa \geq 0, \ -\infty < \psi < \infty, \ P_{\nu}(\cdot) \) associated Legendre function - first kind, degree \(\nu\), order 0.

Useful Mixture

Modelling heavy tails; Uniform “background” & base “foreground”.

\[g(\theta) = pf(\theta) + (1 - p) \frac{1}{2\pi}\]
Symmetric densities with mean direction 0 and mean resultant length 0.45: a) wrapped normal; b) wrapped Cauchy; c) uniform and wrapped normal mixture.
Data Set 2. Orientations of 730 red wood ants in relation to a black target placed at an angle of 180° from the zero direction. Each dot represents the direction followed by five ants.
# 9.1 Summary of Fits

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\psi$ or $\alpha$</th>
<th>$\mu$</th>
<th>$\kappa$ or $\zeta$</th>
<th>$p$</th>
<th>Log-likelihood</th>
<th>$\chi^2$ g-o-f</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP</td>
<td>-1.30</td>
<td>-3.1</td>
<td>1.6</td>
<td>1</td>
<td>-926.55</td>
<td>0.07</td>
</tr>
<tr>
<td>JP+U</td>
<td>-0.10</td>
<td>-3.1</td>
<td>5.01</td>
<td>0.67</td>
<td>-918.72</td>
<td>0.32</td>
</tr>
<tr>
<td>SWS+U</td>
<td>1.95</td>
<td>-3.1</td>
<td>0.93</td>
<td>0.67</td>
<td>-918.89</td>
<td>0.30</td>
</tr>
<tr>
<td>VM+U</td>
<td>0</td>
<td>-3.1</td>
<td>7.54</td>
<td>0.66</td>
<td>-918.80</td>
<td>0.35</td>
</tr>
<tr>
<td>WN+U</td>
<td>2</td>
<td>-3.1</td>
<td>0.93</td>
<td>0.66</td>
<td>-918.90</td>
<td>0.34</td>
</tr>
</tbody>
</table>
9.2 Graphical Representation of Fits

Linear histogram of red ant orientation data together with densities: Jones-Pewsey (Blue); Jones-Pewsey + Uniform (Red); von Mises + Uniform (Green)
Data Set 4. Initial headings of 100 Chinese painted quail on exit from a dog-leg corridor. The zero direction corresponds to the orientation of the last 0.5m section of the corridor.
**Data Set 5.** Linear histogram of 1827 bird-flight headings taken from Bruderer & Jenni (1990).
11. “Problem” of Asymmetry

11.1 Asymmetry as a General Issue

Virtually all of established methodology for the analysis of circular data assumes symmetry.

Symmetry is a tacit assumption in many areas of Statistics (indeed in the Arts and Sciences in general).

Asymmetry not a “problem” unique to the modelling of circular data.
11.2 Asymmetric Data on the Line

**Box-Cox transformation**: bias on transforming back

Increased interest in modelling data on the *scale they were originally observed*.

- Arnold & co-authors (2000, 2002, ...)

11.3 Asymmetric Compositional Data

MOVE/STAY methodology of Aitchinson, Pawlowsky-Glahn and co-workers (Girona/Barcelona).

11.4 Asymmetric Circular Data

Modelling on the original “scale” is an absolute necessity as circle is compact. There does not even exist an equivalent to “standardisation” on the circle. Any form of transformation other than rotation or reflection changes the relative positions of the observations.
12. Models Capable of Modelling Asymmetry

12.1 Papakonstantinou (1979)

\[ f(\theta; \kappa, \nu) = \frac{1}{2\pi} + \frac{\kappa}{2\pi} \sin(\theta + \nu \sin \theta), \]

where \(|\kappa| \leq 1|\) and \(|\nu| < 1\). Shape depends on both \(\kappa\) and \(\nu\), the second determining skewness.

12.2 Batschelet (1981)

Extension of von Mises distribution.

\[ f(\theta; \kappa, \nu) = c \exp\{\kappa \cos(\theta + \nu \cos \theta)\}, \]

\(|\nu| < 1, c\) a normalising constant. Again, \(\nu\) is skewness parameter.
Linear plots of a) Papakonstantinou and b) Batschelet densities. Both pairs of curves correspond to the choices $\kappa = 0.8$ and $\nu = 0.2$ (dashed curve) and $\nu = 0.98$ (solid curve). In b) the normalising constant $c$ has been set equal to 1.
12.3 *Wrapped Skew-normal* Pewsey(2000, 2004a)

*Wrapped normal* and *wrapped half-normal* as special cases.

\[ f(\theta; \xi, \eta, \lambda) = \frac{2}{\eta} \sum_{k=-\infty}^{\infty} \phi \left( \frac{\theta + 2\pi k - \xi}{\eta} \right) \Phi \left( \lambda \frac{\theta + 2\pi k - \xi}{\eta} \right), \]

\( \phi(\cdot) \) and \( \Phi(\cdot) \), density and distribution function of standard normal distribution: \( -\infty < \xi < \infty, \eta > 0, \) and \( -\infty < \lambda < \infty, \) location, scale and skewness parameters, respectively.
\[ \text{WSNC}(0, 1, \lambda) \] densities with: a) \( \lambda = 0 \) (wrapped standard normal); b) \( \lambda = 2 \); c) \( \lambda = 5 \); d) \( \lambda = 20 \).
12.4 Wrapped $\alpha$-stable

Capable of modelling varying degrees of asymmetry.

$$f(\theta; \alpha, \beta, \varsigma) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \exp\left( -\varsigma^a k^a \right) \cos\left( k(\theta - \mu) - \varsigma^a k^a \beta \tan \frac{\alpha \pi}{2} \right) ,$$

$\alpha \in (0,1) \cup (1,2], \ |\beta| \leq 1, \ \varsigma \geq 0$. Work in progress.

12.5 Asymmetric Jones-Pewsey

Work in progress.
13. Modelling Strategies

13.1 Strategy A

<table>
<thead>
<tr>
<th>Test for Symmetry</th>
<th>Pewsey (2002)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symmetric</strong></td>
<td></td>
</tr>
<tr>
<td>Fit individual <em>(symmetric)</em> models or flexible <em>(symmetric)</em> family (e.g. JP or SWS)</td>
<td></td>
</tr>
<tr>
<td><strong>Asymmetric</strong></td>
<td></td>
</tr>
<tr>
<td>Fit individual <em>(asymmetric)</em> models or flexible <em>(asymmetric)</em> family (e.g. WS)</td>
<td></td>
</tr>
</tbody>
</table>
13.2 Strategy B

Fit a flexible family capable of modelling both symmetry and asymmetry
(e.g. WS)

Model refinement using usual likelihood based machinery
14.1 Summaries of Fits

\[ f(\theta; p, \xi, \eta, \lambda) = \frac{(1-p)}{2\pi} + p \sum_{r=-\infty}^{\infty} \left( \frac{\theta + 2\pi \eta r - \xi}{\eta} \right) \Phi \left\{ \lambda \left( \frac{\theta + 2\pi \eta r - \xi}{\eta} \right) \right\} \]

<table>
<thead>
<tr>
<th>Method</th>
<th>( p )</th>
<th>( \xi )</th>
<th>( \eta )</th>
<th>( \lambda )</th>
<th>Log-likelihood</th>
<th>( p )-value ( \chi^2 ) g-o-f</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>1</td>
<td>4.7</td>
<td>1.1</td>
<td>-1.8</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>ML</td>
<td>1</td>
<td>4.7</td>
<td>1.2</td>
<td>-2.2</td>
<td>-2202.1</td>
<td>0</td>
</tr>
<tr>
<td>ML</td>
<td>0.9</td>
<td>4.6</td>
<td>0.9</td>
<td>-2.1</td>
<td>-2128.0</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Test for symmetry \( p \)-value \( \approx 0 \)

Estimated mean direction: 3.8 radians (218 degrees)
95% profile likelihood based CIs for \( p \) and \( \lambda \): (0.876, 0.925) and (-2.63, -1.56)
14.2 Graphical Representation of Fits

Histogram of bird-flight headings together with fitted densities: ML WSNC+U solution (solid red); ML WSNC solution (broken red); MM WSNC solution (broken blue).
References


