Chapter 231

Circular Data Correlation

Introduction

This procedure computes summary statistics, generates rose plots and circular histograms, and computes the circular correlation coefficient for circular data.

Angular data, recorded in degrees or radians, is generated in a wide variety of scientific research areas. Examples of angular (and cyclical) data include daily wind directions, ocean current directions, departure directions of animals, direction of bone-fracture plane, and orientation of bees in a beehive after stimuli.

The usual summary statistics, such as the sample mean and standard deviation, cannot be used with angular values. For example, consider the average of the angular values 1 and 359. The simple average is 180. But with a little thought, we would conclude that 0 is a better answer. Because of this and other problems, a special set of techniques have been developed for analyzing angular data.
Suppose a sample of $n$ angles $a_1, a_2, \ldots, a_n$ is to be summarized. It is assumed that these angles are in degrees. Fisher (1993) and Mardia & Jupp (2000) contain definitions of various summary statistics that are used for angular data. These results will be presented next. Let

$$C_p = \sum_{i=1}^{n} \cos(pa_i), \quad \bar{C}_p = \frac{C_p}{n}, \quad S_p = \sum_{i=1}^{n} \sin(pa_i), \quad \bar{S}_p = \frac{S_p}{n},$$

$$R_p = \sqrt{\bar{C}_p^2 + \bar{S}_p^2}, \quad \bar{R}_p = \frac{R_p}{n}$$

$$T_p = \begin{cases} \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) & \bar{C}_p > 0, \bar{S}_p > 0 \\ \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) + \pi & \bar{C}_p < 0 \\ \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) + 2\pi & \bar{S}_p < 0, \bar{C}_p > 0 \end{cases}$$

To interpret these quantities, it may be useful to imagine that each angle represents a vector of length one in the direction of the angle. Suppose these individual vectors are arranged so that the beginning of the first vector is at the origin, the beginning of the second vector is at the end of the first, the beginning of the third vector is at the end of the second, and so on. We can then imagine a single vector $\vec{a}$ that will stretch from the origin to the end of the last observation.

$R_1$, called the resultant length, is the length of $\vec{a}$. $R_1$ is the mean resultant length of $\vec{a}$. Note that $R_1$ varies between zero and one and that a value of $R_1$ near one implies that there was little variation in values of the angles.

The mean direction, $\theta$, is a measure of the mean of the individual angles. $\theta$ is estimated by $T_1$.

The circular variance, $V$, measures the variation in the angles about the mean direction. $V$ varies from zero to one. The formula for $V$ is

$$V = 1 - \bar{R}_1$$

The circular standard deviation, $v$, is defined as

$$v = \sqrt{-2 \ln(\bar{R}_1)}$$

The circular dispersion, used in the calculation of confidence intervals, is defined as

$$\delta = \frac{1 - T_2}{2\bar{R}_1^2}$$
The skewness is defined as
\[ s = \frac{\overline{R_2} \sin(T_2 - 2T_1)}{(1 - \overline{R_1})^{3/2}} \]

The kurtosis is defined as
\[ k = \frac{\overline{R_2} \cos(T_2 - 2T_1) - \overline{R_1}^4}{(1 - \overline{R_1})^2} \]

**Correction for Grouped Data**

When the angles are grouped, a multiplicative correction for \( R \) may be necessary. The corrected value is given by
\[ \overline{R_p}^* = g \overline{R_p} \]

where
\[ g = \frac{\pi / J}{\sin(\pi / J)} \]

Here \( J \) is the number of equi-sized arcs. Thus, for monthly data, \( J \) would be 12.

**Confidence Interval for the Mean Direction**

Upton & Fingleton (1989) page 220 give a confidence interval for the mean direction when no distributional assumption is made as
\[ T_1 \pm \sin^{-1}(z_{\alpha/2} \hat{\sigma}) \]

where
\[ \hat{\sigma} = \sqrt{\frac{n(1 - H)}{4\overline{R^2}}} \]

\[ H = \frac{1}{n} \left\{ \cos(2T_1) \sum_{i=1}^{n} \cos(2a_i) + \sin(2T_1) \sum_{i=1}^{n} \sin(2a_i) \right\} \]
Circular Uniform Distribution

*Uniformity* refers to the situation in which all values around the circle are equally likely. The probability distribution on a circle with this property is the *circular uniform distribution*, or simply, the uniform distribution. The probability density function is given by

\[ f(a) = \frac{1}{360} \]

The probability between any two points is given by

\[ \Pr(a_1 < a_2 | a_1 \leq a_2, a_2 \leq a_1 + 2\pi) = \frac{a_2 - a_1}{360} \]

Tests of Uniformity

*Uniformity* refers to the situation in which all values around the circle are equally likely. Occasionally, it is useful to perform a statistical test of whether a set of data do not follow the uniform distribution. Several tests of uniformity have been developed. Note that when any of the following tests are rejected, we can conclude that the data were not uniform. However, when the test is not rejected, we cannot conclude that the data follow the uniform distribution. Rather, we do not have enough evidence to reject the null hypothesis of uniformity.

**Rayleigh Test**

The Rayleigh test, discussed in Mardia & Jupp (2000) pages 94-95, is the score test and the likelihood ratio test for uniformity within the von Mises distribution family. The Rayleigh test statistic is \( 2nR^2 \). For large samples, the distribution of this statistic under uniformity is a chi-square with two degrees of freedom with an error of approximation of \( O(n^{-1}) \). A closer approximation to the chi-square with two degrees of freedom is achieved by the modified Rayleigh test. This test, which has an error of \( O(n^{-2}) \), is calculated as follows.

\[ S^* = \left(1 - \frac{1}{2n}\right)2nR^2 + \frac{nR^4}{2} \]

**Modified Kuiper's Test**

The modified Kuiper's test, Mardia & Jupp (2000) pages 99-103, was designed to test uniformity against any alternative. It measures the distance between the cumulative uniform distribution function and the empirical distribution function. It is accurate for samples as small as 8. The test statistic, \( V \), is calculated as follows

\[ V = V_n \left(\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}}\right) \]

where

\[ V_n = \max_{i=1 \to n} \left(\frac{a(i)}{360} - \frac{i}{n}\right) - \min_{i=1 \to n} \left(\frac{a(i)}{360} - \frac{i}{n}\right) + \frac{1}{n} \]
Published critical values of $V$ are

<table>
<thead>
<tr>
<th>$V$</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.537</td>
<td>0.150</td>
</tr>
<tr>
<td>1.620</td>
<td>0.100</td>
</tr>
<tr>
<td>1.747</td>
<td>0.050</td>
</tr>
<tr>
<td>1.862</td>
<td>0.025</td>
</tr>
<tr>
<td>2.001</td>
<td>0.010</td>
</tr>
</tbody>
</table>

This table was used to create an interpolation formula from which the alpha values are calculated.

**Watson Test**

The following uniformity test is outlined in Mardia & Jupp pages 103-105. The test is conducted by calculating $U^2$ and comparing it to a table of values. If the calculated value is greater than the critical value, the null hypothesis of uniformity is rejected. Note that the test is only valid for samples of at least eight angles.

The calculation of $U^2$ is as follows

$$U^2 = \sum_{i=1}^{n} \left[ \frac{u_{(i)} - i - \frac{1}{2}}{n} + \frac{1}{2} \right]^2 + \frac{1}{12n}$$

where

$$\bar{u} = \frac{\sum_{i=1}^{n} u_{(i)}}{n}, \quad u_{(i)} = \frac{a_{(i)}}{360}$$

$a_{(1)} \leq a_{(2)} \leq a_{(3)} \leq \cdots \leq a_{(n)}$ are the sorted angles. Note that maximum likelihood estimates of $\kappa$ and $\theta$ are used in the distribution function. Mardia & Jupp (2000) present a table of critical values that has been entered into NCSS. When a value of $U^2$ is calculated, the table is interpolated to determine its significance level.

Published critical values of $U^2$ are

<table>
<thead>
<tr>
<th>$U^2$</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.131</td>
<td>0.150</td>
</tr>
<tr>
<td>0.152</td>
<td>0.100</td>
</tr>
<tr>
<td>0.187</td>
<td>0.050</td>
</tr>
<tr>
<td>0.221</td>
<td>0.025</td>
</tr>
<tr>
<td>0.267</td>
<td>0.010</td>
</tr>
</tbody>
</table>
**Von Mises Distribution**

The *Von Mises distribution* takes the role in circular statistics that is held by the normal distribution in standard linear statistics. In fact, it is shaped like the normal distribution, except that its tails are truncated.

The probability density function is given by

\[ f(a; \theta, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(a - \theta)] \]

where \( I_p(x) \) (the modified Bessel function of the first kind and order \( p \)) is defined by

\[ I_p(x) = \sum_{r=0}^{\infty} \frac{1}{(r+p)!r!} \left( \frac{x}{2} \right)^{2r+p}, \quad p = 0, 1, 2, \ldots \]

In particular

\[ I_0(x) = \sum_{r=0}^{\infty} \frac{1}{(r!)^2} \left( \frac{x}{2} \right)^{2r} = \frac{1}{2\pi} \int_0^{2\pi} e^{x \cos(\theta)} d\theta \]

The parameter \( \theta \) is the *mean direction* and the parameter \( \kappa \) is the *concentration parameter*.

The distribution is unimodal. It is symmetric about \( A \). It appears as a normal distribution that is truncated at plus and minus 180 degrees. When \( \kappa \) is zero, the von Mises distribution reduces to the uniform distribution. As \( \kappa \) gets large, the von Mises distribution approaches the normal distribution.

**Point Estimation**

The maximum likelihood estimate of \( \theta \) is the sample mean direction. That is, \( \hat{\theta} = T_1 \).

The maximum likelihood of \( \kappa \) is the solution to

\[ A_1(\kappa) = \bar{R} \]

where

\[ A_1(x) = \frac{I_1(x)}{I_0(x)} \]

That is, the MLE of \( \kappa \) is given by

\[ \kappa^* = A_1^{-1}(\bar{R}) \]
This can be approximated by (see Fisher (1993) page 88 and Mardia & Jupp (2000) pages 85-86)

\[
\kappa^* = \begin{cases} 
2\bar{R} + \frac{5\bar{R}^5}{6} & \bar{R} < 0.53 \\
-0.4 + 1.39\bar{R} + \frac{0.43}{1-\bar{R}} & 0.53 \leq \bar{R} < 0.53 \\
\frac{1}{3\bar{R} - 4\bar{R}^2 + \bar{R}^3} & \bar{R} \geq 0.85
\end{cases}
\]

This estimate is very biased. This bias is corrected by using the following modified estimator.

\[
\hat{\kappa} = \begin{cases} 
\max\left(\kappa^* - \frac{2}{n\kappa^*}, 0\right) & \kappa^* < 2 \\
\frac{(n-1)^3\kappa^*}{n(n^2 + 1)} & \kappa^* \geq 2
\end{cases}
\]

Confidence Interval for Mean Direction assuming Von Mises

A general confidence interval for \(\theta\) was given above. When the data can be assumed to follow a von Mises distribution, a more appropriate interval is given by Mardia & Jupp (2000) page 124 and Upton & Fingleton (1989) page 269. This confidence interval is given by

\[
T_1 \pm \cos^{-1}\left(\sqrt{\frac{2n[2\bar{R}^2 - nz^2_{\alpha}]}{\bar{R}^2(4n - z^2_{\alpha})}}\right)
\]

if \(\bar{R} \leq 2/3\)

\[
T_1 \pm \cos^{-1}\left(\sqrt{\frac{n^2 - (n^2 - \bar{R}^2) \exp\left(\frac{z^2_{\alpha}}{n}\right)}{\bar{R}}}\right)
\]

if \(\bar{R} > 2/3\)

Confidence Interval for Concentration of Von Mises

An approximate confidence interval for \(\kappa\) when \(\hat{\kappa} > 2\) was given by Mardia & Jupp (2000) pages 126-127 as

\[
\left(\frac{1 + \sqrt{1 + 3b}}{4b}, \frac{1 + \sqrt{1 + 3d}}{4d}\right)
\]

where

\[
b = \frac{n(1 - \bar{R})}{\chi^2_{n-1,1-\alpha/2}}
\]

\[
d = \frac{n(1 - \bar{R})}{\chi^2_{n-1,\alpha/2}}
\]
Goodness of Fit Tests for the Von Mises Distribution

Stephens Modified Watson’s Test

The following goodness-of-fit test, published by Lockhart & Stephens (1985) as a modification of the Watson test for the circle, is outlined in Fisher (1993) page 84. The test is conducted by calculating $U^2$ and comparing it to a table of values. If the calculated value is greater than the critical value, the null hypothesis of Von Misesness is rejected. Note that the test is only valid for samples of at least 20 angles.

The calculation of $U^2$ is as follows

$$U^2 = \sum_{i=1}^{n} \left[ \hat{p}_{(i)} - \frac{2i - 1}{2n} \right]^2 - n \left( \bar{p} - \frac{1}{2} \right)^2 + \frac{1}{12n}$$

where

$$\bar{p} = \frac{\sum_{i=1}^{n} \hat{p}_{(i)}}{n}$$

$$\hat{p}_{(i)} = F_{\kappa}(a_{(i)} - T_1)$$

$a_{(1)} \leq a_{(2)} \leq a_{(3)} \leq \cdots \leq a_{(n)}$ are the sorted angles and $F_{\kappa}(a - \theta)$ is the cumulative distribution function of the von Mises distribution. Note that maximum likelihood estimates of $\kappa$ and $\theta$ are used in the distribution function. Lockhart & Stephens (1985) present a table of critical values that has been entered into NCSS.

When a value of $U^2$ is calculated, the table is interpolated to determine its significance level.

Cox Test

Mardia & Jupp (2000) pages 142-143 present a von Mises goodness-of-fit test that was originally given by Cox (1975).

The test statistic, $C$, is distributed as a chi-squared variable with two degrees of freedom under the null hypothesis that the data follow the von Mises distribution. It is calculated as follows.

$$C = \frac{s_c^2}{nv_c(\hat{\kappa})} + \frac{s_s^2}{nv_s(\hat{\kappa})}$$

where

$$s_c = \sum_{i=1}^{n} \cos 2(a_i - T_1) - n\alpha_2(\hat{\kappa})$$

$$s_s = \sum_{i=1}^{n} \sin 2(a_i - T_1)$$

$$v_c(x) = \frac{1 + \alpha_4}{2} - \alpha_2^2 - \frac{[\alpha_1/2 + \alpha_3/2 - \alpha_1\alpha_2]^2}{(1 + \alpha_2)/2 - \alpha_1^2}$$

$$v_s(x) = \frac{\alpha_1 - \alpha_4}{2} - \frac{(\alpha_1 - \alpha_3)^2}{1 - \alpha_2}$$
Circular Correlation Measure

This section discusses a measure of the correlation between two circular variables presented by Jammalamadaka and SenGupta (2001). Suppose a sample of \( n \) pairs of angles \((a_{11}, a_{21}), (a_{12}, a_{22}), \ldots, (a_{1n}, a_{2n})\) is available. The circular correlation coefficient is calculated as

\[
r_c = \frac{\sum_{k=1}^{n} \sin(a_{1k} - T_{1,1}) \sin(a_{2k} - T_{2,1})}{\sqrt{\sum_{k=1}^{n} \sin^2(a_{1k} - T_{1,1}) \sum_{k=1}^{n} \sin^2(a_{2k} - T_{2,1})}}
\]

Where \( T_{1,1} \) is the mean direction of the first circular variable and \( T_{2,1} \) is the mean direction of second.

The significance of this correlation coefficient can be test using the fact that \( z_r \) is approximately distributed as a standard normal, where

\[
z_r = r_c \sqrt{\frac{n \lambda_{20} \lambda_{02}}{\lambda_{22}}}
\]

and

\[
\lambda_{ij} = \frac{1}{n} \sum_{k=1}^{n} \sin^i(a_{1k} - T_{1,1}) \sin^j(a_{2k} - T_{2,1})
\]

Data Structure

The data consist of two or more variables. Each variable contains a set of angular values. An example of a dataset containing circular data is Circular3.S0. Missing values are entered as blanks (empty cells).
Rose Plot / Circular Histogram Window Options

This section describes the specific options available on the Rose Plot / Circular Histogram Format window, which is displayed when a Rose Plot / Circular Histogram Format button is clicked. Common options, such as axes, labels, legends, and titles are documented in the Graphics Components chapter.

Rose Plot Tab

Data Type

The data type of the plot is specified independently of the data type specified on the Variables tab of the Circular Data Analysis procedure.
Direction
This option indicates whether the orientation of the plot is in a 'Clockwise' or 'Counter-Clockwise' direction.

Reference Angle (Rotation)
This option lets you indicate the position of 0 degrees by entering an offset angle. On the default circle, 0 degrees is on the right (east), 90 degrees is at the top (north), 180 degrees is on the left (west), and 270 degrees is at the bottom (south). This option lets you add an 'offset' to each angle which moves the position of 0 degrees around the circle.

The offset must be between 0 and 360 degrees.
**Interior Objects**

The two choices for plot styles are Rose Plot and Circular Histogram.

**Group Display**

When the data is grouped data, this option determines whether the petals within a bin are side-by-side, stacked upon each other, or overlaid.

**Side-by-Side**

The bin width is divided equally by the number of groups and the petals are laid out sequentially in the bin. Although the petals are narrower, they still encompass the points of the group that within the boundaries of the whole bin.

**Stacked**

A single petal in each bin is divided by the number of groups. Rose plots with the group display set to Stacked may be misleading because the proportional area is larger for the outside groups.

**Overlaid**

Each petal for each group is overlaid in each bin. Some degree of transparency is recommended when using the Overlaid group display. It is also difficult to distinguish groups when there are more than 2 or 3 groups.

**Radius for Interior Objects**

This option specifies the distance to the outer edge of the bins and petals of the rose plot or circular histogram.

**Petal Width**

Specify the percent of the total width of each bin that is to be used for each petal.
Percent for Histogram Base
This is the percent of Radius for Interior Objects that is used for the base of the circular histogram.

Number of Bins
Specify the number of bins for the circle.

First Bin Start Angle
This permits the user to change the angle at which the binning begins. This is useful, for example when the Data Type is set to Compass, since this option can be used to center the bins on the directions. When Data Type is set to Compass, the recommended First Bin Start Angle is 45 for 4 bins, 22.5 for 8 bins, and 11.25 for 16 bins.

Number of Radial Axes
Select the number of axes that go from the center to the outer edge of the interior region. If this is set to the same number as the number of bins, these axes show the edges of the bins. The radial axes also begin at the First Bin Start Angle.
Alternate Bin Fill

Check this box to show a background fill for each bin. The fills alternate beginning with Fill 1. When the number of bins is odd, the adjacent first and last bins will both have Fill 1.

Data & Means Tab

Raw Data Symbols

Check this box to show the raw data symbols.

Number of Raw Data Symbol Bins

Specify the number of bins for the raw data points. To use no binning set this to 0 or All Uniques.
First Bin Start Angle
This permits the user to change the angle at which the binning begins. This is useful, for example when the Data Type is set to Compass, since this option can be used to center the bins on the directions.

When Data Type is set to Compass, the recommended First Bin Start Angle is 45 for 4 bins, 22.5 for 8 bins, and 11.25 for 16 bins.

Radius for Raw Data Symbols
This is the distance from the center at which the symbols are shown.

Width for Multiple Symbols at One Location
This specifies the width of the band that contains the symbols when there is more than one value at some locations.

Mean Symbols and Lines
Use these options to set up the visual representation of the circular means for each group.

References Tab

Direction References
The options in this section allow you to specify the tick marks and references going around the plot.

Magnitude References
The options in this section allow you to specify the tick marks and references going from the center to the outside of the plot.
Example 1 – Correlation of Circular Data

This section presents an example of how to run this procedure. The data are wind directions at three locations. The data are found in the Circular3 dataset.

Setup

To run this example, complete the following steps:

1. **Open the Circular3 example dataset**
   - From the File menu of the NCSS Data window, select Open Example Data.
   - Select Circular3 and click OK.

2. **Specify the Circular Data Correlation procedure options**
   - Find and open the Circular Data Correlation procedure using the menus or the Procedure Navigator.
   - The settings for this example are listed below and are stored in the Example 1 settings file. To load these settings to the procedure window, click Open Example Settings File in the Help Center or File menu.

   Variables Tab
   
   Data Variables ................................................ Wind1, Wind2, Wind3

3. **Run the procedure**
   - Click the Run button to perform the calculations and generate the output.

Summary Statistics Section

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Mean Direction (Theta)</th>
<th>Mean Resultant Length (R bar)</th>
<th>Circular Variance (V)</th>
<th>Circular Standard Deviation (v)</th>
<th>Circular Dispersion (Delta)</th>
<th>Von Mises Concentration (Kappa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>39.6352</td>
<td>0.9530</td>
<td>0.0470</td>
<td>17.7742</td>
<td>0.0945</td>
<td>10.9134</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>41.6884</td>
<td>0.9514</td>
<td>0.0486</td>
<td>18.0785</td>
<td>0.0976</td>
<td>10.5673</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>42.4911</td>
<td>0.9541</td>
<td>0.0459</td>
<td>17.5658</td>
<td>0.0924</td>
<td>11.1608</td>
</tr>
</tbody>
</table>

Notes:
- Theta is the mean direction.
- R bar is a measure of data concentration. 0 <= R bar <= 1. R bar = 1 implies high concentration.
- V = 1 - R bar, so 0 <= V <= 1. V = 0 implies high concentration.
- v is the circular analog of the linear standard deviation.
- Delta is another measure of circular spread.
- Kappa is the concentration parameter of the Von Mises distribution.

Variable

This is the variable presented on this line.
Sample Size
This is the number of non-missing values in this variable.

Mean Direction
This is estimated mean direction, $T_1$.

Mean Resultant Length
This is the estimated mean resultant length, $R_1$. It is a measure of data concentration. An $R_1$ close to zero implies low data concentration. An $R_1$ close to one implies high data concentration.

Circular Variance
The circular variance, $V$, is a measure of variation in the data. Note that $V = 1 - R_1$.

Circular Standard Deviation
The circular standard deviation is $v = \sqrt{-2 \ln(R_1)}$. Note that it is not the square root of the circular variance.

Circular Dispersion
The circular dispersion, $\delta = \frac{1-T_2}{2R_1^2}$, is another measure of variation.

Von Mises Concentration
This is the estimated concentration parameter of the von Mises distribution, $\kappa$.

---

Circular Correlation Section

<table>
<thead>
<tr>
<th>First Variable</th>
<th>Second Variable</th>
<th>Correlation Coefficient</th>
<th>Z Value</th>
<th>Prob Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>Wind2</td>
<td>0.9956</td>
<td>2.9917</td>
<td>0.0028</td>
</tr>
<tr>
<td>Wind1</td>
<td>Wind3</td>
<td>0.9961</td>
<td>2.9834</td>
<td>0.0029</td>
</tr>
<tr>
<td>Wind2</td>
<td>Wind3</td>
<td>0.9901</td>
<td>2.9823</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

Notes:
This report provides the correlation between two circular variables.
The Z Value tests whether the correlation is zero.
The significance level is accurate for large samples.

This report provides the angular correlation coefficient of each pair of variables as defined in Jammalamadaka and SenGupta (2001). It also provides the results of a large sample significance test of whether the correlation is zero.
Mean Direction Section

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Mean Direction (Theta)</th>
<th>Lower 95.0% Confidence Limit of Theta</th>
<th>Upper 95.0% Confidence Limit of Theta</th>
<th>Standard Error of Mean Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>39.6352</td>
<td>33.3160</td>
<td>45.9543</td>
<td>3.2176</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>41.6884</td>
<td>35.2657</td>
<td>48.1110</td>
<td>3.2701</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>42.4911</td>
<td>36.2446</td>
<td>48.7375</td>
<td>3.1807</td>
</tr>
</tbody>
</table>

Notes: This large-sample confidence interval does not require a Von Mises distribution assumption.

This report provides the large sample confidence interval for the mean direction as described by Upton & Fingleton (1989) page 220. Note that this interval does not require the assumption that the data come from the von Mises distribution.

Variation Statistics Section

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Circular Variance (V)</th>
<th>Circular Standard Deviation (v)</th>
<th>Circular Dispersion (Delta)</th>
<th>Skewness (s)</th>
<th>Kurtosis (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>0.0470</td>
<td>17.7742</td>
<td>0.0945</td>
<td>-1.4417</td>
<td>1.4594</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>0.0486</td>
<td>18.0785</td>
<td>0.0976</td>
<td>-1.4464</td>
<td>1.5370</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>0.0459</td>
<td>17.5658</td>
<td>0.0924</td>
<td>-1.5929</td>
<td>1.4476</td>
</tr>
</tbody>
</table>

Notes: These statistics are designed to assess and compare the variation in the data. 
V = 1 - R bar, so 0 <= V <= 1. V = 0 implies high concentration. 
v is the circular analog of the linear standard deviation. 
Delta is another measure of circular spread. 
For symmetric, unimodal data sets, the skewness is close to zero. 
For Von Mises data sets, the kurtosis is close to zero.

This report provides measures of data variation and dispersion which were defined in the Statistical Summary Report. It also provides measures of the skewness and kurtosis of the data.

Skewness
This is a measure of the skewness (lack of symmetry about the mean) in the data. Symmetric, unimodal datasets have a skewness value near zero.

Kurtosis
This is a measure of the kurtosis (peakedness) in the data. Von Mises datasets have a kurtosis near zero.
NCSS Statistical Software

Circular Data Correlation

Von Mises Distribution Estimation Section

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Mean Direction (Theta)</th>
<th>Lower 95.0% Confidence Limit of Theta</th>
<th>Upper 95.0% Confidence Limit of Theta</th>
<th>Von Mises Conc. (Kappa)</th>
<th>Lower 95.0% Confidence Limit of Kappa</th>
<th>Upper 95.0% Confidence Limit of Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind2</td>
<td>30</td>
<td>41.6884</td>
<td>35.9839</td>
<td>47.3928</td>
<td>10.5673</td>
<td>4.1062</td>
<td>9.2127</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>42.4911</td>
<td>36.9568</td>
<td>48.0253</td>
<td>11.1608</td>
<td>4.2664</td>
<td>9.6664</td>
</tr>
</tbody>
</table>

Notes:
These large-sample confidence intervals assume that the data follow a Von Mises assumption.
The confidence interval for kappa requires the estimated kappa to be > 2.

This report provides estimates and confidence intervals of the parameters (mean direction and concentration) of the von Mises distribution that best fits the data. Note that the von Mises distribution is a symmetric, unimodal distribution. You should check the rose plot or circular histogram to determine if the data are symmetric.

The formulas used in the estimation and confidence intervals were given earlier in this chapter. They come from Mardia & Jupp (2000).

Trigonometric Moments Section

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean Cos(a)</th>
<th>Mean Sin(a)</th>
<th>Mean Cos(2a)</th>
<th>Mean Sin(2a)</th>
<th>R bar</th>
<th>2R bar</th>
<th>Theta</th>
<th>2Theta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>0.7339</td>
<td>0.6079</td>
<td>0.1686</td>
<td>0.8109</td>
<td>0.9530</td>
<td>0.8283</td>
<td>39.6352</td>
<td>78.2548</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>0.7105</td>
<td>0.6328</td>
<td>0.1103</td>
<td>0.8158</td>
<td>0.9514</td>
<td>0.8232</td>
<td>41.6884</td>
<td>82.2994</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>0.7035</td>
<td>0.6445</td>
<td>0.0884</td>
<td>0.8271</td>
<td>0.9541</td>
<td>0.8318</td>
<td>42.4911</td>
<td>83.9028</td>
</tr>
</tbody>
</table>

Notes:
These values are used in the calculation of other statistics.

This report provides summary statistics that are used in other calculations.

Mean Cos(a)
This is \( \bar{C}_1 = \frac{1}{n} \sum_{i=1}^{n} \cos(a_i) \).

Mean Sin(a)
This is \( \bar{S}_1 = \frac{1}{n} \sum_{i=1}^{n} \sin(a_i) \).

Mean Cos(2a)
This is \( \bar{C}_2 = \frac{1}{n} \sum_{i=1}^{n} \cos(2a_i) \).

Mean Sin(2a)
This is \( \bar{S}_2 = \frac{1}{n} \sum_{i=1}^{n} \sin(2a_i) \).
R bar
This is \( R_1 = \frac{1}{n} \sqrt{n(C_1^2 + S_1^2)} \).

2R bar
This is \( R_2 = \frac{1}{n} \sqrt{n(C_2^2 + S_2^2)} \).

Theta, 2 Theta
This is calculated using the following formula with \( p \) set to 1 and then 2, respectively.

\[
T_p = \begin{cases} 
  \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) & \bar{C}_p > 0, \bar{S}_p > 0 \\
  \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) + \pi & \bar{C}_p < 0 \\
  \tan^{-1} \left( \frac{\bar{S}_p}{\bar{C}_p} \right) + 2\pi & \bar{S}_p < 0, \bar{C}_p > 0 
\end{cases}
\]

Uniform Distribution Goodness-of-Fit Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Rayleigh's Test Statistic (S*)</th>
<th>Rayleigh's Test Prob Level</th>
<th>Kuiper's Test Statistic (V)</th>
<th>Kuiper's Test Prob Level</th>
<th>Watson's Test Statistic (U2)</th>
<th>Watson's Test Prob Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>65.9606</td>
<td>0.0000</td>
<td>4.3674</td>
<td>0.0000</td>
<td>1.8130</td>
<td>0.0000</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>65.7007</td>
<td>0.0000</td>
<td>4.3832</td>
<td>0.0000</td>
<td>1.7963</td>
<td>0.0000</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>66.1366</td>
<td>0.0000</td>
<td>4.4147</td>
<td>0.0000</td>
<td>1.8194</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Notes:
The tests in this report assess the goodness-of-fit of the uniform distribution.
The Rayleigh test requires samples of at least 20.
The Kuiper and Watson tests require samples of at least 8.

This section reports the results of three goodness-of-fit tests for the uniform distribution. They were documented earlier in this chapter.

These tests may be viewed as testing whether the data are distributed uniformly around the circle.
Von Mises Distribution Goodness-of-Fit Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size (N)</th>
<th>Watson's Test Statistic (U2)</th>
<th>Watson's Test Prob Level</th>
<th>Cox's Test Statistic (S)</th>
<th>Cox's Test Prob Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind1</td>
<td>30</td>
<td>0.0747</td>
<td>0.2011</td>
<td>0.5160</td>
<td>0.7726</td>
</tr>
<tr>
<td>Wind2</td>
<td>30</td>
<td>0.0538</td>
<td>0.4060</td>
<td>0.5735</td>
<td>0.7507</td>
</tr>
<tr>
<td>Wind3</td>
<td>30</td>
<td>0.0671</td>
<td>0.2597</td>
<td>0.5149</td>
<td>0.7730</td>
</tr>
</tbody>
</table>

Notes:
The tests in this report assess the goodness-of-fit of the Von Mises distribution. Both tests require samples of at least 20.

This section reports the results of two goodness-of-fit tests for the von Mises distribution. They were documented earlier in this chapter. Several hypothesis tests assume that the data follow a von Mises distribution. These tests allow you to check the accuracy of this assumption.

Rose Plots

Circular Plot

[Diagram of circular plot showing data distribution by variable]
Circular Plots By Variable

These plots show the distribution of the data around the circle.
Circular Histograms

Circular Plot

Circular Plots By Variable

© NCSS, LLC. All Rights Reserved.
The circular histograms are generated by setting the Interior Objects on Plot to ‘Circular Histogram’.