Chapter 230

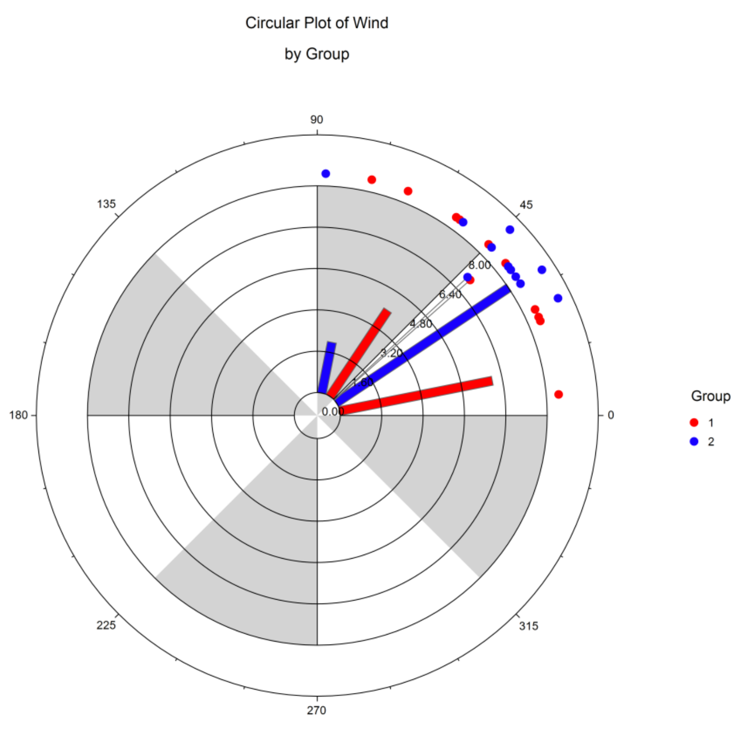
Circular Data Analysis

Introduction

This procedure computes summary statistics, generates rose plots and circular histograms, computes hypothesis tests appropriate for one, two, and several groups, 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 might 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. This procedure implements many of those techniques.



© NCSS, LLC. All Rights Reserved.

Technical Details

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

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

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

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

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 $\vec{a}$. $\bar{R}_1$ is the *mean resultant length* of $\vec{a}$. Note that $\bar{R}_1$ varies between zero and one and that a value of $\bar{R}_1$ near one implies that there was little variation in values of the angles.

The *mean direction*, θ, is a measure of the mean of the individual angles. θ 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

$$s = \frac{\bar{R}_2 \sin(T_2 - 2T_1)}{(1 - \bar{R}_1)^{3/2}}$$

© NCSS, LLC. All Rights Reserved.

The *kurtosis* is defined as

$$k = \frac{\bar{R}_2 \cos(T_2 - 2T_1) - \bar{R}_1^4}{(1 - \bar{R}_1)^2}$$

Correction for Grouped Data

When the angles are grouped, a multiplicative correction for *R* may be necessary. The corrected value is given by

$$\bar{R}_p^* = g\bar{R}_p$$

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

$$T_1 \pm \sin^{-1}\left(z_{\alpha/2}\hat{\sigma}\right)$$

where

$$\hat{\sigma} = \sqrt{\frac{n(1 - H)}{4R^2}}$$

$$H = \frac{1}{n}\left\{\cos(2T_1)\sum_{i=1}^{n}\cos(2a_i) + \sin(2T_1)\sum_{i=1}^{n}\sin(2a_i)\right\}$$

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

$$\Pr(a_1 < a_2 | a_1 \le a_2, a_2 \le a_1 + 2\pi) = \frac{a_2 - a_1}{360}$$

© NCSS, LLC. All Rights Reserved.

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 $2n\bar{R}^2$. 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.

$$S^* = \left(1 - \frac{1}{2n}\right) 2n\bar{R}^2 + \frac{n\bar{R}^4}{2}$$

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

V	**Alpha**
1.537	0.150
1.620	0.100
1.747	0.050
1.862	0.025
2.001	0.010

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

© NCSS, LLC. All Rights Reserved.

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

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

where

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

$a_{(1)} \le a_{(2)} \le a_{(3)} \le \cdots \le a_{(n)}$ are the sorted angles. Note that maximum likelihood estimates of κ and θ 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

U^2	Alpha
0.131	0.150
0.152	0.100
0.187	0.050
0.221	0.025
0.267	0.010

Von Mises Distributions

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

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

© NCSS, LLC. All Rights Reserved.

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 θ is the *mean direction* and the parameter κ 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 κ is zero, the von Mises distribution reduces to the uniform distribution. As κ gets large, the von Mises distribution approaches the normal distribution.

Point Estimation

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

The maximum likelihood of κ 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 κ 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)

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

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

$$\hat{\kappa} = \begin{cases} \begin{cases} \max\left(\kappa^* - \dfrac{2}{n\kappa^*}, 0\right) & \kappa^* < 2 \\ \dfrac{(n-1)^3\kappa^*}{n(n^2+1)} & \kappa^* \geq 2 \end{cases} & n \leq 15 \\ \kappa^* & n > 15 \end{cases}$$

© NCSS, LLC. All Rights Reserved.

Test for a Specified Mean Direction of Von Mises Data

There are several different hypothesis tests that have been proposed for testing $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, where θ_0 is a specific value of the mean direction. The tests presented here require the additional assumption that the data follow the Von Mises distribution, at least approximately.

It will be useful to adopt the following notation.

$$\bar{C}^* = \frac{1}{n}\sum_{i=1}^{n} \cos(a_i - \theta_0)$$

$$\bar{S}^* = \frac{1}{n}\sum_{i=1}^{n} \sin(a_i - \theta_0)$$

$$\bar{R}^* = \sqrt{[\bar{S}^*]^2 + [\bar{C}^*]^2}$$

Score Test

The score test, given by Mardia & Jupp (2000) page 123, is computed as

$$\chi_S^2 = \frac{n\hat{\kappa}}{A_1(\hat{\kappa})}(\bar{S}^*)^2$$

For large *n*, χ_S^2 follows the chi-square distribution with one degree of freedom.

Likelihood Ratio Test

The likelihood ratio test, given by Mardia & Jupp (2000) page 122, is computed as

$$\chi_L^2 = \begin{cases} \dfrac{4n[(\bar{R}^*)^2 - (\bar{C}^*)^2]}{2 - (\bar{C}^*)^2} & \text{if } n \geq 5 \text{ and } \bar{C}^* \leq 2/3 \\ \dfrac{2n^3}{n^2 + (n\bar{C}^*)^2 + 3n} \log\left(\dfrac{1 - (\bar{C}^*)^2}{1 - (\bar{R}^*)^2}\right) & \text{if } n \geq 5 \text{ and } \bar{C}^* > 2/3 \end{cases}$$

The test statistic, χ_L^2, follows a chi-square distribution with one degree of freedom.

Watson & Williams Test

The Watson and Williams test, given by Mardia & Jupp (2000) page 123, is computed as

$$F = \frac{\bar{R}^* - \bar{C}^*}{(1 - \bar{R}^*)/(n-1)} \; if \; \bar{C}^* \geq 5/6$$

The test statistic, *F*, follows an *F* distribution with one and *n*-1 degrees of freedom.

© NCSS, LLC. All Rights Reserved.

Stephens Test

This test, given by Fisher (1993) pages 93-94, is computed as

$$E = \frac{\sin(T_1 - \theta_0)}{\sqrt{1 / (n\hat{\kappa}\bar{R})}}$$

If $\hat{\kappa} \geq 2$, E follows the standard normal distribution.

Confidence Interval for Mean Direction assuming Von Mises

A general confidence interval for θ 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

$$T_1 \pm \cos^{-1}\left(\sqrt{\frac{2n[2R^2 - nz_\alpha^2]}{R^2(4n - z_\alpha^2)}}\right) \qquad \text{if } \bar{R} \leq 2/3$$

$$T_1 \pm \cos^{-1}\left(\frac{\sqrt{n^2 - (n^2 - R^2)\exp\left(\frac{z_\alpha^2}{n}\right)}}{R}\right) \qquad \text{if } \bar{R} > 2/3$$

Test for a Specified Concentration of Von Mises Data

Suppose you want to test a one-sided hypothesis concerning κ, given that the data come from a Von Mises distribution and that the mean direction parameter is unknown. Fisher (1993) page 95 suggests the following procedure when $\hat{\kappa} \geq 2$.

When testing $\kappa = \kappa_0$ versus $\kappa < \kappa_0$, reject the null hypothesis if

$$\bar{R} < 1 - \frac{\chi^2_{n-1;\alpha}}{2n}\left(\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}\right)$$

When testing $\kappa = \kappa_0$ $\kappa = \kappa_0$ versus $\kappa > \kappa_0$ $\kappa > \kappa_0$, reject the null hypothesis if

$$\bar{R} > 1 - \frac{\chi^2_{n-1;1-\alpha}}{2n}\left(\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}\right)$$

These tests are based on the result that

$$\frac{2n(1 - \bar{R})}{\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}} \sim \chi^2_{n-1}$$

© NCSS, LLC. All Rights Reserved.

Confidence Interval for Concentration of Von Mises

An approximate confidence interval for κ 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

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

where

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

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

$a_{(1)} \le a_{(2)} \le a_{(3)} \le \cdots \le a_{(n)}$ 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 κ and θ 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.

© NCSS, LLC. All Rights Reserved.

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.

$$C = \frac{s_c^2}{n v_c(\hat{\kappa})} + \frac{s_s^2}{n v_s(\hat{\kappa})}$$

where

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

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

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

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

Multi-Group Tests

Three multi-group tests are available for testing hypotheses about two or more groups. The nonparametric uniform-scores test tests whether the distributions of the groups are identical. The Watson-Williams F test tests whether a set of mean directions are equal given that the concentrations are unknown, but equal, given that the groups each follow a von Mises distribution. The concentration homogeneity test tests whether the concentration parameters are equal, given that the groups each follow a von Mises distribution.

Mardia-Watson-Wheeler Uniform-Scores Test

Suppose you have *g* populations following any common distribution from which random samples are taken and you wish to test whether these distributions are equal. Fisher (1993) page 122 and Mardia & Jupp (2000) pages 156-157 present a nonparametric test that is calculated as follows

$$W_g = 2\sum_{i=1}^{g} \frac{\left(C_{Ri}^2 + S_{Ri}^2\right)}{n_i}$$

where $C_{Ri} = \sum_{j=1}^{n_i} \cos(\gamma_{ij})$, $S_{Ri} = \sum_{j=1}^{n_i} \sin(\gamma_{ij})$, $n = \sum_{i=1}^{g} n_i$, and γ_{ij} are the circular ranks of the corresponding angles. The circular ranks are calculated using

$$\gamma_{ij} = \frac{2\pi r_{ij}}{n}$$

© NCSS, LLC. All Rights Reserved.

where the r_{ij} are the ranks of the corresponding a_{ij}.

If all n_i are greater than 10, the distribution of W_g is approximately distributed as a chi-square with 2*g*-2 degrees of freedom.

Since ranks are used in this test, ties become an issue. We have adopted the strategy of applying average ranks. Note that little has been done to test the adoption of this strategy within the realm of circular statistics.

Watson-Williams High Concentration F Test

Suppose you have *g* Von Mises populations from which random samples are taken and you wish to test whether their mean directions are equal. That is, you wish to test the null hypothesis

$$H_0: \theta_1 = \theta_2 = \ldots = \theta_g$$

Mardia & Jupp (2000) pages 134-135 present the Watson-William High-Concentration F Test that is calculated as follows

$$F_{WW} = \left(1 + \frac{3}{8\hat{\kappa}}\right) \frac{\left(\sum_{j=1}^{g} R_j - R\right) / (g-1)}{\left(n - \sum_{j=1}^{g} R_j\right) / (n-g)}$$

where $\hat{\kappa}$ is the maximum likelihood estimate of the concentration based on *R* and

$$R_j = \sqrt{C_j^2 + S_j^2},\ C_j = \sum_{i=1}^{n_j} \cos(a_i),\ S_j = \sum_{i=1}^{n_j} \sin(a_i),\ R = \sqrt{C^2 + S^2},$$

$$C = \sum_{j=1}^{g} C_j,\ S = \sum_{j=1}^{g} S_j,\ \text{and } n = \sum_{j=1}^{g} n_j .$$

The distribution of F_{WW} is approximately distributed as an *F* with *g-1* and *n-1* degrees of freedom when the assumptions that $\kappa_1 = \kappa_2 = \ldots = \kappa_g$ and that the distributions are Von Mises are made. The approximation also requires that $\hat{\kappa} \geq 1$.

Multi-Group Concentration Homogeneity Test

Suppose you have *g* groups from which random samples are taken and you wish to test whether the concentrations are equal. That is, you wish to test the null hypothesis

$$H_0: \kappa_1 = \kappa_2 = \ldots = \kappa_g$$

Mardia & Jupp (2000) page 139 presents such a test. It is divided into three cases.

© NCSS, LLC. All Rights Reserved.

Case I. $\bar{R} < 0.45$

*U*1 is approximately distributed as a chi-square with *g-1* degrees of freedom

$$U1 = \sum_{j=1}^{g} w_j f_j^2 - \frac{\left\{\sum_{j=1}^{g} w_j f_j\right\}^2}{\sum_{j=1}^{g} w_j}$$

where $w_j = \frac{4(n_j - 4)}{3}$ and $f_j = \sin^{-1}\left(2\bar{R}_j\sqrt{3/8}\right)$

Case II. $0.45 \leq \bar{R} \leq 0.70$

*U*2 is approximately distributed as a chi-square with *g-1* degrees of freedom

$$U2 = \sum_{j=1}^{g} w_j h_j^2 - \frac{\left\{\sum_{j=1}^{g} w_j h_j\right\}^2}{\sum_{j=1}^{g} w_j}$$

where $w_j = \frac{n_j - 3}{0.797449}$ and $h_j = \sinh^{-1}\left(\frac{\bar{R}_j - 1.089}{0.258}\right)$

Case III. $\bar{R} > 0.70$

*U*3 is approximately distributed as a chi-square with *g-1* degrees of freedom

$$U3 = \frac{1}{1+d}\left\{v \log\left(\frac{n - \sum_{j=1}^{g} R_j}{v}\right) - \sum_{j=1}^{g} v_j \log\left(\frac{n_j - R_j}{v_j}\right)\right\}$$

where $v_j = n_j - 1$, $v = n - g$, and $d = \frac{1}{3(g-1)}\left\{\sum_{j=1}^{g} \frac{1}{v_j} - \frac{1}{v}\right\}$.

© NCSS, LLC. All Rights Reserved.

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 $(a_{11}, a_{21}), (a_{12}, a_{22}), \dots, (a_{1n}, a_{2n})$ is available. The circular correlation coefficient is calculated as

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

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 tested using the fact the z_r is approximately distributed as a standard normal, where

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

and

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

Data Structure

The data consist of one or more variables. Each variable contains a set of angular values. The rows may be separated into groups using the unique values of an optional grouping variable. An example of a dataset containing circular data is Circular1.S0. Missing values are entered as blanks (empty cells).

© NCSS, LLC. All Rights Reserved.

Rose Plot / Circular Histogram Window Options

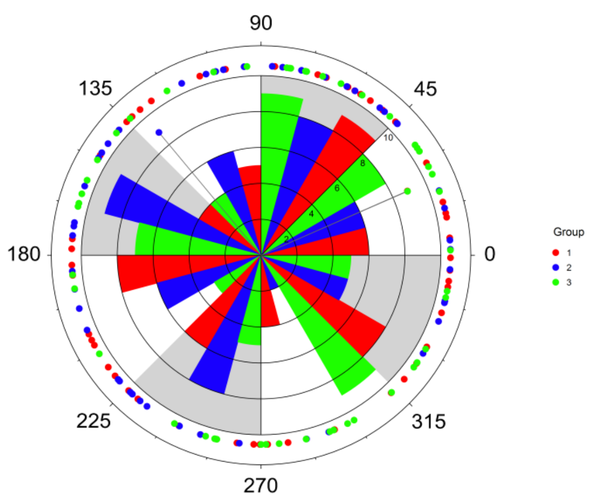
This section describes the specific options available on the Rose Plot / Circular Histogram Format window, which is displayed when a Rose Plot / Circular Histogram Format button is clicked. Common options, such as axes, labels, legends, and titles are documented in the Graphics Components chapter.

Rose Plot Tab

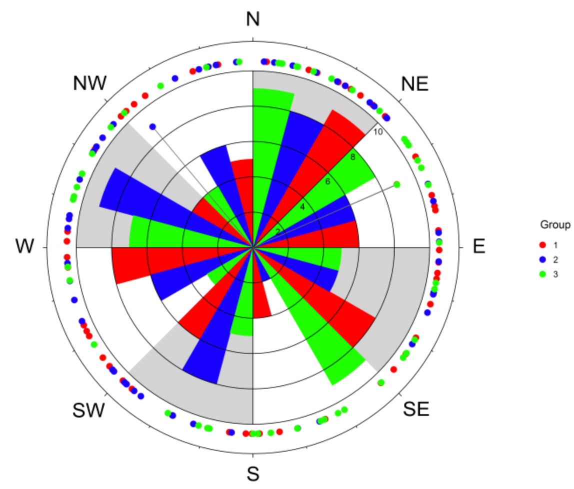
Data Type

The data type of the plot is specified independently of the data type specified on the Variables tab of the Circular Data Analysis procedure.

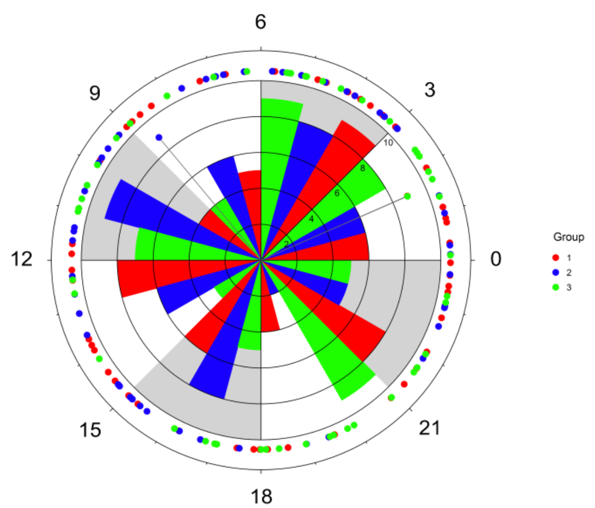
Data Type = Angle (0-360)



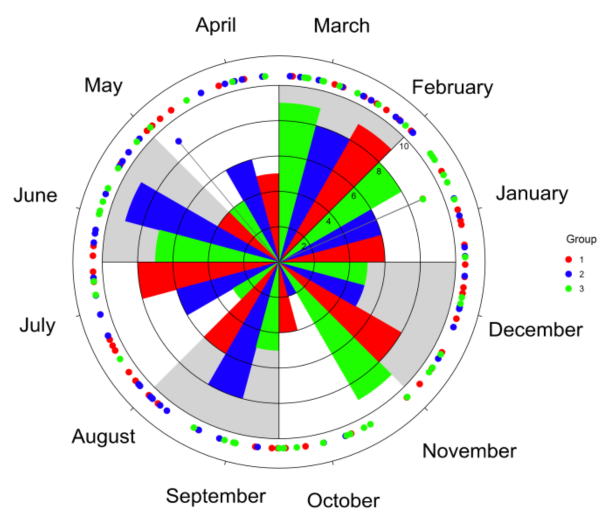
Data Type = Compass (N,E,S,W)



Data Type = Time (0-24)



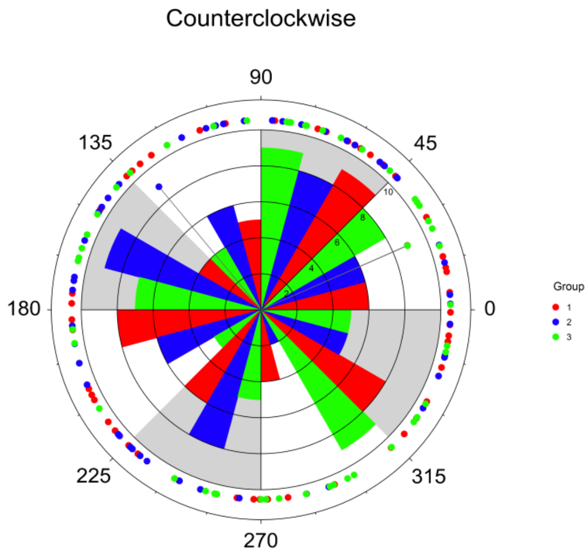
Data Type = Month of Year (1-12)

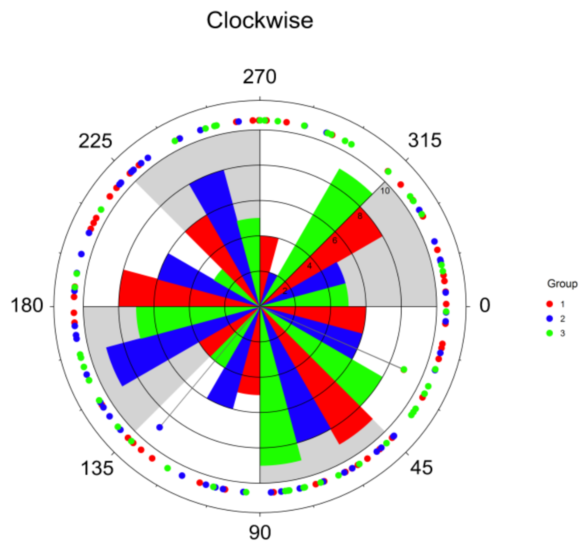


© NCSS, LLC. All Rights Reserved.

Direction

This option indicates whether the orientation of the plot is in a 'Clockwise' or 'Counter-Clockwise' direction.

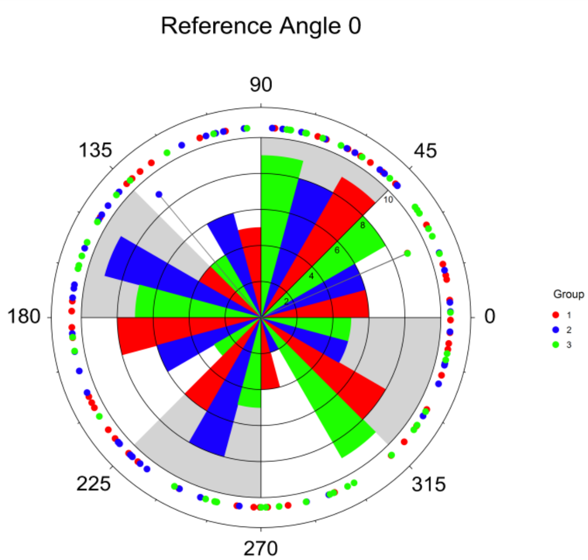


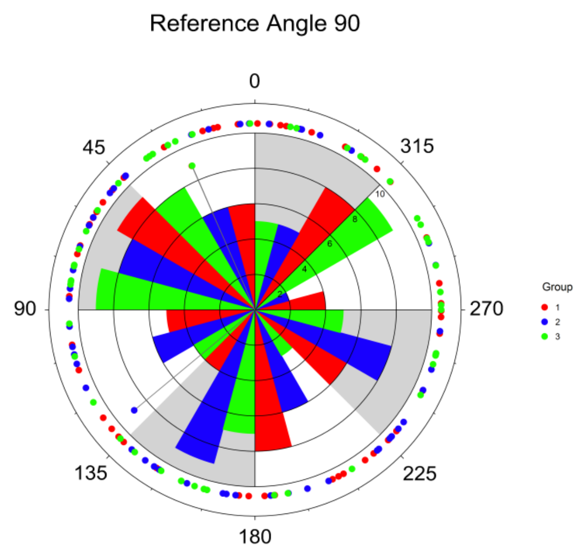


Reference Angle (Rotation)

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.

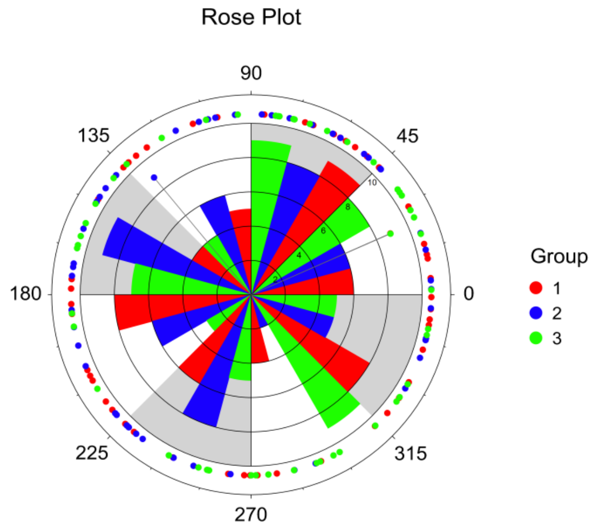


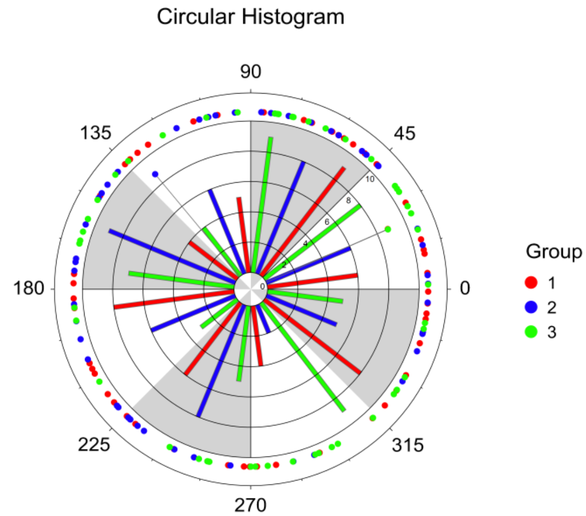


© NCSS, LLC. All Rights Reserved.

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.

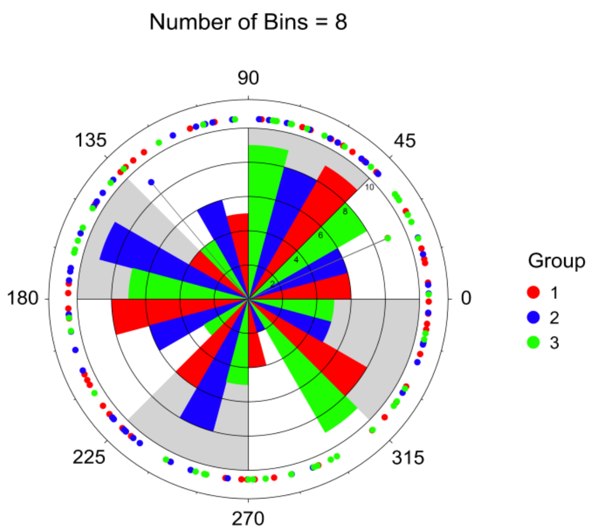
© NCSS, LLC. All Rights Reserved.

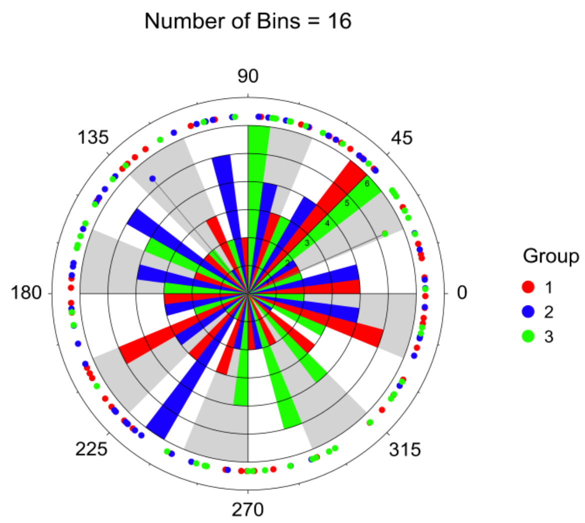
Percent for Histogram Base

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

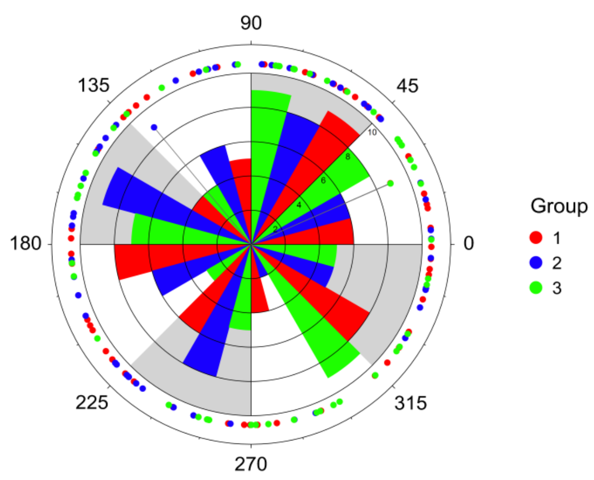
Select the number of axes that go from the center to the outer edge of the interior region. If this is set to the same number as the number of bins, these axes show the edges of the bins. The radial axes also begin at the First Bin Start Angle.

© NCSS, LLC. All Rights Reserved.

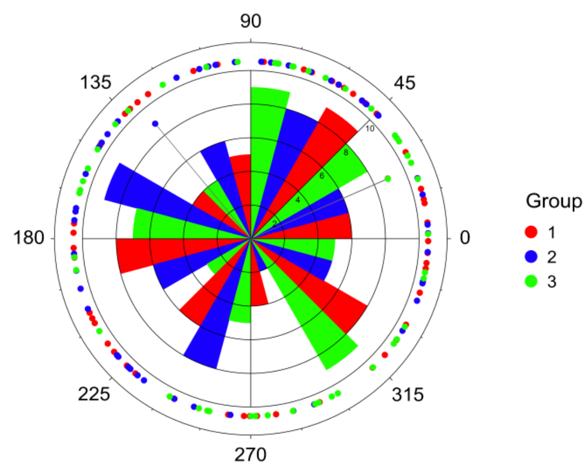
Alternate Bin Fill

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.

Alternate Bin Fill



Don't Alternate Bin Fill

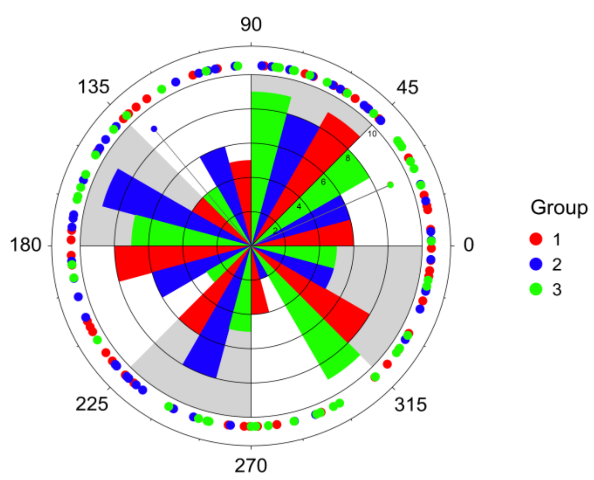


Data & Means Tab

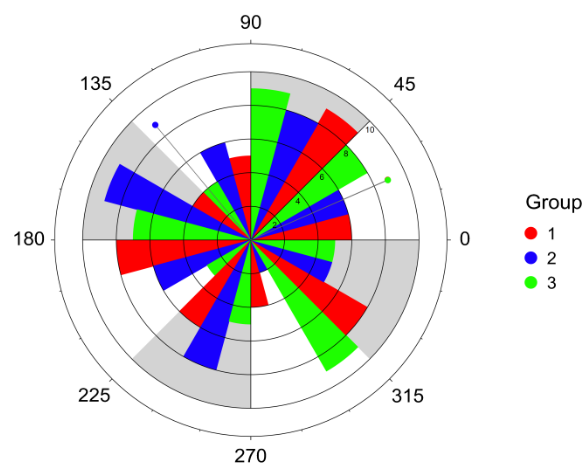
Raw Data Symbols

Check this box to show the raw data symbols.

With Raw Data Symbols



Without 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.

© NCSS, LLC. All Rights Reserved.

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

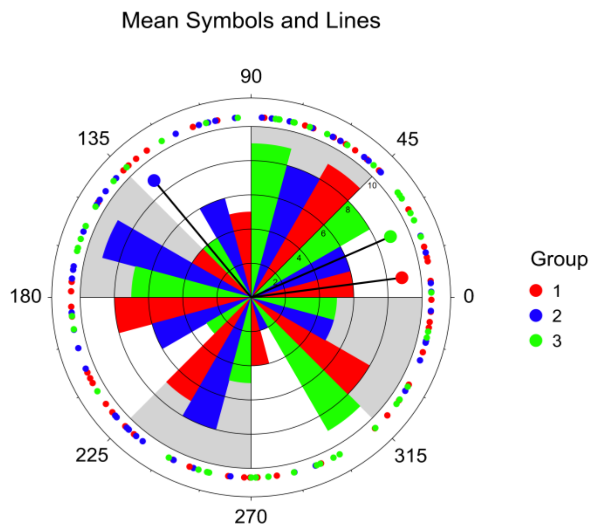
This is the distance from the center at which the symbols are shown.

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.

© NCSS, LLC. All Rights Reserved.

Example 1 – Analysis of Circular Data

This section presents an example of how to run this procedure. The data are wind directions of two groups. The data are found in the Circular1 dataset.

Setup

To run this example, complete the following steps:

1 Open the Circular1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Circular1** and click **OK**.

2 Specify the Circular Data Analysis procedure options

- Find and open the **Circular Data Analysis** 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 .... **Wind**
Grouping Variable.. **Group**
Hypothesized Theta....................................... **40**
Hypothesized Kappa...................................... **2**

3 Run the procedure

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

Summary Statistics

Summary Statistics for Wind

		Mean		Circular			
Group	Sample Size N	Direction θ	Resultant Length R bar	Variance V	Standard Deviation v	Dispersion δ	Von Mises Concentration κ
1	10	41.5869	0.9324	0.0676	21.4299	0.1449	5.5452
2	10	42.6725	0.9599	0.0401	16.3991	0.0768	9.1850

Notes:
θ 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.
δ is another measure of circular spread.
κ is the concentration parameter of the Von Mises distribution.

© NCSS, LLC. All Rights Reserved.

Group

This is the group (or variable) presented on this line.

Sample Size

This is the number of nonmissing values in this group.

Mean Direction (θ)

This is estimated mean direction, T_1.

Mean Resultant Length (R bar)

This is the estimated mean resultant length, $\bar{R}_1$. It is a measure of data concentration. An $\bar{R}_1$ close to zero implies low data concentration. An $\bar{R}_1$ close to one implies high data concentration.

Circular Variance (V)

The circular variance, *V*, is a measure of variation in the data. Note that $V = 1 - \bar{R}_1$.

Circular Standard Deviation (v)

The circular standard deviation is $v = \sqrt{-2\ln(\bar{R}_1)}$. Note that it is not the square root of the circular variance.

Circular Dispersion (δ)

The circular dispersion, $\delta = \frac{1-T_2}{2\bar{R}_1^2}$, is another measure of variation.

Von Mises Concentration (κ)

This is the estimated concentration parameter of the von Mises distribution, κ.

Mean Direction

Mean Direction for Wind

Group	Sample Size N	Mean Direction θ	Standard Error	95% Confidence Interval Limits for θ Lower	Upper
1	10	41.5869	6.8964	27.9417	55.2321
2	10	42.6725	5.0365	32.7516	52.5934

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.

© NCSS, LLC. All Rights Reserved.

Variation Statistics

Variation Statistics for Wind

Group	Sample Size N	Circular Variance V	Circular Standard Deviation v	Circular Dispersion δ	Skewness s	Kurtosis k
1	10	0.0676	21.4299	0.1449	-0.0795	-1.7244
2	10	0.0401	16.3991	0.0768	-4.8582	5.4248

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.
δ 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 Summary Statistics 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.

Von Mises Distribution Estimation

Von Mises Distribution Estimation for Wind

Group	Sample Size N	Mean Direction θ	95% Confidence Interval Limits for θ Lower	95% Confidence Interval Limits for θ Upper	Von Mises Concentration κ	95% Confidence Interval Limits for κ Lower	95% Confidence Interval Limits for κ Upper
1	10	41.5869	29.1135	54.0603	5.5452	2.3222	8.4051
2	10	42.6725	33.2995	52.0455	9.1850	3.0227	13.2191

Notes:
These large-sample confidence intervals assume that the data follow a von Mises assumption.
The confidence interval for κ requires that the estimated κ is at least 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).

© NCSS, LLC. All Rights Reserved.

Trigonometric Moments

Trigonometric Moments for Wind

	Sample Size	Mean				R bar		T	
Group	N	Cos(a)	Sin(a)	Cos(2a)	Sin(2a)	1	2	1	2
1	10	0.6974	0.6189	0.0903	0.7426	0.9324	0.7481	41.5869	83.0670
2	10	0.7057	0.6506	0.1085	0.8516	0.9599	0.8585	42.6725	82.7373

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 1

This is $\bar{R}_1 = \frac{1}{n}\sqrt{n(\bar{C}_1^2 + \bar{S}_1^2)}$.

R bar 2

This is $\bar{R}_2 = \frac{1}{n}\sqrt{n(\bar{C}_2^2 + \bar{S}_2^2)}$.

T1 and T2

This is calculated using the following formula with *p* set to 1 and then 2, respectively.

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

© NCSS, LLC. All Rights Reserved.

Multiple-Group Hypothesis Tests

Multiple-Group Hypothesis Tests for Wind

Null Hypothesis (H0)	Test Name	Test Statistic	P-Value	Reject H0 at α = 0.05?
Equal Distributions	Uniform Scores Test	6.7392	0.0344	Yes
Equal Directions	Watson-Williams F-Test	0.0147	0.9047	No
Equal Concentrations	Concentration Homogeneity Test	0.5717	0.4496	No

Notes:
These statistics test various hypotheses about the parameters of von Mises distributions.
They require that each group follow the von Mises distribution.
The Uniform Scores Test requires a sample size of at least 10.
The Watson and Williams F-Test assumes that all κ's are equal and that their average is greater than 1.

This report provides tests for three hypotheses about the features of several von Mises datasets. That is, it provides a test of whether the distributions are identical, whether the mean directions are identical, and whether the concentrations are identical. These tests are documented in the Technical Details section of this chapter.

Two-Group Hypothesis Tests

Two-Group Hypothesis Tests for Wind

Group		H0: Equal Distributions		H0: Equal Directions		H0: Equal Concentrations	
First	Second	Test Statistic	P-Value	Test Statistic	P-Value	Test Statistic	P-Value
1	2	6.7392	0.0344	0.0147	0.9047	0.5717	0.4496

Notes:
These statistics test various hypotheses about the parameters of von Mises distributions.
They require that each group follow the von Mises distribution.
Equal distributions is tested by the Mardia-Watson-Wheeler uniform scores test. Requires all Ni > 10.
Equal directions is tested by the Watson-Williams F-Test. Assumes Von Mises data with equal κ's, all > 1.
Equal concentrations is tested by the Concentration Homogeneity Test. Assumes Von Mises data.

This report provides the same three tests as the Multiple-Group Hypothesis Tests Section, taken two groups at a time. It allows you to pinpoint where differences occur.

© NCSS, LLC. All Rights Reserved.

Test Statistics and P-Values of Tests for a Specified Mean Direction Assuming Von Mises Data

Test Statistics of Tests for a Specified Mean Direction Assuming Von Mises Data for Wind

	Sample	Mean Direction		Test Statistic			
	Size	Actual	Null	Score Test	Likelihood Ratio Test	Watson and Williams Test	Stephens Test
Group	N	θ	θ0	Z	Chi-Square	F	Z
1	10	41.5869	40	0.0409	0.0470	0.0476	0.0397
2	10	42.6725	40	0.1949	0.2266	0.2341	0.1917

Notes:
These procedures test whether the mean direction is equal to a specified value, when κ (concentration) is unknown.
They assume that the data follow the von Mises distribution.
The Score Test requires a large sample size.
The Likelihood Ratio Test requires a sample size of at least 5.
The Watson and Williams Test requires a large value of κ.
The Stephens Test requires that κ is at least 2.

P-Values of Tests for a Specified Mean Direction Assuming Von Mises Data for Wind

	Sample	Mean Direction		P-Value			
	Size	Actual	Null	Score	Likelihood	Watson and	Stephens
Group	N	θ	θ0	Test	Ratio Test	Williams Test	Test
1	10	41.5869	40	0.8398	0.8284	0.8321	0.8422
2	10	42.6725	40	0.6589	0.6340	0.6400	0.6615

Notes:
This report gives the p-values of the test statistics displayed in the previous report.
Although the p-values of four tests are given, you should use only one of them chosen in advance.

This section reports the results of four tests of the hypothesis that the mean direction of a particular group is equal to a specific value. These are two-sided tests. They were documented earlier in this chapter.

The first table gives the values of the test statistics. The second table gives the probability levels. The null hypothesis is rejected when the probability level is less than 0.05 (or some other appropriate cutoff).

Tests for a Specified Concentration Assuming Von Mises Data

Tests for a Specified Concentration Assuming Von Mises Data for Wind

	Sample	Concentration			P-Value	
	Size	Actual	Null			
Group	N	κ	κ0	Chi-Square	(H1: κ < κ0)	(H1: κ > κ0)
1	10	5.5452	2	2.2756	0.0137	0.9863
2	10	9.1850	2	1.3518	0.0019	0.9981

Notes:
These statistics test whether the κ (concentration) parameter is equal to the specified value (κ0).
The tests require that the estimated κ is at least 2.

© NCSS, LLC. All Rights Reserved.

This section reports the results of two, one-sided tests of the hypothesis that the concentration parameter of each group is equal to a specific value. They were documented earlier in this chapter.

The first probability level is for testing the null hypothesis that kappa is greater than or equal to κ0. The second probability level is for test the null hypothesis that kappa is less than or equal to κ0.

Uniform Distribution Goodness-of-Fit Tests

Uniform Distribution Goodness-of-Fit Tests for Wind

	Sample Size	Rayleigh's Test		Kuiper's Test		Watson's Test	
Group	N	S*	P-Value	V	P-Value	U²	P-Value
1	10	20.2993	0.0000	2.7145	0.0000	0.5657	0.0001
2	10	21.7499	0.0000	2.8088	0.0000	0.6788	0.0000

Notes:
The tests in this report assess the goodness-of-fit of the uniform distribution.
Rayleigh's Test requires a sample size of at least 20.
Kuiper's Test and Watson's Test require a sample size 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

Von Mises Distribution Goodness-of-Fit Tests for Wind

	Sample Size	Watson's Test		Cox's Test	
Group	N	U²	P-Value	S	P-Value
1	10	0.0340	0.5000	0.4030	0.8175
2	10	0.1282	0.0322	2.9309	0.2310

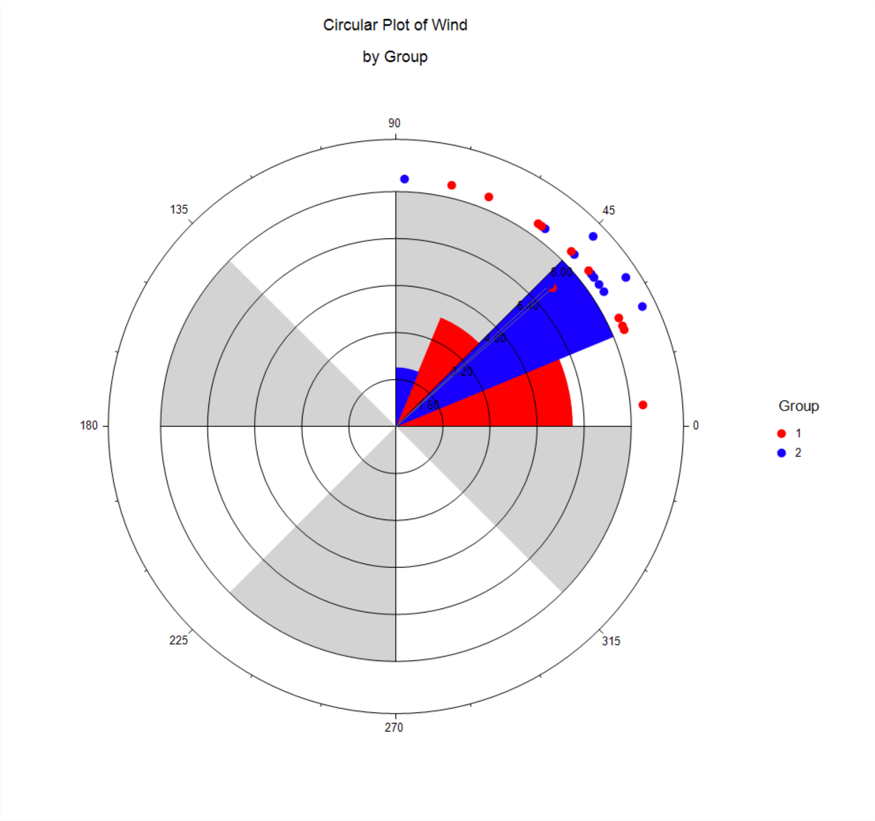
Notes:
The tests in this report assess the goodness-of-fit of the von Mises distribution.
Both tests require a sample size of at least 20.

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.

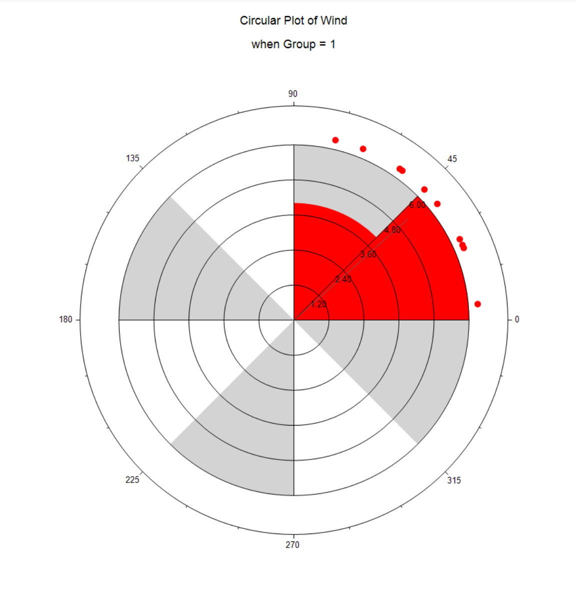
© NCSS, LLC. All Rights Reserved.

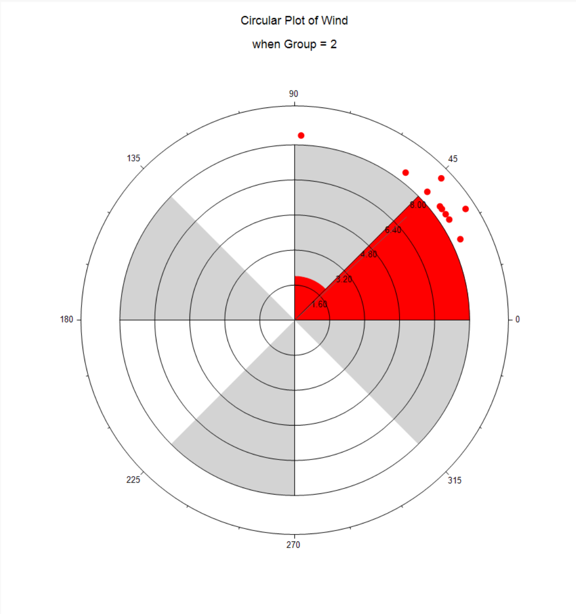
Rose Plots

Combined Rose Plot



Individual Rose Plots



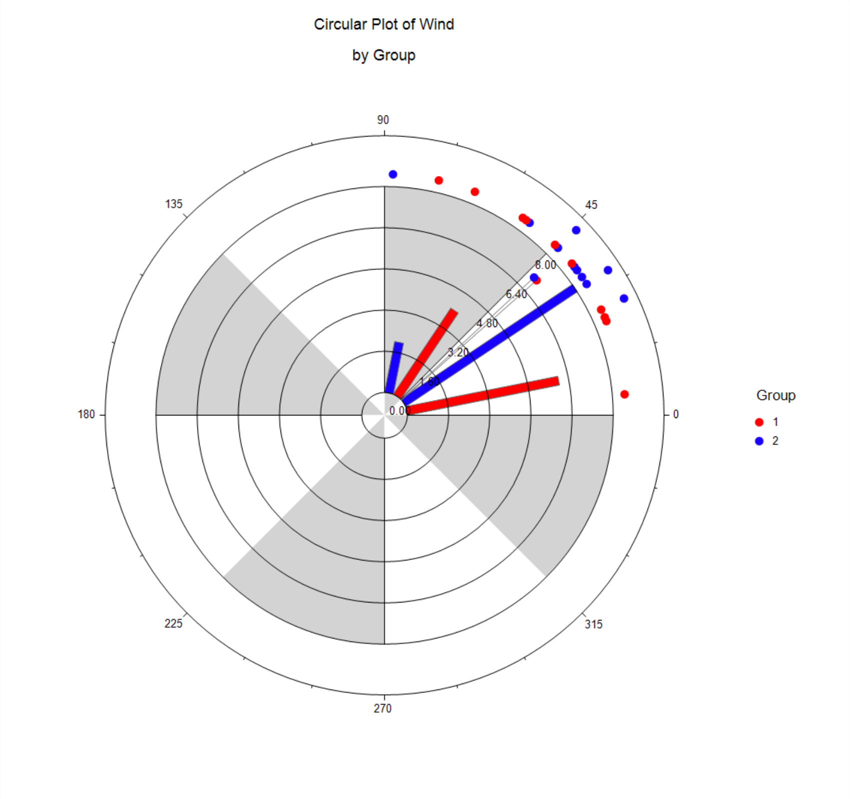


These plots show the distribution of the data around the circle.

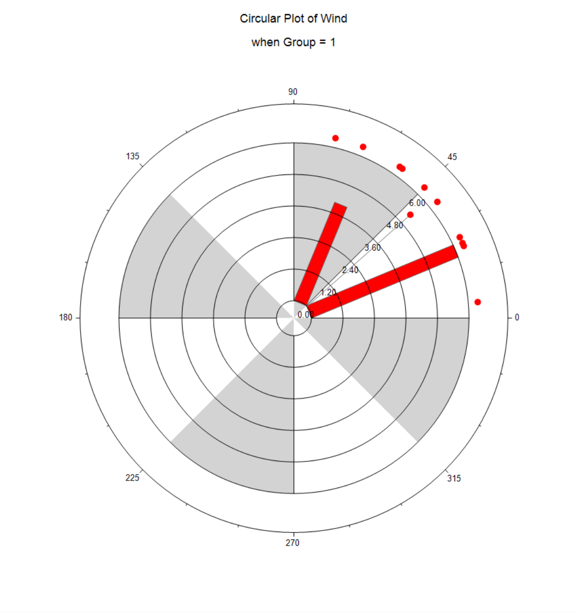
© NCSS, LLC. All Rights Reserved.

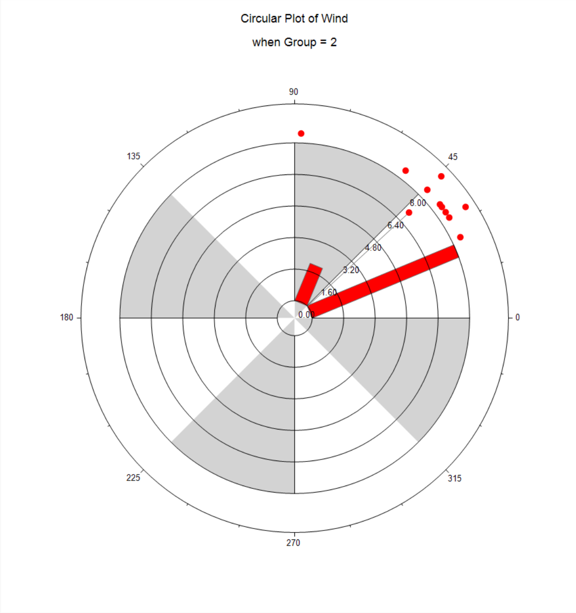
Circular Histograms

Combined Rose Plot



Individual Rose Plots





The circular histograms are generated by clicking the Plot Format buttons and setting Interior Objects to "Circular Histogram."

© NCSS, LLC. All Rights Reserved.