Circular Statistics

Karen Grace-Martin

Circular variables, which indicate direction or cyclical time, can be of great interest to biologists, geographers, and social scientists. The defining characteristic of circular variables is that the beginning and end of their scales meet. For example, compass direction is often defined with true North at 0 degrees, but it is also at 360 degrees, the other end of the scale. A direction of 5 degrees is much closer to 355 degrees than it is to 40 degrees. Likewise, times that represent cycles, such as times of day (best expressed on a 24 hour clock), day in a reproductive cycle, or month of a year are also circular. January, month 1 is closer to December, month 12, than it is to June, month 6.

Examples of circular variables are abundant in biology, geography, and the social sciences. One experiment we saw recently in consulting compared the distance and direction flown by male moths in comparison to unmated and mated female moths under different weather conditions. Another compared the direction and distance traveled by beetles under two different conditions of crowding. Other examples include measures of wind and water flow direction to understand the movement of pollutants and the timing of events within a cycle, such as the when the number of heart attacks peaks within a week or how body temperature fluctuates over a day. Note that time can be considered either circular or linear. Time is circular when it measures part of a cycle, such as the timing of a daily event. It is linear when it measures length of time, such as the number of days since an event.

Most familiar statistics do not work with circular variables because they assume that variables are linear–the lowest value is farthest from the highest value. For example, the average of 5 degrees, 60 degrees and 340 degrees (which are all northerly directions) is 135 degrees–a southerly direction. This is clearly an unreasonable average. Changing 340 degrees to 20 degrees (an equivalent value) changes the mean to 15 degrees, which is more reasonable. But 5 degrees could also be changed to 365 degrees, giving a mean of 255 degrees, also reasonable. Which is right?

Because classical statistical analysis does not work for circular variables, an entire field of circular statistics has been developed. In circular statistics, each datum is defined by its length and its angle from a chosen point on the circle. In the case of the moths, each moth's final location would be designated by the distance it traveled from the release point and the angle in degrees from true north. The mean location of all the moths can be found using the sine and cosine of the angle then adjusting for the length. Because the sine of 0 degrees and 360 degrees is the same, this solves the original problem of ends of the scale being near each other.

Circular statistics include tests of uniform direction around the circle, confidence intervals, tests for comparing two groups of directions, circular graphs, correlations, and regression, among others. Although the theory behind these statistics is not new, there have been no mainstream

statistical packages that could implement them until recently. Now, both Stata and S-Plus have implemented comprehensive circular statistics modules within the last year.

References: Batschelet, E. (1981). Circular Statistics in Biology. Academic Press: London. Jammalamadaka, S.R. & Sengupta, A. (2001). Topics in Circular Statistics. World Scientific, River Edge, N.J.

Created November 2002. Last updated April 2022.