Tutorial 1 - An introduction to circular data

Arjan Huizing

2 April 2018

What is circular data?

Most researchers will be familiar with the concept of measurement levels of variables. The most popular two being continuous data and categorical data. Circular data is another type of measurement with properties that distinguish it from other data types. The main property of circular data is the fact that it is periodical. What this means is that there is no true start or end point, but rather a continuous scale that loops back on itself.

An example would be the time of day as represented on the clock. As time passes the hands on the face of a clock will inevitably be bound to pass the same 12 hours over and over. This also means that distances can not be interpreted as linear. The difference in time between 1 AM and 3 AM is the same as the distance between 11 PM and 1 AM, that is, two hours. Another quality of circular data is that every value can be seen to have an opposite value. When we analyze directions on a compass, we would say that north is the opposite of south, and east is the opposite of west. Circumplex models also use this quality to create a continuum of related values, where opposite values are 180 degrees apart. Examples of this from social and behavioural science are Schwartz' theory of basic values, Russel's mood scale and Leary's interpersonal circumplex.

[Image of circumplex here?]

Circular data can also be used to analyse directions. In a study by Warren et al. (2017), 39 participants were asked to navigate a Euclidian and a non-Euclidian maze using a virtual reality headset. The Euclidian and non-Euclidian mazes were identical except for the presense of two 'wormholes' that would transport participants instantly. The movement of participants through these mazes was tracked for all eight trials. The mazes were designed in a way that the angular error would be higher if participants opted to use wormholes present in the non-Euclidian maze.

The researchers then used this data to test whether participants were capable of learning to navigate a space using these wormholes.
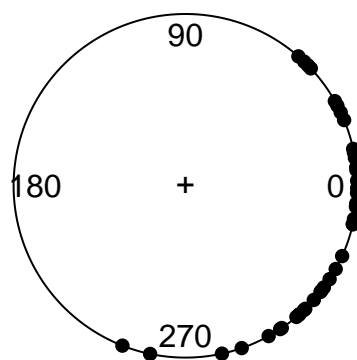
[Image of Mazes here?]

A subset of the data is displayed below, containing the angular error in degrees of 20 participants in the standard trials. Here we see a noticable difference in direction between the Euclidian maze and the non-Euclidian maze.
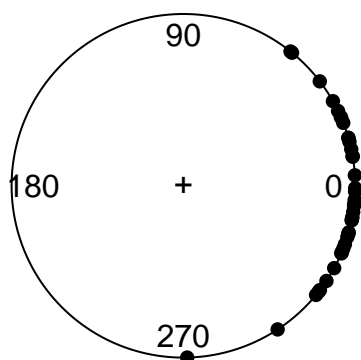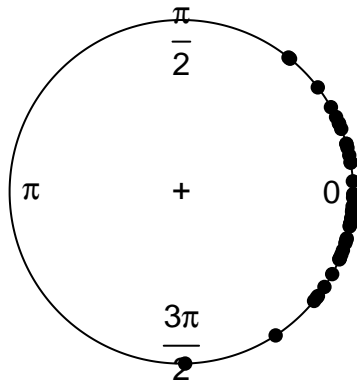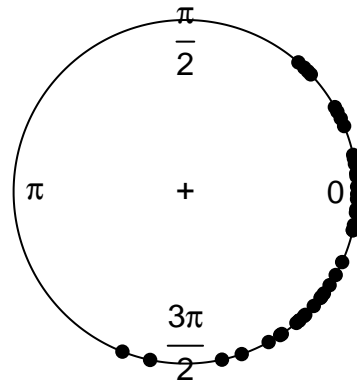
Euclidian maze
Non–Euclidian maze

Note how the output indicates the values in degrees. Measuring circular data in degrees is one of the two common measurement types. A second way to express angles on a circle is by using radians. A circle has a total of 360 degrees, or equivalently, a total of $2\pi \approx 6.28$ radians.

Euclidian maze   **Non–Euclidian maze**

Values in degrees are easily converted to radians and vice versa.

```r
#Converting 90 degrees into radians
degrees <- 90
radians <- degrees*(pi/180)
radians
```

```
## [1] 1.570796
```

```r
#Converting the radians back into degrees
180*(radians/pi)
```

```
## [1] 90
```

So as we can see $90° = \pi/2$ radians.

Depending on the type of data we want to analyze, we use different methods. For example, if we want to predict a categorical outcome a binomial or multinomial logistic regression is prefered to using a linear regression. In the same vein, circular data should be analyzed using methods tailored to the unique attributes of such data. Analyzing circular data as either categorical or continuous data may lead to issues of validity, as even something as simple as a mean value can result in very different outcomes if analyzed as continous instead of circular as we will see in tutorial 2.