Tutorial 3 - Circular distributional assumptions

Arjan Huizing

16 April 2018

## Circular distributions

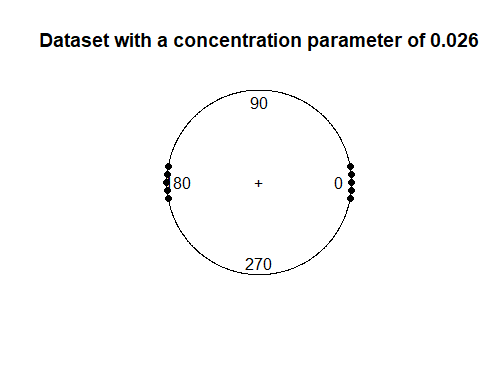
It is important to investigate the way that data is distributed. Not only do many models assume that data has a certain distribution, even simple descriptives such as the median or mean may paint a incorrect picture of our data if we take them at face value.

When we analyse circular data it is just as important that we visualize the distribution of data points. In the previous tutorial we noted it was possible for mean resultant lengths to indicate a very low concentration even when in our data this may not be the case. For example, take the dataset specified below with a mean resultant length of . This value would lead us to believe that the data is spread entirely across the circle.

example <- as.circular(c(350, 355, 0, 5, 10,   
 185, 190, 180, 170, 175), units = "degrees")  
rho.circular(example)

## [1] 3.861386e-17

However, when inspecting the actual distribution we can see that there are two modal groups on opposite ends of the circle.



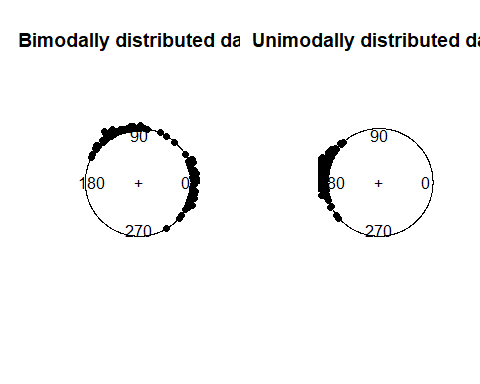
One can also imagine what would happen if we asked for the mean direction of this dataset. As all data points have an exact opposite value, the mean will be undefined.

mean(example)

## Circular Data:   
## Type = angles   
## Units = degrees   
## Template = none   
## Modulo = asis   
## Zero = 0   
## Rotation = counter   
## [1] NA

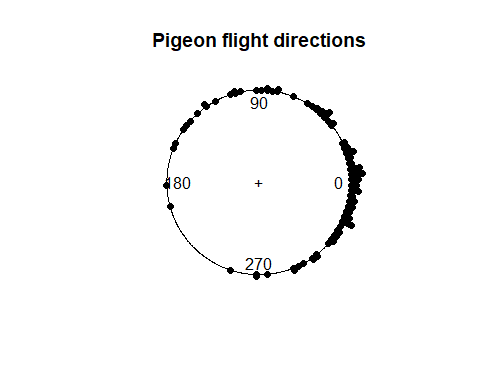
## Distributional assumptions

Many of the statistical models we use also assume the data were obtained from a certain probability distribution.



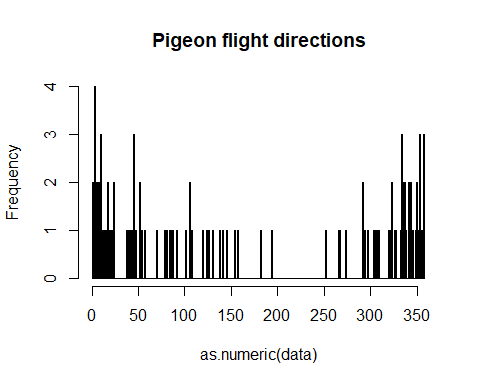
## Plotting circular data in R

It is thus always a good idea to plot the dataset before any analysis takes place. The circular package contains the function plot.circular for this purpose. An example of this is given below, using a dataset on flight direction of pigeons. Note how the argument is passed onto the function. The default value for this is , which would result in similar values overlapping.



It can sometimes also be useful to inspect circular data on a continues scale. This can be done by transforming the data from circular to continues using as.numeric. Keep in mind that values that are outside of the circular range will need to be calculated down to their remainder using .

hist(as.numeric(data),  
 breaks = 360,  
 xlim = c(0,360),  
 main = "Pigeon flight directions")



#Testing symmetry using a symmetry plot?

## Uniformity

Uniformally distributed data indicates an even spread of observations across the entire circular range. This can sometimes be difficult to observe visually, which is why it is recommended to also assess uniformity using uniformity plots and statistical tests.

#test for uniformity.  
  
#uniformity plot  
#kuiper.test(data)  
#rao.test(data)