Estadística circular aplicada a la Ecología

Irene Mendoza1

1. Departamento de Ecología Integrativa. Estación Biológica de Doñana (CSIC).

Autor para correspondencia: Irene Mendoza [[irene.mendoza@ebd.csic.es](mailto:irene.mendoza@ebd.csic.es)]

## Palabras clave

fenología; actividad diaria; test de Rayleigh; test de Mardia-Watson-Wheeler

## Keywords

phenology; diel activity; Rayleigh test; Mardia-Watson-Wheeler test

La estadística circular es una poderosa herramienta que permite analizar variables que tienen una naturaleza cíclica y para los que la estadística lineal no es apropiada. Algunos ejemplos clásicos en Ecología incluyen el estudio de la fenología de una especie a lo largo del año o los patrones de actividad diarios de animales. En estos casos, la designación de un cero (y de valores mayores o menores) es completamente arbitraria. Por ejemplo, la diferencia entre el mes de enero y diciembre es de solo un mes en una escala circular, mientras que en una escala lineal, dicha diferencia sería de 12 meses.

# Transformación en ángulos de nuestra variable de naturaleza circular

La primera transformación que debemos hacer con nuestros datos de naturaleza circular es convertirlos en ángulos mediante una simple ecuación (Zar, 1999):

*X* representa la variable temporal que queremos convertir en un valor angular (en °) y *k* es la cantidad total de unidades de tiempo de nuestra circunferencia (*k* = 365 para días del año o *k* = 12 para meses, por ejemplo). De esta manera, el 14 de febrero, que es el 45º día del año, corresponde a

a <- 360\*45/365

44.38°. En el caso de medidas mensuales, los 360° se dividen en 12 sectores de 30°. Por convención, se considera el punto medio del sector como el correspondiente a cada mes, es decir 15° para enero, 45° para febrero, … y 345° para diciembre.

# Ejemplos usados

Usaré dos ejemplos en esta nota, uno referido a fenología de la fructificación en Nouragues (Guyana francesa) extraído de Mendoza *et al.* (2018) y el otro a los patrones de actividad del agutí y del pecarí de la Isla de Barro Colorado (Rowcliffe et al., 2014), extraídos del paquete activity (Rowcliffe, 2019). Me centraré en tres de los muchos paquetes de R (R Core Team, 2019) que permiten analizar datos circulares: Directional (Tsagris et al., 2020), CircStats (Ulric Lund y Claudio Agostinelli, 2018), y circular (Agostinelli y Lund, 2017), así como el paquete ggplot2 (Wickham, 2016) para las representaciones gráficas. Debido a las limitaciones de espacio de esta nota, se recomienda revisar otras referencias para tener más detalle de los métodos, fórmulas exactas y aplicaciones a la Ecología de la estadística circular (Batschelet, 1981; Zar, 1999; Mardia y Jupp, 2000; Morellato et al., 2010; Staggemeier et al., 2020) así como otros paquetes de R disponibles para análisis circulares (p.ej. cplots, Xu y Wang, 2019).

library(activity)  
library(circular)  
library(CircStats)  
library(Directional)  
library(ggplot2)

aguti <- subset(activity::BCItime, species == "agouti")  
pecari <- subset(activity::BCItime, species == "peccary")  
#se seleccionan los datos del agutí y pecarí de la base de datos BCItime del paquete "activity". El tiempo de muestreo está expresado como una proporción relativa a las 24 horas de un día.  
  
frutofrq <- data.frame(meses = month.abb, meses.ang = seq(15, 350, 30), numspp = c(21, 23, 32, 31, 29, 20, 19, 18, 15, 17, 21, 16))   
  
#estos datos representan el número de especies en fruto por mes en Noura\_gues (extraído de Mendoza et al. 2018).  
#"meses.ang" es el valor angular de cada mes

# Representaciones gráficas de datos circulares

Cualquier estadística descriptiva pasa por la representación gráfica de los datos. En el caso de datos circulares, el paquete circular (Agostinelli y Lund, 2017) permite hacer *diagramas de rosa* convirtiendo los datos circulares en un objeto circular primero y usando luego rose.diag (Fig. 1).

aguti.circ <- circular::circular(aguti$time\*24, units = "hours", template = "clock24")  
pecari.circ <- circular::circular(pecari$time\*24, units = "hours", template = "clock24")  
#convertimos en horas los valores temporales expresados como proporción  
  
#Fig. 1:  
circular::rose.diag(aguti.circ, main = "agutí", bins = 24, prop = 2)  
circular::rose.diag(pecari.circ, main = "pecarí", bins = 24, prop = 2)  
#el argumento "bins" especifica el número de grupos en los que se divide la circunferencia. En nuestro caso son 24, uno por cada hora del día.

En los diagramas de rosa cada sección del histograma (con forma de cuña) muestra el número de observaciones para cada rango de valores angulares (o temporales). El área del sector es proporcional a la frecuencia de cada grupo y el radio, a la raíz cuadrada de la frecuencia relativa de cada grupo. La función coord\_polar del paquete ggplot2 permite muchas más opciones gráficas para representarlos (Fig. 2).

fig2 <- ggplot(frutofrq, aes(x = meses.ang, y = sqrt(numspp))) + geom\_bar(stat='identity') + coord\_polar() + ylim(-7,35) #gracias al parámetro "ylim" controlamos el círculo interno.

# Principales métricas circulares: media angular, desviación estándar angular y vector *r*

La estadística linear no sirve para describir adecuadamente datos circulares. Para eso, se han desarrollado métricas circulares como la media, varianza y desviación estándar circulares. Se calculan fácilmente con la función circ.summary del paquete Directional. Se recuerda que los datos se convierten de grados a radianes usando la fórmula .

aguti.sum <- Directional::circ.summary(aguti$time\*2\*pi, plot = F)  
  
aguti.sum$mesos #media angular en radianes

## [1] 2.985616

aguti.sum$circstd #desviación estándar circular en radianes

## [1] 0.0185485

En el caso de datos resumidos en forma de frecuencias (como ocurre para los datos de fruto de Nouragues), se deben convertir las frecuencias en un vector único que repita cada valor angular el número de veces de su frecuencia. Una vez conseguido este vector, se podrá aplicar la función circ.summary como en el caso anterior.

fruto <- c(rep(15,21), rep(45,23), rep(75,32), rep(105,31), rep(135,29), rep(165,20), rep(195,19), rep(225,18), rep(255, 15), rep(285,17), rep(315,21), rep(345,16))  
  
fruto.sum <- Directional::circ.summary(fruto, plot = F, rads = F)   
  
fruto.sum$MRL

## [1] 0.1558043

De estas métricas, el vector *r* representa la longitud del vector medio (llamado *MRL* por la función circ.summary). Es un valor adimensional que varía entre 0 y 1 y mide la dispersión de los datos (0 indica dispersión uniforme en todas las direcciones y 1, todas las observaciones apuntando en la misma dirección). El vector *r* se utiliza en fenología como una estima de la concentración de la actividad o la estacionalidad de la fenofase, siendo mayor cuanto más cercano a 1. En el caso de la fructificación en Nouragues, al estar los datos agrupados por meses, debemos multiplicar *r* por un factor de corrección , resultando en 0.06.

# Test de uniformidad de una muestra: test de Rayleigh como medida de estacionalidad

La forma de verificar estadísticamente si una distribución circular es significativamente estacional es usando el *test de Rayleigh* (Zar, 1999). La hipótesis nula es que la distribución circular es uniforme, es decir, que no hay una dirección angular preferida. El test de Rayleigh asume que la distribución de los datos es unimodal. Para distribuciones multimodales, existen otros test (ver Landler et al., 2018).

CircStats::r.test(fruto, degree = T)

## $r.bar  
## [1] 0.1558043  
##   
## $p.value  
## [1] 0.001729289

Gracias a este test podemos ver que, a pesar del bajo valor de *r* para Nouragues, la fructificación es significativamente estacional, con el valor máximo en marzo (Fig. 2).

# Test entre dos distribuciones circulares: test de Mardia-Watson-Wheeler aplicado a los patrones actividad de los animales

Muchos trabajos (ver Monterroso et al., 2014 como ejemplo) analizan los patrones de actividad diaria de diferentes especies animales usando el *test de Mardia-Watson-Wheeler* (Batschelet, 1981). Este test permite detectar si dos muestras circulares difieren significativamente entre sí de una forma no paramétrica. La forma más sencilla de calcularlo es usando la función watson.wheeler.test del paquete circular.

circular::watson.wheeler.test(list(aguti.circ, pecari.circ))

##   
## Watson-Wheeler test for homogeneity of angles  
##   
## data: 1 and 2  
## W = 2144.8, df = 2, p-value < 2.2e-16

En el ejemplo de los datos de actividad del agutí y el pecarí, tal y como se observaba en la Fig. 1, el test nos demuestra que las dos especies difieren en su patrón de actividad diaria.

# Consideraciones finales

En esta nota apenas se han esbozado algunas de las aplicaciones básicas de la estadística circular a la Ecología, especialmente para la fenología o el estudio de patrones de actividad diaria de animales, resaltando la utilidad de esta herramienta estadística. Obviamente, hay muchas métricas y otros análisis, como las correlaciones circular-linear o circular-circular, que han quedado sin explicar por falta de espacio. Gracias a la funcionalidad de los paquetes de *R* especializados en estadística circular, estas métricas y análisis están fácilmente disponibles para el usuario interesado que quiera seguir avanzando.

El código necesario para reproducir este documento se puede consultar en [GitHub](https://github.com/ecoinfAEET/Notas_Ecosistemas/tree/master/estadistica_circular).

# Agradecimientos

Esta nota es un producto del proyecto TEMPNET, financiado con una beca Marie-Sklodowska Curie (798269 - TEMPNET - H2020-MSCA-IF-2017).

### REFERENCIAS

Agostinelli, C., Lund, U. 2017. *R package circular: Circular Statistics (version 0.4-93)*. CA: Department of Environmental Sciences, Informatics; Statistics, Ca’ Foscari University, Venice, Italy. UL: Department of Statistics, California Polytechnic State University, San Luis Obispo, California, USA.

Batschelet, E. 1981. *Circular Statistics in Biology*. 1.ª ed. Academic Press, London.

Landler, L., Ruxton, G.D., Malkemper, E.P. 2018. Circular data in biology: advice for effectively implementing statistical procedures. *Behavioral Ecology and Sociobiology* 72:.

Mardia, K.V., Jupp, P.E. 2000. *Directional Statistics*. 1.ª ed. John Wiley & Sons, Chichester.

Mendoza, I., Condit, R.S., Wright, S.J., Caubère, A., Châtelet, P., Hardy, I., Forget, P.-M. 2018. Inter-annual variability of fruit timing and quantity at Nouragues (French Guiana): insights from hierarchical Bayesian analyses. *Biotropica* 50: 431-441.

Monterroso, P., Alves, P.C., Ferreras, P. 2014. Plasticity in circadian activity patterns of mesocarnivores in Southwestern Europe: implications for species coexistence. *Behavioral Ecology and Sociobiology* 68: 1403-1417.

Morellato, L.P.C., Alberti, L.F., Hudson, I.L. 2010. Applications of Circular Statistics in Plant Phenology: a Case Studies Approach. En Hudson, I. L., Keatley, M. R. (eds.), *Phenological Research: Methods for Environmental and Climate Change Analysis*, pp. 339-359. Springer Netherlands, Dordrecht.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rowcliffe, J.M., Kays, R., Kranstauber, B., Carbone, C., Jansen, P.A. 2014. Quantifying levels of animal activity using camera trap data. *Methods in Ecology and Evolution* 5: 1170-1179.

Rowcliffe, M. 2019. *activity: Animal Activity Statistics*.

Staggemeier, V.G., Camargo, M.G.G., Diniz-Filho, J.A.F., Freckleton, R., Jardim, L., Morellato, L.P.C. 2020. The circular nature of recurrent life cycle events: a test comparing tropical and temperate phenology. *Journal of Ecology* 108:.

Tsagris, M., Athineou, G., Sajib, A., Amson, E., Waldstein, M.J. 2020. *Directional: Directional Statistics*.

Ulric Lund, S.-p. original by, Claudio Agostinelli, R. port by. 2018. *CircStats: Circular Statistics, from "Topics in Circular Statistics" (2001)*.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Xu, D., Wang, Y. 2019. *cplots: Plots for Circular Data*.

Zar, J.H. 1999. *Biostatistical Analysis*. 4.ª ed. Prentice Hall, New Jersey.

###### PIES DE FIGURA

**Figura 1**. Diagramas de rosa representando la actividad diaria del agutí y el pecarí en la Isla de Barro Colorado. (Datos extraídos de Rowcliffe *et al.* 2014).

**Figura 2**. Diagrama de rosa representando el número de especies con fruto en cada mes del año en la reserva de Nouragues (Guyana Francesa). Datos extraídos de Mendoza *et al.* 2018.

###### FIGURA 1

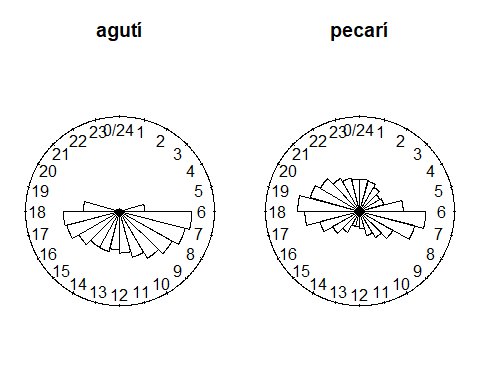


Figura 1.

###### FIGURA 2

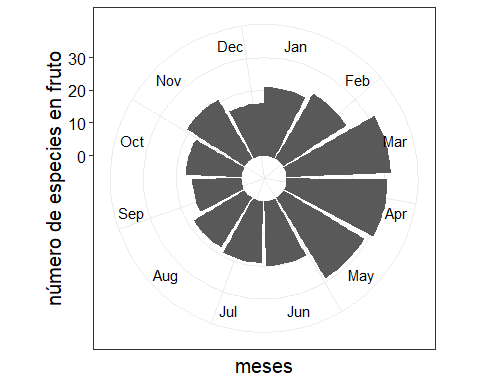


Figura 2.