

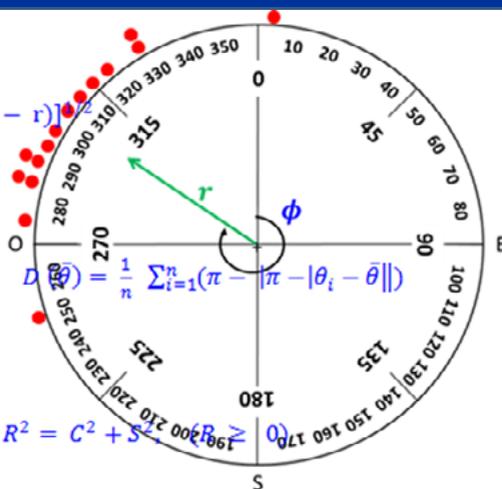
# Introducción a la Estadística Circular

$$\bar{x} = \frac{\sum f_i \cdot \cos \phi_i}{n};$$

$$\bar{y} = \frac{\sum f_i \cdot \text{sen} \phi_i}{n}$$

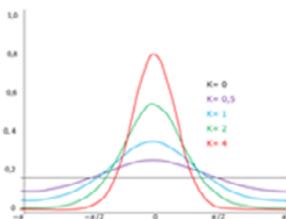
$$s = \frac{180^\circ}{\pi} [2 \cdot (1 - r)]$$

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} = \sqrt{(0,2740)^2 + (0,2358)^2} = 0,3616$$



$$C = \sum_{i=1}^n \cos \theta_i, \quad S = \sum_{i=1}^n \text{sen} \theta_i, \quad R^2 = C^2 + S^2$$

$$F(\theta) = \int_0^{\theta_2} f(\theta) d(\theta)$$



José Luis Pérez Bote

UNIVERSIDAD DE EXTREMADURA



2020



---

# Introducción a la Estadística Circular



---

José Luis Pérez Bote

# Introducción a la Estadística Circular



Cáceres  
2019



Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) si necesita fotocopiar o escanear algún fragmento de esta obra.



© José Luis Pérez Bote, para esta edición  
© Universidad de Extremadura, para esta edición

Tipografía utilizada: Palatino Linotype (para cubierta), Bembo Std (para páginas iniciales) y Palatino LT Std (para el texto de la obra)

Edita:  
Universidad de Extremadura. Servicio de Publicaciones  
Plaza de Caldereros, 2. 10071 Cáceres (España)  
Tel. 927 257 041; Fax 927 257 046  
[publicac@unex.es](mailto:publicac@unex.es)  
<http://www.unex.es/publicaciones>

I.S.B.N.: 978-84-09-15443-2

Impreso en España - *Printed in Spain*

*Impresión:* Dosgraphic, s. l.

*Para Maite y Natalia*



# ÍNDICE

	<u>Páginas</u>
Prólogo .....	13
Abreviaturas .....	15
<b>Capítulo 1. Introducción.....</b>	<b>17</b>
1.1. Introducción.....	17
1.2. Antecedentes históricos .....	21
1.3. Obras de referencia .....	24
1.4. Programas informáticos.....	25
1.5. Casos prácticos.....	26
<b>Capítulo 2. Los datos .....</b>	<b>31</b>
2.1. Introducción.....	31
2.1.1. El Ángulo .....	31
2.1.2. Vectores .....	33
2.2. Funciones trigonométricas .....	36
2.3. Tipos de datos .....	37
2.3.1. Ángulos .....	37
2.3.2. Radianes .....	37
2.3.3. Datos axiales.....	37
2.3.4. Direcciones.....	39
2.3.5. Tiempo en horas.....	40
2.3.6. Días de la semana .....	42
2.3.7. Meses .....	42
2.3.8. Días del año.....	42
2.3.9. Datos agrupados.....	42
2.3.10. Datos apareados .....	42
2.3.11. Datos empatados.....	43

<b>Capítulo 3. Estadística descriptiva .....</b>	<b>45</b>
3.1. Introducción.....	45
3.2. Medidas de centralización.....	45
3.2.1. Ángulo medio.....	45
3.2.2. Media de los ángulos medios .....	50
3.2.3. Mediana angular.....	52
3.2.4. Moda angular .....	53
3.2.5. Rango circular .....	54
3.3. Medidas de dispersión.....	55
3.3.1. Longitud del vector medio como medida de concentración...	55
3.3.2. Varianza angular y desviación angular .....	56
3.3.3. Distancia angular.....	57
3.3.4. Límites de confianza para la media.....	57
3.4. Medidas de forma.....	58
3.4.1. Asimetría.....	58
3.4.2. Curtosis.....	58
3.5. Representación de los datos.....	59
3.5.1. Introducción .....	59
3.5.2. Gráficos de distribución.....	60
3.5.3. Gráfico de dispersión .....	62
3.5.4. Histogramas lineales .....	62
3.5.5. Gráfico de datos originales.....	64
3.5.6. Histogramas circulares (o angulares) .....	65
3.5.7. Diagramas de rosa.....	66
3.5.8. Densidad de Kernel .....	67
3.5.9. Diagramas de tallo y hoja.....	68
3.6. Distribuciones circulares.....	68
3.6.1. Introducción .....	68
3.6.2. Distribución uniforme .....	72
3.6.3. Distribución de von Mises.....	73
3.6.4. Distribución cardioide .....	77
3.6.5. Distribución wrapped.....	78
3.6.6. Distribución wrapped normal .....	78
<b>Capítulo 4. Inferencia estadística .....</b>	<b>79</b>
4.1. Introducción.....	79
4.2. Una muestra .....	79
4.2.1. Test de Rayleigh .....	80

4.2.2.	Test de Hodges-Ajne .....	83
4.2.3.	Test V .....	84
4.2.4.	Test de Rao (Test del espaciado de Rao) .....	86
4.2.5.	Test de Kuiper.....	89
4.2.6.	Test de la chi-cuadrado.....	91
4.2.7.	Test $U^2$ de Watson.....	93
4.2.8.	Test de Hotelling para una muestra.....	94
4.2.9.	Test de Moore para una muestra.....	97
4.2.10.	Bootstrapping.....	98
4.3.	Dos o más muestras.....	99
4.3.1.	Test de Watson-Williams .....	100
4.3.2.	Test de Mardia-Watson-Wheeler .....	108
4.3.3.	Test $U^2$ de Watson.....	113
4.3.4.	Test de la chi-cuadrado .....	115
4.3.5.	Test de Hotelling para dos muestras apareadas.....	116
4.3.6.	Test de Moore para dos muestras.....	118
4.3.7.	Test no paramétrico de Wallraff .....	121
4.4.	ANOVA Circular.....	122
<b>Capítulo 5.</b>	<b>Correlación.....</b>	<b>125</b>
5.1.	Introducción.....	125
5.2.	Correlación circular-circular.....	125
5.3.	Correlación circular-lineal .....	128
<b>Capítulo 6.</b>	<b>Bibliografía.....</b>	<b>131</b>
<b>Anexo I.</b>	<b>Tablas .....</b>	<b>133</b>



# PRÓLOGO

La Estadística Circular se ocupa del análisis de datos circulares, unos datos con unas características muy particulares. En primer lugar, la escala en la que se desenvuelven es finita, ya que tiene una amplitud de  $0^\circ$  a  $360^\circ$ . En segundo lugar, la posición del cero es arbitraria, ya que no existe argumento científico para no poder situarlo en cualquier punto del círculo. En tercer lugar, no existen magnitudes cuando tratamos con datos circulares, ya que  $270^\circ$  no son *más* que  $180^\circ$ , sino que *señalan* direcciones diferentes. Estas (y algunas más) hacen que los procedimientos utilizados en estadística lineal no sean aplicables a los datos circulares, por ello nace la Estadística Circular, para que se puedan tratar este tipo de datos.

La mayoría de los datos circulares se obtienen en forma de direcciones u orientaciones, pero también son datos circulares fenómenos cíclicos, es decir, aquellos que pueden acontecer en una hora, un día, una semana o un año. Dada la versatilidad de los datos circulares son muchas las disciplinas en las que se utilizan datos circulares, quizás en la que más se utilicen sea en Biología o, por ser más exactos, en las diferentes ramas de la Biología: Botánica, Etología, Fisiología, Geología, Biología Molecular, etc., pero también son empleados en Medicina, Meteorología, Ciencias de la Computación, Sociología, Ciencias Políticas, etc. Puede decirse, por tanto, que la Estadística Circular es una herramienta que es utilizada cada vez con mayor frecuencia por investigadores de, prácticamente, todas las ramas de la Ciencia y la Tecnología.

Esta obra surge con la intención de dar a conocer al lector, preferentemente estudiantes e investigadores, los principios básicos de la Estadística Circular. Se asume que el usuario posee unos conocimientos mínimos de estadística, por ello se han evitado los complejos y, a veces difíciles de entender, procedimientos matemáticos, para lo cual existen excelentes obras de referencia que el lector podrá encontrar en el capítulo de bibliografía. Se ha pretendido, por tanto, desarrollar una obra en la que predomine la claridad y la sencillez siguiendo un procedimiento “paso a paso” para que el interesado, de este modo, pueda entender y aplicar los distintos conceptos que se exponen. Para facilitar este proceso se ha optado por incluir ejercicios resueltos. En este sentido es conveniente aclarar que los ejercicios se han desarrollado de modo que el lector pueda realizarlos con la ayuda de una calculadora o una hoja de cálculo, evitando el empleo de programas informáticos (que existen para tal fin), que pueden suponer el estudio y aprendizaje de complejos manuales o la compra de los mismos.

El acceso a los métodos estadísticos se ve dificultado a veces por el reducido número de libros de texto sobre el tema, generalmente todos ellos en inglés y además principalmente dirigidos hacia aspectos teóricos más que a su empleo práctico. Es importante señalar también que este trabajo contribuye de forma notable a las publicaciones en castellano sobre Estadística Circular, circunstancia que puede alentar a muchos lectores a iniciarse en esta disciplina. Los precedentes de la Estadística Circular en castellano habría que situarlos en la obra de R. Lahoz-Beltrá, J. Ortega-Escobar y C. Fernández-Montraveta (1994), *Métodos estadísticos en biología del comportamiento*, Madrid, Editorial Complutense, donde hay un capítulo dedicado a la estadística circular. Además, cabe señalar la obra de M. Mena (2005), *Aplicaciones de Estadística Circular a problemas de Ciencias Naturales*, México, Akadia. En esta obra los diferentes temas no se tratan de un modo exhaustivo sino que se exponen empleando un enfoque matemático simple con el objeto de facilitar su aplicación práctica. Para una mayor comprensión en cada tema se desarrollan ejemplos como problemas resueltos. Como puede comprobarse las dos obras citadas son antiguas y es complicado conseguirlas.

Este trabajo ha sido dividido en cinco capítulos. El primero es un capítulo introductorio donde se explica que es la Estadística Circular, sus orígenes e historia, para finalizar con una breve reseña sobre las principales obras publicadas sobre el tema, los principales programas informáticos que se pueden utilizar y una serie de ejemplos obtenidos de la bibliografía, que permitirán evaluar la importancia y aplicabilidad de esta disciplina. En capítulo dos está dedicado a los datos empleados en estadística circular, aspecto sumamente importante ya que, aunque parezca lo contrario, son numerosas las formas de poder expresar datos circulares. El tercer capítulo está dedicado a la estadística descriptiva (poco hay que decir de ella), a las representaciones gráficas (muy importantes) y a las distribuciones circulares. En este sentido hay que señalar que, dada complejidad matemática del desarrollo de las distribuciones circulares, se ha optado por hacer una presentación básica, a la vez que sencilla, de las mismas. En este caso vuelvo a remitir al lector a obras especializadas. El capítulo cuarto está dedicado a la inferencia estadística, básicamente a los test de uniformidad y a los test de bondad del ajuste. En el capítulo cinco se aborda otro aspecto interesante de las relaciones entre variables circulares y lineales: la correlación circular-circular y circular-lineal. Finalmente el apartado de bibliografía recoge los libros y artículos citados en el texto (obviamente) y que constituyen, sin embargo, las principales obras de referencia sobre Estadística Circular. Dado que a lo largo del texto se ha seguido un procedimiento "paso a paso" como ya se ha comentado, se ha incluido un anexo con las tablas de los valores críticos de los test estadísticos utilizados en el libro.

El seguimiento de la obra no requiere de unos conocimientos estadísticos o matemáticos de gran nivel, simplemente unas nociones básicas de estadística, álgebra y trigonometría serán suficientes para que el lector se introduzca en este interesante y cada vez más demandado campo de la Estadística.

# ABREVIATURAS

$\arctan$	arcotangente
$a$	dirección angular
B	estadístico del test de Mardia-Watson-Wheeler
$\cos$	coseno
C	suma de los cosenos de una muestra de n datos angulares
D	límite de confianza para la media
$D(\bar{\theta})$	desviación circular estándar
$D_{(\alpha)}$	desviación respecto a un ángulo dado
$df$	grados de libertad
$D_{\alpha_1-\alpha_2}$	distancia angular entre dos ángulos
$f_{(\theta)}$	función de probabilidad de densidad
F	estadístico del test de Hotelling, estadístico del test de Watson-Williams
$F_{(\theta)}$	función de distribución
$\hat{f}_{(\theta)}$	estimador de Kernel
H	ventana o banda del estimador de Kernel
k	parámetro de Kernel
$\hat{K}$	curtosis
$L_1, L_2$	límites de confianza para la media
$m$	vector medio
$r$	longitud del vector medio
$r_c$	longitud del vector medio corregida
$\hat{r}$	asimetría
$r_{aa}$	coeficiente de correlación entre dos variables circulares
$r_{al}$	coeficiente de correlación entre una variable circular y otra lineal
R	longitud del vector resultante
R	R de Rayleigh
$\bar{R}$	longitud resultante media
$\bar{R}_2$	longitud resultante media de los ángulos doblados
$R'$	estadístico del test de Moore
$(r_p, \phi)$	coordenadas polares

$sen$	seno
$S$	suma de los senos de una muestra de $n$ datos angulares
$s$	desviación angular media
$s^2$	varianza angular
$S^2$	varianza circular
$tan$	tangente
$T$	longitud del arco entre dos observaciones adyacentes
$(x, y)$	coordenadas rectangulares
$U$	estadístico del test de Rao
$U^2$	estadístico de la $U$ de Watson
$V$	varianza circular
$V_n^*$	estadístico del test de Kuiper
$w$	rango circular
$X$	eje de abscisas
$\bar{x}$	media de las coordenadas rectangulares $x$ de $n$ datos angulares
$\bar{X}$	media de las coordenadas rectangulares $x$ de $n$ datos angulares
$Y$	eje de ordenadas
$\bar{y}$	media de las coordenadas rectangulares $y$ de $n$ datos angulares
$\bar{Y}$	media de las coordenadas rectangulares $y$ de $n$ datos angulares
$z$	estadístico del test de Rayleigh
$\alpha$	nivel de significación (normalmente 0,05)
$\bar{\phi}$	ángulo medio
$\tilde{\phi}$	mediana
$\bar{\theta}$	dirección media
$\lambda$	longitud del arco de una circunferencia
$\pi$	número pi: 3,1415192653589
$\kappa$	parámetro de concentración
$\rho$	longitud resultante media
$\kappa$	desviación circular estándar
$\hat{\delta}$	dispersión circular
$u$	estadístico del test $V$
$\chi^2$	chi cuadrado

# CAPÍTULO 1

## INTRODUCCIÓN

### 1.1. INTRODUCCIÓN

La Estadística se ha convertido en una herramienta indispensable para la mayoría de los científicos. Pero ¿qué es la estadística y como pueden utilizarse las técnicas estadísticas para responder a las cuestiones prácticas propuestas por los científicos? La Estadística es un conjunto de teorías y métodos científicos que han sido desarrollados para tratar la recolección, el análisis y la descripción de datos con el objeto de extraer conclusiones útiles para la solución de algún problema en particular de algún universo colectivo. Comúnmente la Estadística apoya al investigador a inferir sobre los parámetros de la población a partir de estadísticos muestrales.

Podemos decir, por tanto, que en Estadística se trabaja con datos de diversa naturaleza. Ahora es importante definir que es un dato. Según Sokal y Rohlf (1979) un *dato* es “una observación o medida tomada sobre la más pequeña unidad muestral”. Si, por ejemplo, medimos el peso de 100 ratas, entonces el peso de cada rata es una observación individual. El conjunto de los 100 pesos representa una *muestra*, definida como “una colección de observaciones individuales seleccionada por un procedimiento determinado” (Sokal y Rohlf, 1979). Finalmente una *variable* es una “propiedad con respecto a la cual los elementos de una muestra se diferencian en algo verificable” (Sokal y Rohlf, 1979).

En Ciencia y Tecnología los datos son sorprendentemente variados en lo que respecta a sus medición, confiabilidad y replicabilidad. Esto ha permitido que se hayan desarrollado métodos más efectivos para enfrentar estos desafíos, entre los que se incluyen aquellos que permiten cuantificar efectos temporales o espaciales. En disciplinas como Medicina, Biología, Geología, Meteorología, Sociología, etc. se utilizan los denominados *datos circulares* para poder interpretar fenómenos y procesos relacionados con dispersiones o movimientos de animales, la dirección del viento, el campo magnético terrestre, ciclos biológicos, estructura de proteínas, etc. La Estadística Circular es la rama de la Estadística que tiene como objeto de estudio aquellos datos que pueden ser representados en un círculo de radio 1, conocido como *círculo unidad*. A este tipo de datos se les denomina *datos circulares* para diferenciarlos de los denominados datos lineales. Por tanto, el círculo unidad representa para los datos circulares lo mismo que la recta real para los datos lineales.

Los datos circulares se obtienen a partir de variables circulares o direccionales. Estas variables se caracterizan por no tener un cero verdadero y porque la designación

de valores máximos y mínimos es arbitraria. Estas dos características hacen que la estadística lineal no sea adecuada para tratar este tipo de datos, ya que las medidas son angulares, no escalares y, por tanto, debemos aplicar la estadística circular, también conocida como estadística de datos circulares (Fisher *et al.*, 1987). Los datos circulares se pueden obtener de diversas formas. Las dos principales corresponden a los instrumentos de medición circular clásicos como la brújula y el reloj (Fig. 1), pero también podemos utilizar el transportador de ángulos o el teodolito. Entre las observaciones típicas medidas por la brújula se encuentran la dirección de viento, las direcciones migratorias de las aves o el vuelo de las abejas hacia la colmena. Ejemplos que incluyen el reloj como instrumento de medida pueden ser los tiempos de llegada de los pacientes a una unidad de urgencias de un hospital, las acumulaciones de tráfico en las carreteras, etc. El caso del reloj nos sirve para hacer constar que no todos los datos circulares son medidos o estimados como ángulos, como acontece con el tiempo, por ejemplo. Sin embargo, sabemos que cada “tiempo” corresponde a una posición en la esfera del reloj y estas posiciones pueden ser convertidas en ángulos que, a su vez, pueden ser usados para definir puntos en el círculo unidad. Es fácil comprobar como, por ejemplo, un periodo de 24 horas equivale a un ciclo de  $360^\circ$ , correspondiendo cada hora a unos  $15^\circ$ . Es posible, por tanto, convertir cualquier unidad de tiempo en un ángulo (Fig. 1). Otros conjuntos de datos similares surgen al considerar las veces que un evento acontece (frecuencia) en un año, en un mes, una semana o un día. Por tanto, los denominados *ritmos biológicos* pueden considerarse variables circulares, de modo que los métodos estadísticos aplicados para el estudio de las direcciones son válidos para el análisis de los biorritmos. No obstante, hay que ser precavidos en el análisis de este tipo de datos para no confundir biorritmos con los análisis de series temporales.

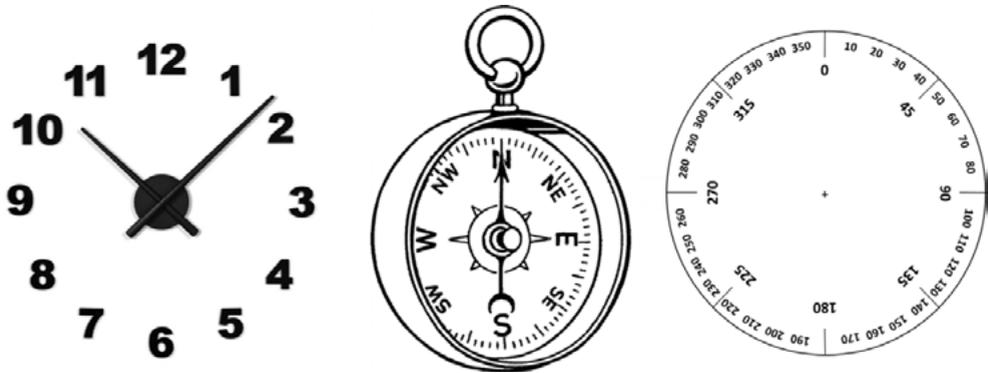
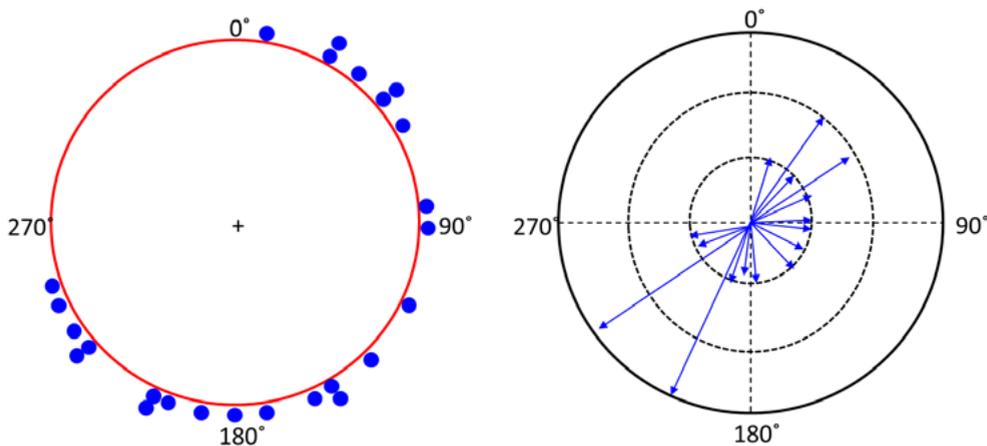


Figura 1. Algunos objetos para la medición de datos angulares y su “equivalente”: un círculo dividido en  $360^\circ$ .

Cuando obtenemos un dato con una brújula, por ejemplo, lo que “medimos” es una dirección que expresamos en grados sexagesimales o radianes en sentido positivo o negativo respecto a una *dirección cero*. Cuando medimos datos circulares es obligatorio establecer previamente el sentido positivo y negativo y la posición del origen respecto a la que se van a tomar los datos. En estadística circular “podemos” establecer

arbitrariamente el sentido positivo y el origen, circunstancia que no ocurre con los datos lineales, pues sabemos que el origen es 0 y a la izquierda de cero se sitúan los valores negativos, mientras que a la derecha lo hacen los valores positivos. Para los datos circulares cada ángulo define un punto en el círculo unidad, de igual manera que cada valor de una variable lineal define un punto en la recta real.

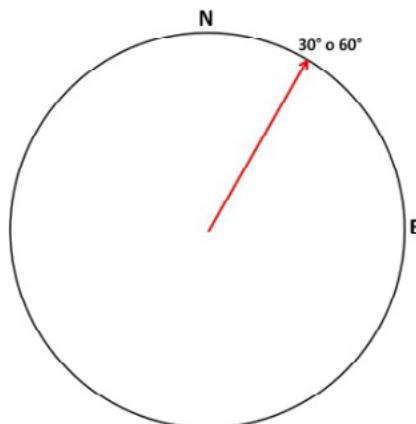
En el párrafo anterior se ha señalado que para expresar una *dirección* necesitamos un sistema de referencia que nos permita ubicar un objeto en el espacio y así poder comparar con otras direcciones. Las direcciones bidimensionales pueden ser representadas como ángulos respecto a una “dirección cero” (punto de inicio) y un “sentido de rotación” siguiendo la dirección de las agujas del reloj (considerado como sentido positivo) o en sentido contrario a las agujas del reloj (dirección negativa). En estudios de orientación la *dirección cero* suele establecerse en el norte ( $0^\circ$ ), aunque también es frecuente que se utilice como *dirección cero* el este. Debido a que la *direcciones* no tiene magnitud ( $350^\circ$  no son más que  $125^\circ$ , solo señalan direcciones diferentes) pueden ser representadas como puntos en una circunferencia de radio 1 y centrada en el origen, o como vectores que conectan el centro con los puntos (Fig. 2). Esta circunferencia de radio 1 es el ya conocido círculo unidad, ¿por qué la circunferencia tiene radio 1? Muy sencillo: al carecer las direcciones de magnitud podemos simplificar los procedimientos matemáticos utilizando el valor más bajo posible, en este caso el 1.



**Figura 2.** Dos formas de representar los mismos datos circulares (puntos y vectores).

Los datos direccionales tienen características únicas en lo que respecta a su representación y tratamiento estadístico, que los hacen distintos a los datos lineales como ya se ha señalado. Por ejemplo, la representación de una dirección en forma de ángulo o de vector no es única, ya que va a depender de donde se establezca el cero y del sentido de giro, como se ha comentado con anterioridad. De este modo, al observar la figura 3 podemos obtener conclusiones diferentes en función de quien realice la observación: para un matemático el vector indica  $60^\circ$ , mientras que para un ornitólogo señala  $30^\circ$ , debido a que la posición del cero difiere: para un matemático el cero está en la posición este, mientras que para un ornitólogo está en posición norte

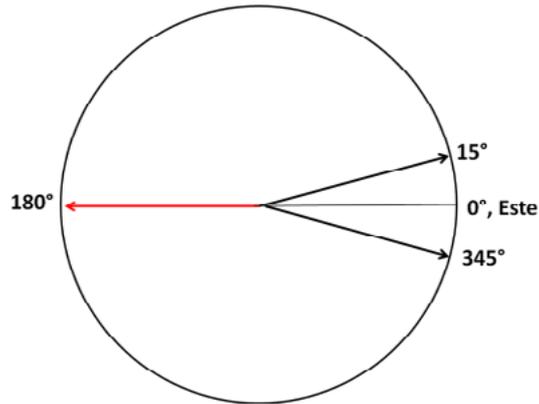
(en estudios con direcciones se acepta generalmente que el norte esté a  $0^\circ$  como ya se ha comentado). Por tanto, la dirección cero es arbitraria. No existe una justificación científica que permita posicionar el cero en un lugar determinado del círculo.



**Figura 3.** La flecha puede indicar dos direcciones diferentes en función de donde se establezca el punto de origen.

Es interesante destacar también que la “escala” en la que se miden los ángulos es finita, oscilando entre  $0^\circ$  y  $360^\circ$ , lo que obliga a “reducir” muchos datos circulares en operaciones algebraicas. Es decir, si nos excedemos de  $360^\circ$ , al doblar ángulos por ejemplo, nos veremos “obligados” a ajustar los datos hasta poder incluirlos en la escala  $0-360^\circ$  (esto se hace con una simple resta). Además, para los datos circulares el “principio” y el “final” de la escala coinciden, ya que  $0^\circ$  y  $360^\circ$  (el origen y final de la escala para los datos circulares) coinciden en el mismo punto de la circunferencia siendo, por tanto, el mismo ángulo. En el caso de los números reales 355 está cerca de 358 y ambos quedan lejos de 0, pero en una circunferencia  $355^\circ$  está también próximo a  $358^\circ$  pero, a diferencia de lo que ocurre con los números reales, los dos están próximos a  $0^\circ$ . Finalmente, tenemos que recordar que en el análisis de datos direccionales las magnitudes no existen, es decir,  $250^\circ$  no son más que  $75^\circ$ , simplemente indican direcciones distintas.

Las características anteriormente señaladas hacen que el análisis de datos direccionales sea diferente al de los datos lineales. Muchos investigadores, generalmente por no estar familiarizados con la estadística circular, aplican erróneamente procedimientos de la estadística lineal a datos circulares, obteniendo conclusiones erróneas ya que, si bien algunos tipos de análisis de la estadística lineal pueden aplicarse a datos circulares, la mayoría no son apropiados. La media es un estadístico que ilustra muy bien este caso. Supongamos que un ave vuela en una dirección de  $15^\circ$  y que otra lo hace en una dirección de  $345^\circ$  ¿Cuál es la dirección media? (hemos tomado el este como la dirección cero y el sentido contrario a las agujas del reloj como el sentido positivo de la rotación). Si utilizamos la media aritmética la dirección media es de  $180^\circ$ , pero este dato es totalmente erróneo como se aprecia en la figura 4, ya que si las aves vuelan hacia el este la media no puede salir hacia el oeste.



**Figura 4.** La media aritmética no es apropiada para determinar la “dirección media”.

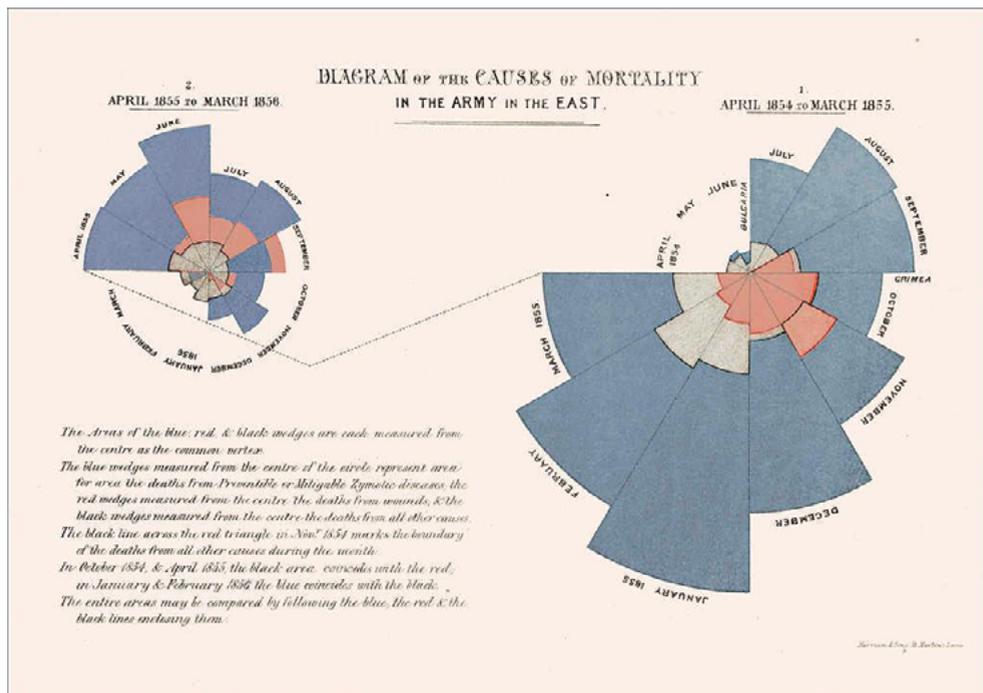
## 1.2. ANTECEDENTES HISTÓRICOS

Aunque aún no es muy utilizado, el análisis de datos circulares es una herramienta matemática con muchos años de historia. Los primeros gráficos que representan “datos” circulares están datados a finales del primer milenio. Funkhouser (citado por Fisher, 1993) comenta que en el siglo X u XI se realizaban estudios astrológicos con base en el movimiento de los planetas a través del zodiaco. Estos movimientos ya quedaron reflejados en gráficos circulares. También existen gráficos de este tipo en estudios realizados sobre el campo magnético de la tierra datados en 1510 y 1701. En este último caso, sir Edmund Halley publicó un mapa que mostraba las líneas de la declinación magnética para todas las aguas navegables del mundo.

El análisis de datos circulares se inicia a mediados del siglo XVIII. En 1767, el reverendo John Mitchell estudiando las separaciones angulares entre las estrellas habla ya de distribuciones uniformes de trayectorias. Medio siglo después John Playfair señala la necesidad de analizar los datos direccionales de forma diferente a como se hace con los “datos ordinarios”. Playfair ya recomienda el método del vector resultante para obtener direcciones medias en estudios geológicos.

En 1858, la enfermera Florence Nightingale, de la Armada Británica, durante la guerra de Crimea, aplicó una serie de métodos que le ayudaron a registrar y analizar de una manera más sencilla los registros sanitarios. En esa época el uso de la estadística no era muy común, pero se necesitaba avanzar en las reformas médicas y considerar que los fenómenos sociales podían ser objetivamente medidos haciendo análisis matemáticos. Uno de estos fue una nueva gráfica, a la que llamó CoxComb, la cual consistía de un círculo dividido en doce partes iguales, con diferentes colores que representaban los meses del año (Fig. 5). Este diagrama es una primera referencia de lo que actualmente se conoce como diagrama de rosa.

Desde estos primeros ejemplos de análisis de datos circulares en Astronomía y Medicina han aparecido aplicaciones en otras disciplinas científicas como Biología, Geografía, Geología, Meteorología y Oceanografía, siendo la Biología y la Geología las disciplinas donde se han realizado más estudios.



**Figura 5.** Diagrama diseñado por Florence Nightingale, precursor del diagrama de rosa actual (<<http://economiafeminista.com/florence-nightingale/>>).

Batschelet (1981) expone en su libro *Circular Statistics in Biology* una amplia metodología para analizar datos circulares en un contexto biológico. También aparece capítulos dedicados a este tema en las publicaciones de Zar (1984, 1996), Upton y Fingleton (1989) y Cabrera *et al.* (1991). Son particularmente interesantes los estudios sobre orientación en libélulas, salamandras, peces, aves y crustáceos que aparecen en la obra de Upton y Fingleton (1989).

En el campo de la Geología las cosas son bien diferentes. Fisher (1993) atribuye los escasos métodos de análisis apropiados para esta disciplina a la falta de interés de los estadísticos por las Ciencias de la Tierra. No obstante, en este campo se hicieron importantes contribuciones en el análisis de datos, como fueron las aportaciones de Lord Rayleigh (1880) en el desarrollo de los test de uniformidad, las de Schmidt (1971) sobre la representación de datos circulares o la de von Mises (1918) la distribución de probabilidad que hoy lleva su nombre.

En 1939 Krumbein introduce el análisis de datos axiales, ampliamente utilizados. En 1940-1942 Salvemini desarrolla lo que hoy llamamos varianza circular.

El campo de la estadística direccional ha sido objeto de mucha atención en las últimas dos décadas debido no solamente a los requerimientos de disciplinas como la Bioinformática o la Cosmología, sino también a los avances tecnológicos y a la disponibilidad de grandes cantidades de datos, la mayoría de los cuales requieren técnicas

estadísticas adaptadas. La estadística direccional es una rama de la estadística que se ocupa de analizar observaciones (datos) que son direcciones. Este tipo de datos son típicos en disciplinas como la Meteorología (dirección del viento), Astronomía (dirección de los rayos cósmicos), Ciencias de la Tierra (localización de los epicentros de los terremotos) o Biología (ritmos circadianos, migraciones, orientaciones) o las ya mencionadas Geología y Cosmología por citar algunas. La mayoría de estas observaciones quedan representadas en una circunferencia de radio uno (estadística circular) o en una esfera también de radio uno (estadística esférica). El principal problema relacionado con este tipo de datos es la curvatura del espacio muestral ya que el círculo no es un sistema lineal. Esto implicó que cada concepto de la estadística circular debía ser reinventado a partir de la estadística clásica.

En los años 50 del siglo XX comenzaron a desarrollarse las técnicas probabilísticas apropiadas para distribuciones de datos circulares. Gumbel investigó la conveniencia de aplicar distribuciones en modelación de grupos de datos de diferente tipo. Sin embargo, el desarrollo de métodos estadísticos para analizar datos circulares se vuelven razonablemente acertados en 1956, con los trabajos pioneros de Watson y Williams. Estos autores no solamente unificaron los problemas inferenciales que planteaba von Mises, sino que también aportaron novedosas ideas y resultados. Watson introdujo el análisis de varianza, pruebas paramétricas y no paramétricas entre otras. Mardia en 1972 realizó avances en pruebas tales como la correlación y la regresión y recientemente series de análisis temporal de datos circulares.

Como hemos visto el análisis de los datos direccionales se remonta a principios del siglo XIX con autores como C. F. Gauss, pero con el trabajo de Fisher (1953) los investigadores fueron plenamente conscientes de la necesidad considerar la naturaleza curvilínea del espacio muestral. El impacto del documento de Fisher (1953) propició que se desarrollaran los estudios metodológicos y sistemáticos de los datos direccionales, teniendo en cuenta su topología real. Se propusieron y estudiaron numerosos procedimientos y distribuciones direccionales, principalmente extendiéndose a los conceptos clásicos del ajuste direccional a partir del análisis multivariante, como la estimación de puntos, los procedimientos de prueba de una y varias muestras, o la regresión.

Después de este activo periodo hubo casi una pausa en la investigación de los datos circulares hasta la década de 1990. Afortunadamente, la situación cambió desde el comienzo del nuevo milenio. Ley y Verdebout (2017) señalan tres razones principales para el resurgimiento del interés por la estadística circular. En primer lugar, el libro de Mardia y Jupp (2000) hizo que las estadísticas direccionales fueran muy populares. La facilidad de exposición, combinada con la versatilidad de los resultados presentados, atrajo el interés tanto de los estadísticos teóricos como de los profesionales. Además, un año más tarde apareció la obra de Jammalamadaka y SenGupta (2001) que se centra en las estadísticas circulares. En segundo lugar, y en parte como consecuencia lo anterior, los investigadores de las ciencias de la vida, la ecología y el aprendizaje automático, entre otros, reconocieron la importancia de la estadística circular para sus trabajos. Esto ha generado nuevas demandas y de ahí la necesidad de nuevos métodos y procedimientos. En tercer y último lugar, los avances tecnológicos han reconfigurado el campo de la estadística. El aumento exponencial de la potencia

de cálculo y de la disponibilidad de grandes cantidades de datos han traído métodos intensivos de computación y descriptivos más potentes a la vanguardia de la investigación moderna en estadística y, por lo tanto, también en estadística direccional.

El hecho de que la Estadística Circular esté siendo cada vez más utilizada se refleja en que los trabajos publicados cada año se ven incrementando poco a poco. En una consulta realizada en marzo de 2019 en Web of Science (WOS) y Scopus con un criterio de búsqueda generalizado (“Circular statistics”) arrojó un resultado de más de 1400 trabajos publicados en más de 20 disciplinas científicas (Fig. 6).

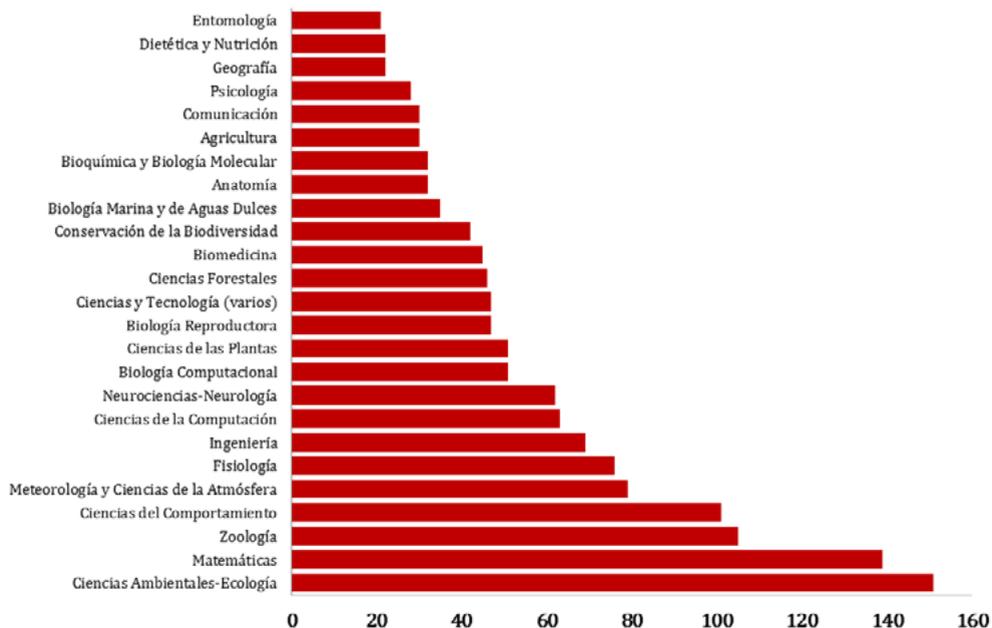


Figura 6. Trabajos publicados en el periodo 2015-2019 localizados en WOS con el criterio de búsqueda “circular statistics”.

### 1.3. OBRAS DE REFERENCIA

Como resultado de todo ello hoy en día podemos disponer de obras que permiten introducirnos en el análisis de las direcciones tales como las de Mardia (1972), Batschelet (1981), Upton y Fingleton (1989), Fisher (1993), Zar (1996), Mardia y Jupp (2000), Jammalamadaka y SenGupta (2001), Pewsey *et al.* (2013), Ley y Verdebout (2017). Las tres primeras obras, lamentablemente, están fuera de catálogo y solo pueden consultarse en librerías universitarias.

Si se desea profundizar en los fundamentos teóricos de la Estadística Circular son recomendables las obras de Mardia (1972) y especialmente la de Jammalamadaka y SenGupta (2001) para lo cual es necesario disponer de unos sólidos conocimientos matemáticos dada la naturaleza de estas obras.

Las obras de Batschelet (1981) y Fisher (1993) tienen un mayor enfoque práctico, especialmente la primera, donde se explican con detalle los fundamentos de la estadística circular acompañados de ejemplos ilustrativos. La principal ventaja de este trabajo es que todos los cálculos se pueden realizar de forma sencilla utilizando una calculadora científica o con la ayuda de una hoja de cálculo. Sin duda, para todos aquellos que deseen iniciarse en este campo de la estadística circular la obra de Batschelet debe ser la primera en consultarse. El trabajo de Fisher (1993) sigue las mismas pautas que el de Batschelet (muchos casos prácticos, posibilidad de realizar cálculos de forma sencilla), pero es un poco más complejo, lo cual no es impedimento para considerarla otra obra fundamental para iniciarse en la estadística circular.

El trabajo de Pewsey *et al.* (2013) sigue, en parte, la estructura de las obras citadas con anterioridad, pero presenta la novedad de estar enfocada para trabajar con R, un paquete estadístico bien conocido y cada vez más empleado por investigadores de todo el mundo. Para aquellos investigadores acostumbrados a trabajar en entorno R esta es, sin duda, una obra de referencia, ya que además de introducir los conceptos básicos de estadística circular aporta ejemplos de su aplicación.

La obra de Ley y Verdebout (2017) aborda la estadística direccional presentando los métodos más robustos que se aplican actualmente en este campo. Es una obra compleja que también requiere de sólidos conocimientos de estadística.

#### 1.4. PROGRAMAS INFORMÁTICOS

El avance en las ciencias de la computación también ha posibilitado que se hayan desarrollado paquetes estadísticos que permiten analizar datos circulares:

Oriana (Kovach, 2011) un programa de estadística circular, desarrollado bajo el entorno de Windows, que permite analizar orientaciones y otros datos circulares. Pueden estimarse estadísticos descriptivos, test para determinar si los datos se distribuyen de forma uniforme y si se ajustan a una distribución determinada, correlaciones, regresiones etc. Pero lo más destacable del programa es la gran disponibilidad de gráficos para representar los datos. Hay una versión de demostración, pero se necesita licencia para trabajar con él una vez finalizado el periodo de prueba gratuito.

PAST (Paleontological Statistics Software Package for Education and Data Analysis) fue desarrollado por Hammer *et al.* (2001) para el análisis de datos paleontológicos, pero su uso se ha extendido a otras disciplinas científicas. Posee un menú (Geometry) donde podemos encontrar un submenú (Directional analysis) con el que pueden realizarse análisis de datos circulares para una o dos muestras y correlaciones circulares. Es más limitado que Oriana, pero permite realizar los análisis básicos de datos circulares. Esta desarrollado bajo el entorno de Windows y es gratuito.

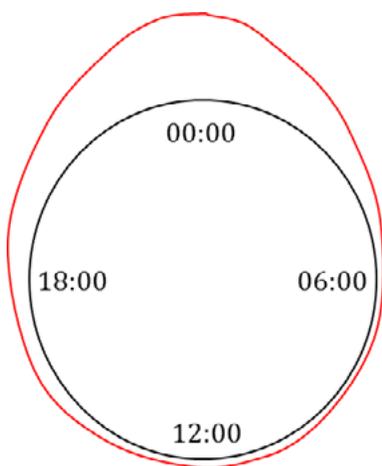
R (<[www.R-project.org](http://www.R-project.org)>) es un software libre diseñado para el análisis estadístico y la producción de gráficos. El paquete *circular* permite analizar y graficar datos circulares de forma sencilla para aquellos que estén familiarizados con R.

Otros programas que pueden utilizarse para el análisis de datos circulares son CIRCSTAT Stata (<[www.stata.com](http://www.stata.com)>) o Matlab (<[www.mathworks.com](http://www.mathworks.com)>).

## 1.5. CASOS PRÁCTICOS

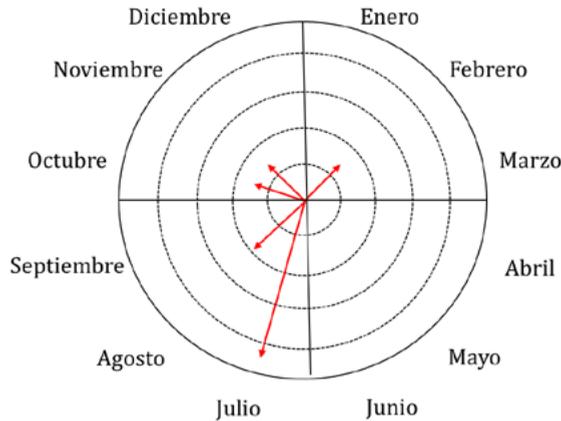
Una forma de familiarizar a los lectores con la estadística circular es describir algunos casos publicados, por tanto se presentan a continuación algunos ejemplos que, a su vez, nos permiten mostrar las diferentes áreas donde se puede aplicar la estadística circular.

**Ciencias Políticas:** Gill, J. y Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, 18: 316-336. Se trata de un estudio pionero que puso de relevancia la importancia de la estadística circular en Ciencias Políticas. Según estos autores no se había prestado restado atención a los datos circulares (puramente cíclicos) en la investigación en ciencias políticas. Estos autores demostraron que tales datos existían y que eran manejados incorrectamente por modelos que no tienen en cuenta la naturaleza cíclica de algunos fenómenos. En este estudio aplican modelos circulares a casos de terrorismo, a las elecciones y diversos tipos de sucesos como ejemplos de fenómenos cíclicos (Fig. 7).



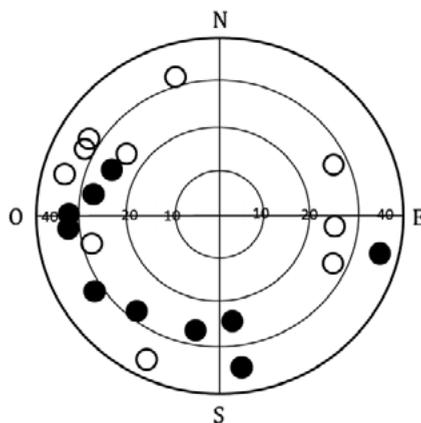
**Figura 7.** Densidad de Kernel (ver más adelante) para explicar la concentración de homicidios con arma de fuego a lo largo del día. Se observa como es hacia medias noche cuando se producen más actos violentos (modificado de Gill y Hangartner, 2010).

**Botánica:** Patricia, L.; Morellato, C.; Alberti, L.F. y Hudson, I.L. (2010). Applications of circular statistics in plant phenology: a case studies approach. En: I.L. Hudson y M.R. Keatley (eds.), *Phenological Research*: 339-359. Springer Science + Business Media. La fenología se ocupa del estudio de fenómenos biológicos recurrentes y su relación con el clima. En este trabajo los autores evalúan las aplicaciones de la estadística circular en fenología vegetal y su potencial en el análisis de datos fenológico, concluyendo que la estadística circular es una herramienta muy apropiada para determinar las relaciones entre la floración y otros procesos fenológicos (Fig. 8).



**Figura 8.** Diagrama que representa el periodo de pérdida de las hojas de un árbol, comprobándose como la época de mayor caída de hojas es en junio para luego ir decreciendo hasta febrero (modificado de Patricia *et al.*, 2010).

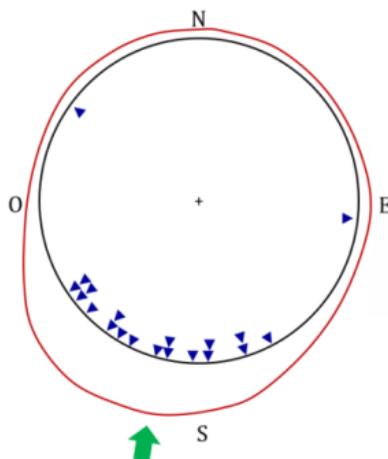
**Ornitología:** Schaaf, A.A.; García, C.G.; Puechagut, P.B.; Silvetti, L.E. y Quaglia, A.I.E. (2018). Effect of geographical latitude and sun exposure on Rufous Hornero (*Furnarius rufus*) nest orientation. *Journal of Ornithology*, 159: 967-974. La orientación de los nidos de las aves es determinada por variables ambientales, que a su vez son debidas a su posición geográfica. Las aves orientan sus nidos para evitar tales factores adversos como la exposición solar, el viento o la lluvia. Los autores utilizan la estadística circular para analizar la orientación de los nidos de *F. rufus*, en función de las variables anteriormente citadas (Fig. 9).



**Figura 9.** Orientación de los nidos de *F. rufus* (modificado de Schaaf *et al.*, 2018).

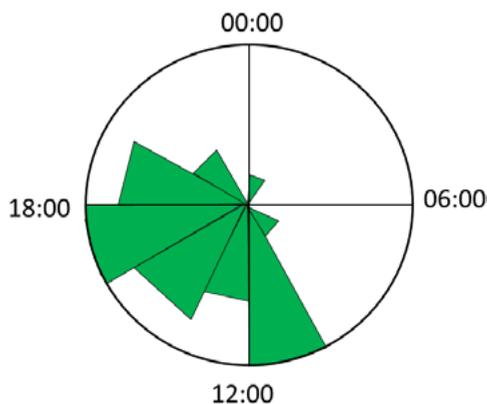
**Entomología:** Sopini, F.; Rossano, C.; Marchetti, G.M. y Morgan, E. (2005). The role of the biological clock in the sun compass orientation of free-running individuals of *Talitrus saltator*. *Animal Behaviour*, 69: 835-843. Estos autores investigan la relación

entre la orientación del sol y la actividad locomotora diaria en el saltamontes *T. saltator* utilizando modelos de regresión circular-circular, encontrando una estrecha relación entre la orientación del sol y la orientación de los saltamontes (Fig. 10).



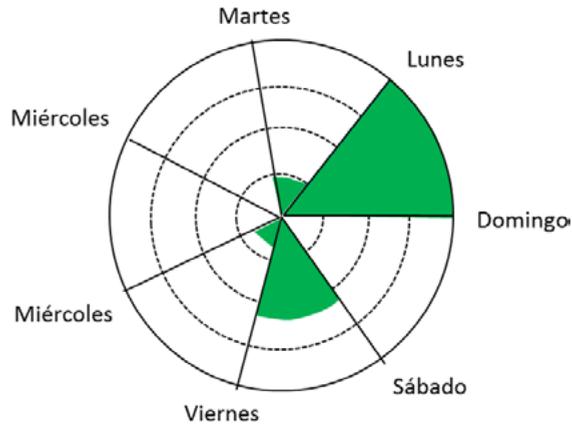
**Figura 10.** Distribución de *T. saltator* a las 13:00 horas. Se señala la posición del sol (flecha) y la densidad de Kernel (ver más adelante) (modificado de Sopini *et al.*, 2005).

**Ciencias de la Atmósfera:** Borycka, K. y Kasprzyk, I. (2018). Hourly pattern of allergenic alder and birch pollen concentrations in the air: Spatial differentiation and the effect of meteorological conditions. *Atmospheric Environment*, 182: 179-192. El objetivo de este estudio fue el de estudiar la concentración horaria del polen de dos plantas en una gran ciudad. Con la ayuda de la estadística circular determinan las concentraciones de polen horarias, concluyendo que las horas de menor exposición son las de la mañana (Fig. 11).



**Figura 11.** Concentración diaria de polen en una de las localidades estudiadas. Se observa como las mayores concentraciones de polen se producen en torno a las 12:00 y las 18:00 horas (modificado de Boryck y Kasprzyk, 2018).

**Medicina:** Karoly, Ph.J.; Goldenholz, D.M.; Freestone, D.R.; Mass, R.E.; Grayden, D.B.; Theodore, W.T. y Cook, M.J. (2018). Circadian and circaseptan rhythms in human epilepsy: a retrospective cohort study. *Lancet Neurology*, 17: 977-988. Estos autores señalan que la epilepsia se manifiesta de forma periódica y utilizan la estadística circular determinar patrones cíclicos en los brotes de la enfermedad (Fig. 12).



**Figura 12.** Histograma circular donde se representa en ciclo de temblores semanales ocasionados por la epilepsia en los pacientes estudiados. Se comprueba cómo el viernes y el domingo son los días donde se producen más temblores (modificado de Karoly *et al.*, 2018).



## CAPÍTULO 2

# LOS DATOS

### 2.1. INTRODUCCIÓN

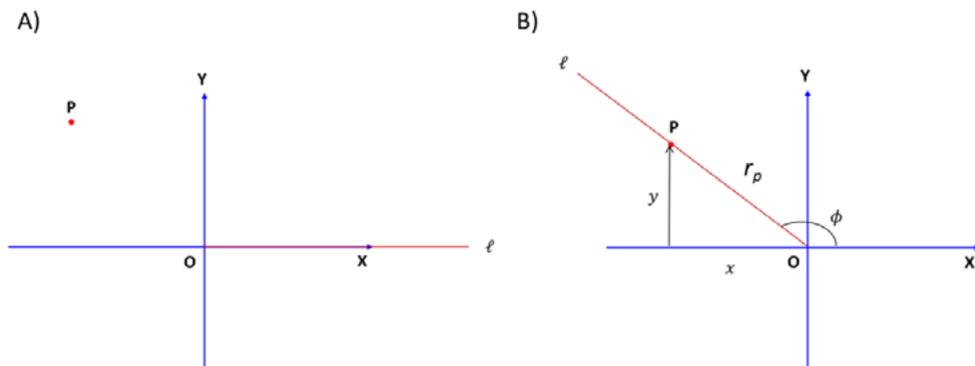
El hecho de que los datos circulares sean medidos en una escala cerrada y cíclica hace que la mayoría de los estadísticos lineales más comunes sean inapropiados para tratar este tipo de datos. En la naturaleza existen muchos fenómenos que se pueden expresar en forma de datos circulares, siendo los más frecuentes los direccionales o aquellos que determinan orientaciones. Para expresar una dirección en un plano podemos utilizar vectores unitarios o puntos sobre un círculo, como vimos en el capítulo anterior, aunque hay otras dos formas muy útiles de representar las direcciones: como ángulos o como números complejos, para lo cual será necesario establecer un sistema de coordenadas que nos permitan situar la posición respecto a este sistema de referencia.

En este capítulo se describirán los diferentes “modos” en los que podemos expresar datos circulares, pero antes se hará un breve apunte sobre ángulos, vectores y las principales funciones trigonométricas, que ayudarán en la comprensión de los procedimientos matemáticos empleados en estadística circular.

#### 2.1.1. El Ángulo

La posición de un punto  $P$  en un plano solo puede ser determinada por dos coordenadas:  $x$  y  $y$ . El sistema de *coordenadas rectangulares* o *cartesianas* con centro en  $O$  y dos ejes perpendiculares,  $X$  e  $Y$ , que pasan por  $O$ , es el que nos permite ubicar estos puntos (Fig. 13A). No obstante, la localización de  $P$  también puede caracterizarse por un ángulo  $\phi$ . Vamos a introducir una línea  $\ell$  (Fig. 13A) y asumiremos que, en principio, coincide con el eje  $X$  en su parte positiva. Asumimos también que el eje  $X$  marca la *dirección cero* y lo denominaremos *eje polar*. Ahora rotamos  $\ell$  en el sentido contrario a las agujas del reloj hasta  $P$ . A la magnitud de la rotación es lo que llamamos *ángulo*, en este caso el *ángulo polar* de  $P(\phi)$  (Fig. 13B). Para determinar la posición de  $P$  en el plano combinamos el ángulo  $\phi$  con la distancia  $r_p$  desde el origen hasta la posición de  $P$ . Ambos,  $\phi$  y  $r_p$ , conforman las *coordenadas polares* del punto  $P$ .

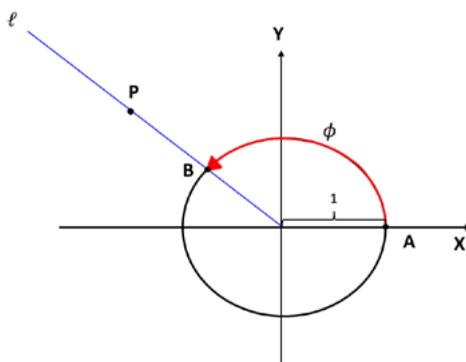
Cuando  $\ell$  coincide por primera vez con  $Y$  en la zona positiva se forma un ángulo de  $90^\circ$ . Al continuar con la rotación llegamos a la parte negativa de  $X$ , a la parte negativa de  $Y$  y, de nuevo, a la parte positiva de  $X$ : se han formado ángulos de  $180^\circ$ ,  $270^\circ$  y  $360^\circ$  y así se podría continuar formando ángulos de  $450^\circ$ ,  $540^\circ$ , etc. Si rotamos  $\ell$  en sentido de las agujas del reloj encontremos en primer lugar con  $Y$  en su parte negativa, por tanto tenemos un ángulo de  $-90^\circ$ , luego uno de  $-180^\circ$ , etc. En definitiva,



**Figura 13.** A) Para poder posicionar  $P$  en el plano necesitamos establecer un sistema de coordenadas:  $x$  e  $y$ ; B) Coordenadas polares,  $r_p$  y  $\phi$  y coordenadas rectangulares  $X$  e  $Y$  del punto  $P$ . El ángulo  $\phi$  se genera al rotar  $\ell$  en sentido contrario a las agujas del reloj.

los ángulos pueden ser negativos o positivos dependiendo del sentido de rotación. En este sentido, se denomina *azimut* al ángulo usado para medir la dirección en el plano horizontal, siempre que la dirección cero esté dirigida al norte y la rotación se haga en sentido de las agujas del reloj.

A veces es necesario cambiar la unidad en la que son medidos los ángulos para, por ejemplo, hacer representaciones gráficas o por facilitar los cálculos. Imaginemos que tenemos un círculo de radio 1 y centro en  $O$  (Fig. 14). Si, como en el caso anterior, rotamos  $\ell$  desde el origen hasta  $P$  vemos como  $A$  y  $B$  definen los puntos de intersección de  $\ell$  con el círculo unidad. Ahora podemos medir la longitud del arco que une  $A$  con  $B$  (trazo rojo en la figura 14). Esta medida de longitud es lo que se denomina *radian*. Un radián se define, por tanto, como el ángulo central de una circunferencia que abarca un arco de igual longitud que el radio de la misma. El radián, en definitiva, nos permite medir ángulos. Dado que la circunferencia de un círculo de radio  $r$  es  $2\pi r$ , el ángulo  $2\pi$  radianes, o 6,28 radianes, equivalen a los  $360^\circ$  de la circunferencia. Por tanto, los ángulos pueden ser convertidos a radianes y los radianes a ángulos, equivaliendo un radián a  $57,2958^\circ$ .



**Figura 14.** El círculo unidad y la medida de un ángulo expresada en radianes (línea roja).

2.1.2. Vectores

Los vectores están ligados a conceptos como velocidad y fuerza en los que no solo es importante la dirección sino también la magnitud. Son muy utilizados en el análisis de las direcciones.

Si tenemos dos puntos,  $Q$  y  $P$ , la línea que une  $Q$  con  $P$  es un vector y se denota por  $\overrightarrow{QP}$  o por una letra en negrita,  $v$  (Fig. 15).

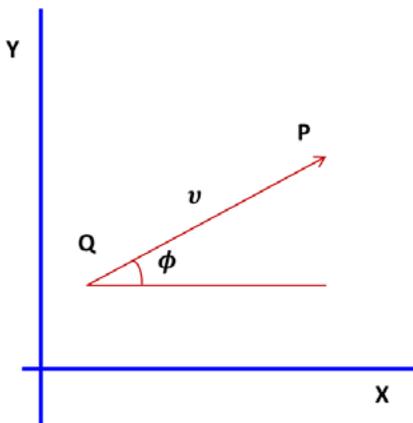


Figura 15. Representación del vector  $v$ .

En estadística circular los vectores son empleados para localizar puntos en el plano. En tal caso la base del vector  $Q$ , coincide con el origen  $O$  del sistema de coordenadas, de modo que el vector  $\overrightarrow{QP}$  puede quedar definido por sus coordenadas rectangulares  $(x, y)$  o por sus coordenadas polares  $(r_p, \phi)$  (Fig. 16):

$$\overrightarrow{QP} = v = \begin{bmatrix} x \\ y \end{bmatrix} = (r_p, \phi)$$

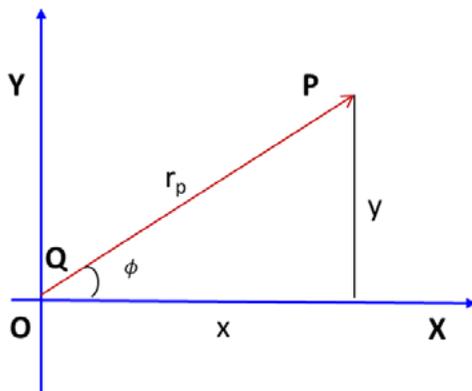


Figura 16. Coordenadas cartesianas y polares del vector  $\overrightarrow{QP}$ .

La longitud  $r$  es denominada *valor absoluto* de  $v$  y se denota:

$$r = |v|$$

El vector cero es un caso especial, ya que no se puede asociar ningún ángulo polar a este vector:

$$0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = (0, 0); \quad |0| = 0$$

En el análisis de datos direccionales es frecuente el uso de vectores unitarios, es decir, vectores que tienen una longitud igual a 1. Los vectores unitarios determinan puntos en el círculo unidad y, al contrario, cada punto del círculo unidad está asociado a un vector unitario.

Una de las operaciones algebraicas más utilizadas en estadística circular es la suma de vectores. Si tenemos dos vectores unitarios  $e_1$  y  $e_2$  la suma de los dos vectores se estima del siguiente modo:

$$e_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}; \quad e_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

$$e_1 + e_2 = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$$

También es frecuente necesitar multiplicar un vector por un número real,  $k$ :

$$k \cdot e = k \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k \cdot x \\ k \cdot y \end{bmatrix}$$

Otro importante concepto relacionado con vectores es el de *centro de masas*. Tengamos dos puntos con masa,  $M_1$  y  $M_2$  ( $M_1 > M_2$ ), localizados en el extremo de dos vectores  $e_1$  y  $e_2$ . El centro de masas,  $C$  (Fig. 17), se sitúa en la línea que une las dos masas, estando más próximo a la mayor de ellas. El vector,  $m$ , asociado al centro de masas tiene la siguiente expresión:

$$m = \frac{1}{M_1 + M_2} (M_1 \cdot e_1 + M_2 \cdot e_2)$$

Este vector  $m$  se denomina *vector medio* y puede interpretarse como la media aritmética de los vectores  $e_1$  y  $e_2$ .

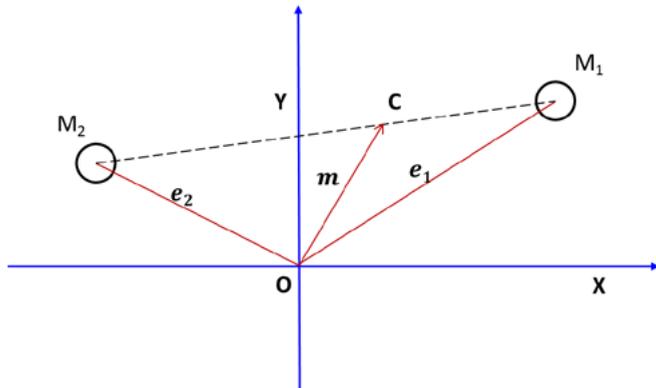


Figura 17. Representación de centro de masas (C) de dos puntos y el vector medio,  $m$ .

Del mismo modo, el vector medio de  $n$  puntos de masa  $M_1, M_2, \dots, M_n$  vendrá dado por la expresión:

$$m = \frac{1}{\sum M_i} (M_1 \cdot e_1 + M_2 \cdot e_2 + \dots + M_n \cdot e_n)$$

Si designamos  $R$  a la longitud del vector resultante y  $r$  la longitud del vector medio, tendremos que (Fig. 18):

$$|\sum e_i| = R, \quad |m| = r$$

Por tanto:

$$r = \frac{R}{n}$$

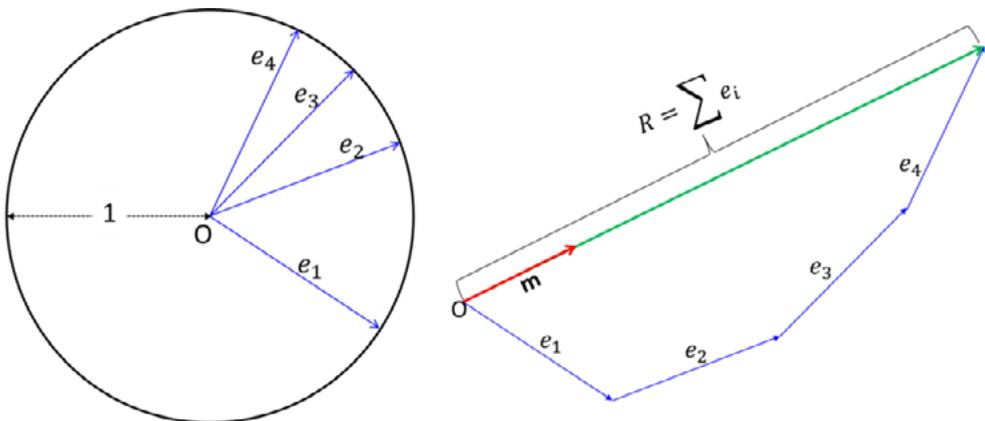


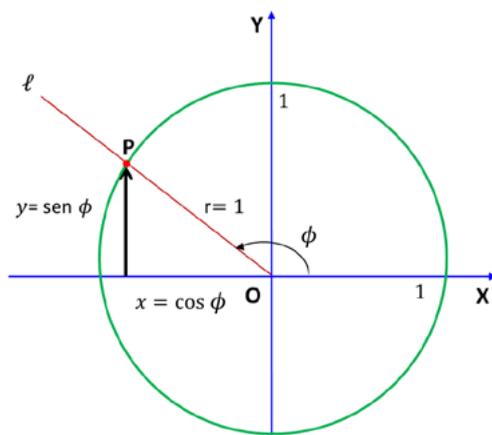
Figura 18. Suma de vectores unitarios para formar un polígono. El vector resultante ( $R$ ) es la suma de todos los vectores unitarios. Tras dividir este vector por el número de vectores unitarios obtenemos el vector medio,  $m$ .

## 2.2. FUNCIONES TRIGONOMÉTRICAS

Partamos de un sistema de coordenadas rectangulares con  $x = y = 1$ . Ahora introducimos una línea  $\ell$  para generar un ángulo  $\phi$  (Fig. 19). Introducimos también una circunferencia de radio  $r = 1$  y un punto  $P$ , cuyas coordenadas polares  $(r, \phi)$  son:  $(1, \phi)$  y con coordenadas rectangulares  $(x, y)$ .

Con estos antecedentes el coseno y el seno de  $\phi$  son:

$$\cos \phi = x; \quad \text{sen } \phi = y$$



**Figura 19.** Definición de las funciones trigonométricas seno y coseno de  $\phi$ .

Por tanto, el coseno y el seno de  $\phi$  son funciones de  $\phi$ . Por otro lado, si el dominio de  $\phi$  lo constituyen todos los números reales,  $x$  e  $y$  son funciones periódicas de  $\phi$ , dado que cualquier nueva rotación alrededor del círculo unitario genera los mismos valores de  $x$  e  $y$ . El periodo es  $360^\circ$ .

Una tercera función trigonométrica utilizada en estadística circular es la tangente:

$$\tan \phi = \frac{\text{sen } \phi}{\cos \phi}$$

Conociendo las funciones trigonométricas podemos convertir coordenadas polares  $(r, \phi)$  en coordenadas rectangulares  $(x, y)$  y viceversa:

De coordenadas polares a rectangulares. Si  $r$  es la distancia polar tendremos que:

$$x = r \cdot \cos \phi; \quad y = r \cdot \text{sen } \phi$$

De coordenadas rectangulares a polares:

$$r = (x^2 + y^2)^{1/2}$$

Si:

$$\tan \phi = \frac{\text{sen } \phi}{\text{cos } \phi}$$

Entonces:

$$\phi = \arctan (y/x) \quad \text{si } x > 0$$

$$\phi = 180^\circ + \arctan (y/x) \quad \text{si } x < 0$$

## 2.3. TIPOS DE DATOS

### 2.3.1. Ángulos

Es la forma más frecuente de representar datos circulares. El ángulo de un dato circular oscila entre  $0^\circ$  y  $360^\circ$ . Los datos circulares no pueden sobrepasar, por tanto, los  $360^\circ$  y en el caso de que esto ocurra deben ser reducidos a la escala  $360^\circ$ , restando esta magnitud a la que sobrepase de la misma. Por ejemplo,  $450^\circ$  habría que reducirlos hasta la escala de  $360^\circ$ , del siguiente modo:  $450^\circ - 360^\circ = 90^\circ$ . Los ángulos negativos, por ejemplo  $-10^\circ$ , también pueden ser convertidos en ángulos positivos, solo hay que restar a  $360^\circ$  el valor negativo. Por tanto  $-10^\circ$  equivaldrían a  $350^\circ$ .

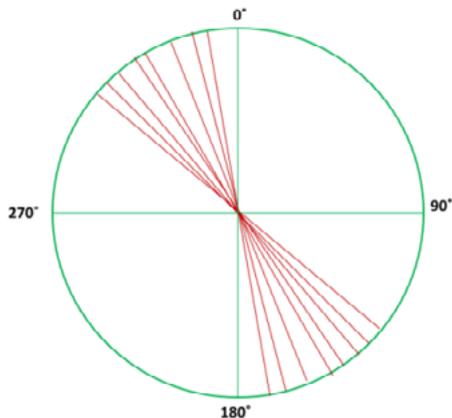
### 2.3.2. Radianes

Los datos también pueden expresarse en radianes, en un rango de entre 0 y  $2\mu$  radianes. Podemos convertir grados en radianes multiplicando por  $\pi/180^\circ$ . Por ejemplo  $45^\circ$  equivaldrían a  $45^\circ \cdot 3,141516/180^\circ = 0,7853$  radianes.

### 2.3.3. Datos axiales

Los datos que expresan direcciones pueden ser de dos tipos: uni o bidireccionales, que son conocidos como *datos vectoriales* y *datos axiales*, respectivamente. Los datos vectoriales quedan expresados por una línea direccionada, es decir, hay un sentido, como ocurre, por ejemplo, al desplazarnos en el sentido de la marcha. Los datos axiales vienen expresados por una línea en la que no se define una dirección como ocurre, por ejemplo, con una fractura horizontal en el terreno, que puede tener una dirección de  $90$  o  $270$  grados. En el caso de los datos axiales la escala varía entre  $0^\circ$  y  $180^\circ$  debido a que los datos axiales se caracterizan por poseer direcciones opuestas y equivalentes (Fig. 20). Si calculamos la media angular de los datos representado en la figura 20 obtendríamos un valor no válido, por ello es necesario reducir el módulo (escala) a  $180^\circ$  mediante la *duplicación de los ángulos*, para así obtener una muestra circular unimodal que nos permita aplicar los procedimientos estadísticos que consideremos oportunos.

**Ejercicio 1** (tomado de Mardia y Jupp, 2000). Tenemos una serie de ángulos:  $20^\circ$ ,  $135^\circ$ ,  $145^\circ$ ,  $165^\circ$ ,  $170^\circ$ ,  $200^\circ$ ,  $300^\circ$ ,  $325^\circ$ ,  $335^\circ$ ,  $350^\circ$ ,  $350^\circ$  y  $355^\circ$ , que son claramente axiales (Fig. 20).



**Figura 20.** Representación de datos axiales.

Para reducir estos datos al módulo  $180^\circ$  se procede al duplicado de los ángulos:

<i>Ángulo</i>	<i>Ángulo · 2</i>
$20^\circ$	$40^\circ$
$135^\circ$	$270^\circ$
$145^\circ$	$290^\circ$
$165^\circ$	$330^\circ$
$170^\circ$	$340^\circ$
$200^\circ$	$400^\circ$
$300^\circ$	$600^\circ$
$325^\circ$	$650^\circ$
$335^\circ$	$670^\circ$
$350^\circ$	$700^\circ$
$350^\circ$	$700^\circ$
$350^\circ$	$700^\circ$
$355^\circ$	$710^\circ$

Vemos como a partir del sexto dato los ángulos superan los  $360^\circ$ , por lo que debemos reducirlos al módulo  $360^\circ$ , para lo cual a cada uno de ellos le restamos  $360^\circ$ :

Ángulo	Ángulo · 2
20°	40°
135°	270°
145°	290°
165°	330°
170°	340°
200°	$400° - 360° = 40°$
300°	$600° - 360° = 240°$
325°	$650° - 360° = 290°$
335°	$670° - 360° = 310°$
350°	$700° - 360° = 340°$
350°	$700° - 360° = 340°$
350°	$700° - 360° = 340°$
355°	$710° - 360° = 350°$

Ahora se pueden aplicar los estadísticos circulares y comprobamos como la distribución de los datos se ha convertido en unimodal (Fig. 21).

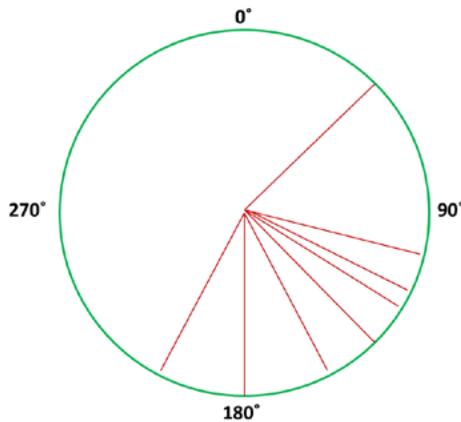


Figura 21. Ángulos doblados. Como puede apreciarse el módulo queda reducido a 180°.

### 2.3.4. Direcciones

Son datos que se representan en función de la dirección que marca la brújula. Es generalmente aceptado que 0° señale la dirección norte (N), el este (E) corresponderá a 90°, el sur (S) a 180° y el oeste (O) a 270° (Fig. 22). Este tipo de datos son muy utilizados cuando se quieren expresar direcciones u orientaciones.

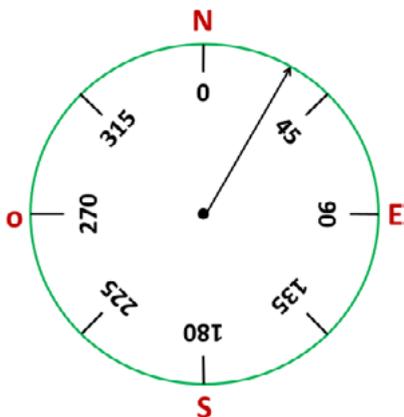


Figura 22. Representación de datos direccionales.

### 2.3.5. Tiempo en horas

Son datos circulares que se expresan en forma de tiempo en una escala de 24 horas. De este modo un día queda dividido en 24 intervalos iguales de 1 hora, que equivale a  $15^\circ$  ( $360^\circ/24 = 15^\circ$ ), de manera que  $1^\circ$  de la circunferencia equivalen a 4 minutos.

En general podemos convertir  $X$  unidades de tiempo en una dirección angular ( $a$ , en grados) empleando una escala circular con  $k$  unidades de tiempo:

$$a = \frac{360^\circ \cdot X}{k}$$

Por ejemplo ¿cuál es la medida en ángulos de las 6:25 horas?

$$a = \frac{360^\circ \cdot X}{k} = \frac{360^\circ \cdot 6,25 \text{ horas}}{24 \text{ horas}} = 93,75^\circ$$

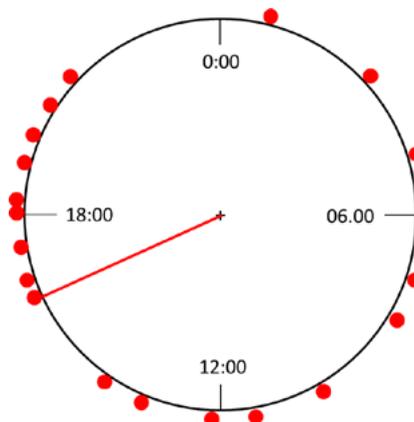
**Ejercicio 2** (tomado de Batschelet, 1981). En la siguiente tabla se presentan las horas en las que se producen los accidentes de tráfico más graves en una ciudad durante varios días:

00:56	12:08	17:24
03:08	13:28	18:08
04:52	14:16	18:16
07:16	16:20	18:56
08:08	16:44	19:32
10:00	17:04	20:52
11:24	17:20	22:08

Utilizando el conversor y con la ayuda de Excel obtenemos la siguiente tabla:

<i>Hora</i>	<i>Hora (en formato decimal)</i>	$a = \frac{360 \cdot \text{horas}}{24 \text{ horas}}$
0:56:00	0,56	8,4
3:08:00	3,08	46,2
4:52:00	4,52	67,8
7:16:00	7,16	107,4
8:08:00	8,08	121,2
10:00:00	10,00	150
11:24:00	11,24	168,6
12:08:00	12,08	181,2
13:28:00	13,28	199,2
14:16:00	14,16	212,4
16:20:00	16,20	243
16:44:00	16,44	246,6
17:04:00	17,20	258
17:20:00	17,20	258
17:24:00	17,24	258,6
18:08:00	18,08	271,2
18:16:00	18,16	272,4
18:56:00	18,56	278,4
19:32:00	19,32	289,8
20:52:00	20,52	307,8
22:08:00	22,08	331,2

Con estos datos podemos construir el gráfico (Fig. 23):



**Figura 23.** Representación de los datos del ejercicio 2. Se ha representado también el valor medio, que constituye la denominada *hora punta* o *acrofase*, que representa el momento en el que se producen más incidencias.

### 2.3.6. Días de la semana

Según el estándar internacional el primer día de la semana es el lunes. Para expresar los días de la semana de forma angular se divide el círculo ( $360^\circ$ ) por 7, por tanto cada día suponen  $25,7143^\circ$ . Las semanas completas también se pueden expresar en forma angular, considerando que el año tiene 52 o 53 semanas. Cada semana abarca  $6,9231^\circ$  del círculo unitario.

### 2.3.7. Meses

Igual que en el caso anterior, para expresar los meses de forma angular dividimos  $360^\circ$  entre 12 y obtenemos el valor que correspondería a un mes, es decir, cada sector tendría una amplitud de  $30^\circ$  (ver ejemplo Botánica capítulo 1).

### 2.3.8. Días del año

Como en casos anteriores para expresar un día en forma angular basta con dividir  $360^\circ$  entre 365 días, por lo que a cada día le corresponden  $0,9863^\circ$ . El primer día es el 1 de enero y el último el 31 de diciembre.

Podemos utilizar la fórmula anterior para determinar la correspondencia de un día cualquiera del año con su ángulo. Por ejemplo ¿Cuál es la correspondencia angular del 14 de febrero?

$$a = \frac{360^\circ \cdot X}{k} = \frac{360^\circ \cdot \text{día 45 del año}}{365 \text{ días}} = 44,38^\circ$$

### 2.3.9. Datos agrupados

El agrupamiento de datos angulares ocurre cuando la circunferencia es subdividida en arcos de igual longitud y en cada uno de esos sectores aparecen más de un dato. Por ejemplo, las direcciones mensuales del viento estimadas a diario. En cada sector, que corresponden a cada mes, aparecerán 30 o 31 medidas, siendo el número de arcos de 12. El círculo, por tanto, quedaría dividido en 12 sectores que corresponderían a los 12 meses del año.

### 2.3.10. Datos apareados

Se generan datos apareados cuando sobre un mismo individuo se hace más de una estimación. Por ejemplo, si queremos determinar cuándo se producen los picos de fiebre en los pacientes de una unidad hospitalaria. En este caso, a cada paciente se la tomarán más de una medida de temperatura al día. Estos son datos apareados, como se muestra en la siguiente tabla:

<i>Paciente</i>	<i>Temperatura 06:00 h</i>	<i>Temperatura 12:00 h</i>	<i>Temperatura 18:00 h</i>	<i>Temperatura 00:00 h</i>
1	37,5	37,6	37,9	38,2
2	36,0	38,6	37,9	38,1
3	36,7	37,5	36,9	36,2

### 2.3.11. Datos empatados

Algunos test no paramétricos requieren que no existan datos empatados entre las muestras que se van a analizar. Diremos que tenemos datos empatados cuando un mismo valor se repite en las dos muestras. Para romper el empate podemos recolocar los datos en el proceso de asignación de rangos o modificar los valores, pero esto puede generar alteraciones en los resultados.



# CAPÍTULO 3

## ESTADÍSTICA DESCRIPTIVA

### 3.1. INTRODUCCIÓN

La Estadística Descriptiva nos permite ordenar, describir y sintetizar la información recogida. En este proceso será necesario establecer medidas cuantitativas que reduzcan a un número manejable de parámetros el conjunto de datos obtenidos. La realización de gráficos (visualización de los datos en diagramas) también forma parte de la Estadística Descriptiva dado que proporcionan una manera directa de organizar la información. La finalidad de la Estadística Descriptiva no es extraer conclusiones generales sobre el fenómeno que ha producido los datos bajo estudio, sino solamente su descripción (Gorgas García *et al.*, 2011).

Entre las medidas características de una distribución destacan las llamadas medidas de centralización, que nos indicarán el valor promedio de los datos, o en torno a qué valor se distribuyen estos. Las medidas de centralización reducen la información recogida de la muestra a un solo valor. Sin embargo, dicho valor central, o medio, será más o menos representativo de los valores de la muestra dependiendo de la dispersión que las medidas individuales tengan respecto a dicho centro. Para analizar la representatividad de las medidas de centralización se definen las llamadas medidas de dispersión, que nos indicarán la variabilidad de los datos en torno a su valor promedio, es decir si se encuentran poco o muy esparcidos en torno a su centro (Gorgas García *et al.*, 2011). Finalmente, las medidas de forma nos dan una idea precisamente de la forma que adquiere la distribución de los datos, al medir el aplastamiento lateral y vertical que pueden experimentar la distribución de datos.

### 3.2. MEDIDAS DE CENTRALIZACIÓN

#### 3.2.1. Ángulo medio

El ángulo medio es una medida de tendencia central que indica la orientación preferida en una muestra de  $n$  datos angulares. Partamos de un sistema de coordenadas rectangulares con ejes  $X$  e  $Y$  y centro en  $O$ , donde  $\phi_i$  es uno de  $n$  ángulos y  $e_i$  su vector unitario. Sabemos que las coordenadas rectangulares de  $e_i$  son  $x_i$  e  $y_i$ . Por la definición de seno y coseno tendremos que:

$$x_i = \cos \phi_i; \quad y_i = \text{sen } \phi_i$$

Si denominamos  $\bar{x}$  e  $\bar{y}$  a la media de  $x_i$  e  $y_i$  (también es frecuente utilizar  $X$  e  $Y$  para designar a la media) respectivamente se cumplirá que:

$$\bar{x} = \frac{\sum_i^n \cos \phi_i}{n}; \quad \bar{y} = \frac{\sum_i^n \sen \phi_i}{n}$$

A partir de estas coordenadas se obtiene la *longitud del vector medio*,  $r$ , que nos indica el grado de concentración de los datos sobre algún punto de la circunferencia (ver más adelante):

$$r = \sqrt{\bar{x}^2 + \bar{y}^2}$$

También podemos calcular el *ángulo medio*,  $\phi$ :

$$\phi = \arctan (\bar{y}/\bar{x}), \quad \text{si } \bar{y} > 0$$

$$\phi = 180^\circ + \arctan (\bar{y}/\bar{x}), \quad \text{si } \bar{y} < 0$$

De las ecuaciones anteriores se deduce que:

$$\cos \bar{\phi} = \bar{x}/r; \quad \sen \bar{\phi} = \bar{y}/r$$

**Ejercicio 3** (tomado de Batschelet, 1981). Vamos a calcular la longitud del vector medio y el ángulo medio para una serie medidas angulares: 250°, 275°, 285°, 285°, 290°, 290°, 295°, 300°, 310°, 315°, 320°, 330°, 330° y 5°, que corresponden a la dirección que toman 15 palomas.

En primer lugar realizaremos una representación gráfica de los datos para tener una idea de su distribución. En este caso utilizaremos un diagrama circular y observamos cómo tienen una orientación hacia el noroeste (Fig. 24).

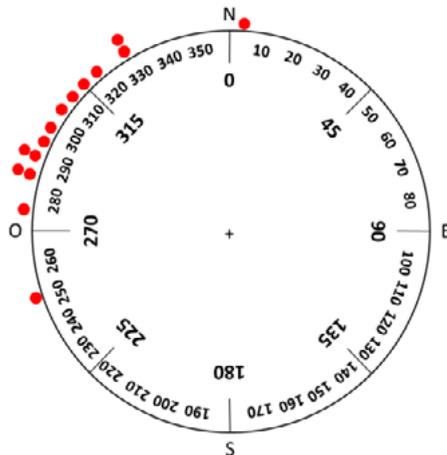


Figura 24. Direcciones que toman las 15 palomas del ejercicio 3.

A continuación lo que se debe hacer es calcular el coseno y el seno de cada ángulo, para ello utilizaremos las siguientes fórmulas de Excel: =cos(radianes(ángulo)) y =seno(radianes(ángulo)). Obtendremos la siguiente tabla:

Ángulo ( $\phi_i$ )	Coseno	Seno	Ángulo ( $\phi_i$ )	Coseno	Seno
250°	-0,3420	-0,9397	305°	0,5736	-0,8192
275°	0,0872	-0,9962	310°	0,6428	-0,7660
285°	0,2588	-0,9659	315°	0,7071	-0,7071
285°	0,2588	-0,9659	320°	0,7660	-0,6428
290°	0,3420	-0,9397	330°	0,8660	-0,5000
290°	0,3420	-0,9397	330°	0,8660	-0,5000
295°	0,4226	-0,9063	5°	0,9962	0,0872
300°	0,5000	-0,8660			

A partir de aquí obtenemos:

$$\sum \cos \phi_i = 7,2872; \quad \sum \text{sen } \phi_i = -11,3674$$

Por tanto:

$$\bar{x} = \frac{\sum_i^n \cos \phi_i}{n} = \frac{7,2872}{15} = 0,4858; \quad \bar{y} = \frac{\sum_i^n \text{sen } \phi_i}{n} = \frac{-11,3674}{15} = -0,7578$$

La longitud del vector medio es:

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} = \sqrt{(0,4858)^2 + (-0,7578)^2} = 0,9001$$

El ángulo medio es:

$$\bar{\phi} = \arctan (\bar{y}/\bar{x}) = \arctan (-0,7578/0,4858) = -53,7^\circ$$

Como  $\bar{\phi}$  es  $< 0$  calculamos el ángulo equivalente a  $-53,3^\circ$ , que es  $360^\circ - 53,7^\circ = 302,7^\circ$ . Por tanto:

$$\bar{\phi} = 302,7^\circ$$

que representa la dirección media que toman las palomas (Fig. 25).

A menudo los datos circulares están agrupados por frecuencias, de modo que las medias se calculan a partir de la siguiente fórmula:

$$\bar{x} = \frac{\sum f_i \cdot \cos \phi_i}{n}; \quad \bar{y} = \frac{\sum f_i \cdot \text{sen } \phi_i}{n}$$

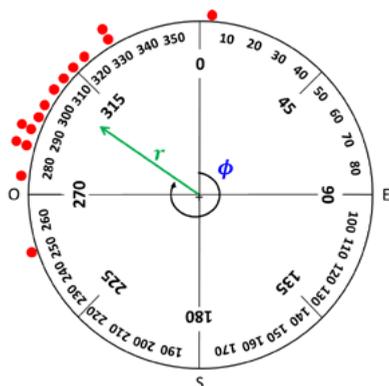


Figura 25. Direcciones de las 15 palomas del ejercicio 3 y su vector medio,  $r$ .

**Ejercicio 4** (tomado de Batschelet, 1981). En la siguiente tabla se muestran las direcciones que toman 100 moscas (se dispersaron por 12 sectores) y los cálculos necesarios para estimar  $r$  y el ángulo medio:

Ángulo ( $\phi$ )	Frecuencia ( $f$ )	Coseno	Seno	$f \cdot \text{Coseno}$	$f \cdot \text{Seno}$
20°	15	0,9397	0,3420	14,0954	5,1303
50°	15	0,6428	0,7660	9,6418	11,4907
80°	13	0,1736	0,9848	2,2574	12,8025
110°	9	-0,3420	0,9397	-3,0782	8,4572
140°	5	-0,7660	0,6428	-3,8302	3,2139
170°	7	-0,9848	0,1736	-6,8937	1,2155
200°	5	-0,9397	-0,3420	-4,6985	-1,7101
230°	1	-0,6428	-0,7660	-0,6428	-0,7660
260°	2	-0,1736	-0,9848	-0,3473	-1,9696
290°	8	0,3420	-0,9397	2,7362	-7,5175
320°	7	0,7660	-0,6428	5,3623	-4,4995
350°	13	0,9848	-0,1736	12,8025	-2,2574

Gráficamente tendríamos la siguiente representación (Fig. 27):

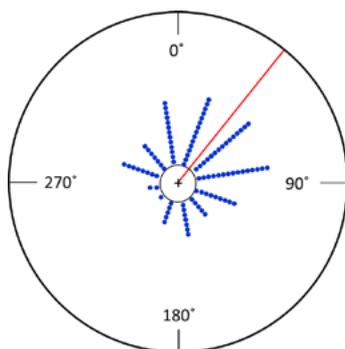


Figura 26. Representación de los datos del ejercicio 4.

Entonces tendremos que:

$$\bar{x} = \frac{\sum f_i \cdot \cos \phi_i}{n} = \frac{27,4049}{100} = 0,2740; \quad \bar{y} = \frac{\sum f_i \cdot \text{sen } \phi_i}{n} = \frac{23,5899}{100} = 0,2358$$

La longitud del vector medio es:

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} = \sqrt{(0,2740)^2 + (0,2358)^2} = 0,3616$$

El ángulo medio es:

$$\bar{\phi} = \arctan (\bar{y} / \bar{x}) = \arctan (0,2358 / 0,2740) = 40,71^\circ$$

Como ya vimos, el agrupamiento de datos angulares ocurre cuando la circunferencia es subdividida en arcos de igual longitud y en cada uno de esos sectores aparecen más de un dato. Es frecuente que cuando los datos están agrupados  $r$  tienda a ser muy pequeño por lo que necesita ser corregido. La media angular no necesita corrección. Veamos cómo se realiza la corrección. Si llamamos  $k$  al número de arcos de la circunferencia cada arco tendrá una longitud ( $\lambda$ ) que es igual a:

$$\lambda = \frac{2 \cdot \pi}{k} \text{ (radianes) o } \lambda = \frac{360^\circ}{k} \text{ (grados)}$$

El valor de  $r$  corregido ( $r_c$ ) es igual a:

$$r_c = c \cdot r$$

siendo

$$c = \frac{\lambda / 2}{\text{seno } \lambda / 2}$$

Batschelet (1981) señala que el agrupamiento de datos proporciona buenos resultados para muestras unimodales y con un buen grado de simetría. Además, indica que si el número de grupos excede de 12 la corrección tiene efectos mínimos y puede obviarse. Veamos un ejemplo.

En el ejercicio 4 obtuvimos que  $\bar{\phi} = 40,71$  y  $k = 12$ , por tanto:

$$\lambda = \frac{360^\circ}{12} = 30^\circ$$

De la siguiente tabla (Batschelet, 1981) obtenemos el valor de  $c$ :

$k$	$\lambda$ (grados)	$c$
4	90	1,1107
5	72	1,0690
6	60	1,0472
8	45	1,0262
9	40	1,0206
10	36	1,0166
12	30	1,0115
15	24	1,0073
18	20	1,0051
20	18	1,0041
24	15	1,0029
30	12	1,0018
36	10	1,0013
40	9	1,0010
45	8	1,0008
60	6	1,0005

Vemos que a  $30^\circ$  le corresponde un valor de  $c = 1,0115$ , por lo que:

$$r_c = c \cdot r = 1,0115 \cdot 0,3616 = 0,3658$$

Efectivamente, se comprueba que la corrección es mínima.

### 3.2.2. Media de los ángulos medios

Si determinamos el ángulo medio para varios grupos de ángulos obtendremos un grupo de ángulos medios. Es decir, obtendremos el ángulo medio  $\bar{\phi}$  de  $k$  muestras de datos circulares. En un momento dado puede interesar calcular la media de las medias si, por ejemplo, tenemos varias muestras de datos angulares. El cálculo se realiza de forma similar a cuando solo tenemos una muestra, pero modificando las coordenadas rectangulares que son:

$$\bar{X} = \frac{\sum_{j=1}^k r_j \cdot \cos \bar{\phi}_j}{k}; \quad \bar{Y} = \frac{\sum_{j=1}^k r_j \cdot \sen \bar{\phi}_j}{k}$$

**Ejercicio 5** (tomado de Zar, 1996). Supongamos que tenemos 7 poblaciones con 10 individuos cada una a los que se les ha determinado su dirección de dispersión ( $\phi_j$ ) y la longitud del vector medio ( $r_j$ ). Para calcular el ángulo medio de las medias procedemos del siguiente modo:

Población (j)	$\phi_j$	$r_j$	$X_j = r_j \cdot \cos \phi_j$	$Y_j = r_j \cdot \sen \phi_j$
1	160°	0,8954	-0,8414	0,3062
2	169°	0,7747	-0,7605	0,1478
3	117°	0,4696	-0,2132	0,4184
4	140°	0,8794	-0,6737	0,5653
5	186°	0,3922	-0,3901	-0,0410
6	134°	0,6952	-0,4829	0,5001
7	171°	0,3338	-0,3297	0,0522
		Suma =	-3,6914	1,9491

Entonces:

$$\bar{X} = \frac{\sum_{j=1}^7 r_j \cdot \cos \bar{\phi}_j}{7} = \frac{-3,6914}{7} = -0,5273$$

$$\bar{Y} = \frac{\sum_{j=1}^7 r_j \cdot \sen \bar{\phi}_j}{7} = \frac{1,9491}{7} = 0,2784$$

La longitud del vector medio es:

$$r = \sqrt{\bar{X}^2 + \bar{Y}^2} = \sqrt{(-0,5273)^2 + (0,2784)^2} = 0,5963$$

El ángulo pedido es:

$$\bar{\phi} = \arctan (\bar{Y} / \bar{X}) = \arctan (0,2784 / -0,5273) = -27,83^\circ$$

Como  $\bar{X} < 0$  calculamos  $180^\circ - 27,83^\circ = 152,17^\circ$  (Fig. 27).

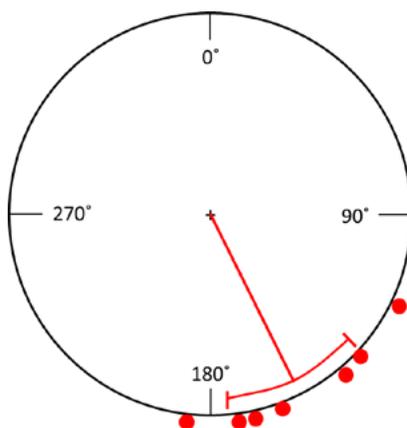


Figura 27. Representación de los datos del ejercicio 5.

### 3.2.3. Mediana angular

Según Batschelet (1981) la mediana es el ángulo ( $\tilde{\phi}$ ) que deja la mitad de los ángulos de una muestra a un lado y la otra mitad al otro lado (Fig. 28). Si  $n$  es impar la mediana coincidirá con un ángulo de la muestra, si  $n$  es par la mediana se localizará entre dos ángulos.

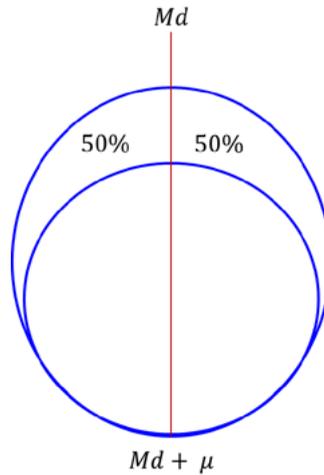


Figura 28. Representación de la mediana,  $Md$ .

En el caso del ejercicio 5 vemos que hay 15 datos en un rango de  $205^\circ$  a  $365^\circ$ . La mediana es, por tanto  $\tilde{\phi}=300^\circ$ , que es el ángulo que deja siete ángulos a la derecha y siete a la izquierda (Fig. 29).

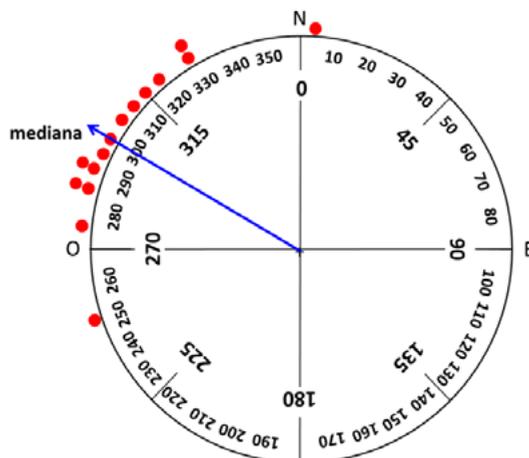


Figura 29. Representación de la mediana del ejercicio 5.

La mediana es la principal medida de centralización ya que divide la muestra en dos mitades, cada una de las cuales contiene el 50% de los datos. La mediana también es el segundo cuartil ( $Q_2$ ) y, además, el percentil 50. Recordamos que los cuartiles dividen la muestra, una vez ordenada, en cuatro partes iguales, mientras que los percentiles dividen la muestra en 100 partes. Finalmente, el *rango intercuartílico* es igual a  $Q_3 - Q_1$ .

Zar (1996) propone un estadístico para testar la simetría alrededor de la mediana, que está basado en el test no paramétrico de Wilcoxon (test pareado de Wilcoxon). Veamos cómo se calcula este estadístico.

**Ejercicio 6** (tomado de Zar, 1996). Tenemos una serie de ángulos:  $97^\circ$ ,  $104^\circ$ ,  $121^\circ$ ,  $159^\circ$ ,  $164^\circ$ ,  $172^\circ$ ,  $195^\circ$  y  $213^\circ$ , y queremos testar la siguiente hipótesis:

$H_0$ : la distribución de datos es simétrica respecto a la mediana.

$H_A$ : la distribución de datos no es simétrica respecto a la mediana.

Como  $n$  es par (8), la mediana es igual a:

$$\text{mediana} = \frac{159^\circ + 164^\circ}{2} = 161,5^\circ$$

Ahora construimos la siguiente tabla:

$X_i$	$d_i = X_i - \text{mediana}$	Rango de $ d_i $	Rango de $ d_i $ con signo
$97^\circ$	$-64,5^\circ$	8	-8
$104^\circ$	$-57,5^\circ$	7	-7
$21^\circ$	$-40,5^\circ$	5	-5
$159^\circ$	$-2,5^\circ$	1,5	-1,5
$164^\circ$	$2,5^\circ$	1,5	1,5
$172^\circ$	$10,5^\circ$	3	3
$195^\circ$	$33,5^\circ$	4	4
$213^\circ$	$51,5^\circ$	6	6

Ahora calculamos la suma de los rangos positivos y negativos:

$$T_+ = 1,5 + 3 + 4 + 6 = 14,5; \quad T_- = 8 + 7 + 5 + 1,5 = 21,5$$

Consultando las tablas para el test de Wilcoxon comprobamos que  $T_{0,05, (2), 8} = 3$ . Con este valor no podemos rechazar  $H_0$ , ya que ni  $T_+$  ni  $T_-$  son menores que  $T_{0,05, (2), 8}$ . Por tanto, la distribución es simétrica respecto a la mediana.

### 3.2.4. Moda angular

La moda angular es el ángulo que más se repite en una muestra, pudiendo haber más de una. En el caso del ejercicio 3 hay tres modas:  $285^\circ$ ,  $290^\circ$  y  $330^\circ$  (estos tres

ángulos están repetidos dos veces). En Biología, por ejemplo, lo más frecuente es la mayoría de las muestras sean unimodales (tengan un sola moda), pero también pueden ser bimodales, tri o polimodales. Un ejemplo típico de distribución bimodal ocurre cuando a un grupo de animales se le ofrece la posibilidad de elegir entre dos opciones: algunos elegirán una y otros la contraria. Si representamos las direcciones que siguen los animales observaremos dos “picos”, que corresponden a las modas.

Uno de los casos más frecuentes de distribución bimodal en Biología, por ejemplo, son los denominados datos axiales, en los que aparece una “moda” y otra idéntica en el sentido opuesto. Este tipo de muestras necesitan un tratamiento especial, como veremos más adelante.

### 3.2.5. Rango circular

El *rango circular* ( $w$ ) es la longitud del menor arco que contiene todas las observaciones que hay en una muestra. Supongamos que tenemos una serie de ángulos:  $\theta_1, \theta_2, \dots, \theta_n$  y que  $\theta_1 \leq \theta_2, \dots \leq \theta_n$ . Entonces la longitud del arco entre observaciones adyacentes es:

$$T = \theta_{(i+1)} - \theta_i, \text{ con } i = 1, 2, 3, \dots, n-1$$

por tanto:

$$T_n = 2\pi - \theta_{(n)} - \theta_1$$

Siendo el rango circular igual a:

$$w = 2\pi - \max(T_1 \dots T_n)$$

**Ejercicio 7** (tomado de Mardia y Jupp, 2000). Tenemos los siguientes ángulos ordenados:

Ángulo	$T$
43°	2
45°	7
52°	9
61°	14
75°	13
88°	0
88°	191
279°	78
357°	98

Estimamos ahora la longitud del arco entre observaciones adyacentes:  $T_1 = 45^\circ - 43^\circ = 2^\circ$ ;  $T_2 = 52^\circ - 45^\circ = 7^\circ$ ;  $T_3 = 61^\circ - 52^\circ = 9^\circ$ .....  $T_8 = 279^\circ - 88^\circ = 191^\circ$ , y comprobamos que  $T$  máximo es  $T_8$ , con un valor de  $191^\circ$ , de manera que el rango angular es:

$$w = 2\pi - \max(T_1 \dots T_n) = 2 \cdot 180^\circ - 191^\circ = 169^\circ$$

Esto significa que en un arco de  $169^\circ$  se incluyen los 9 datos de nuestra muestra (Fig. 30).

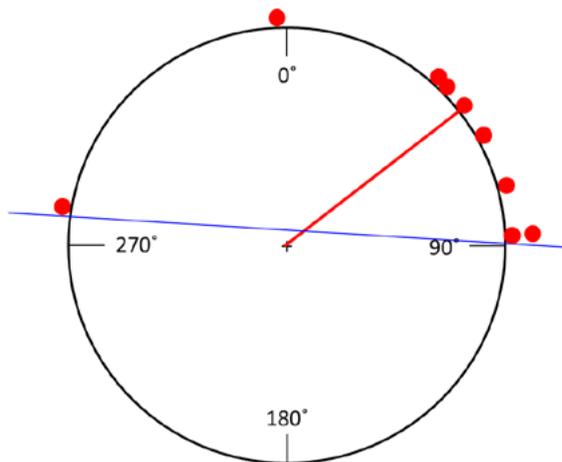


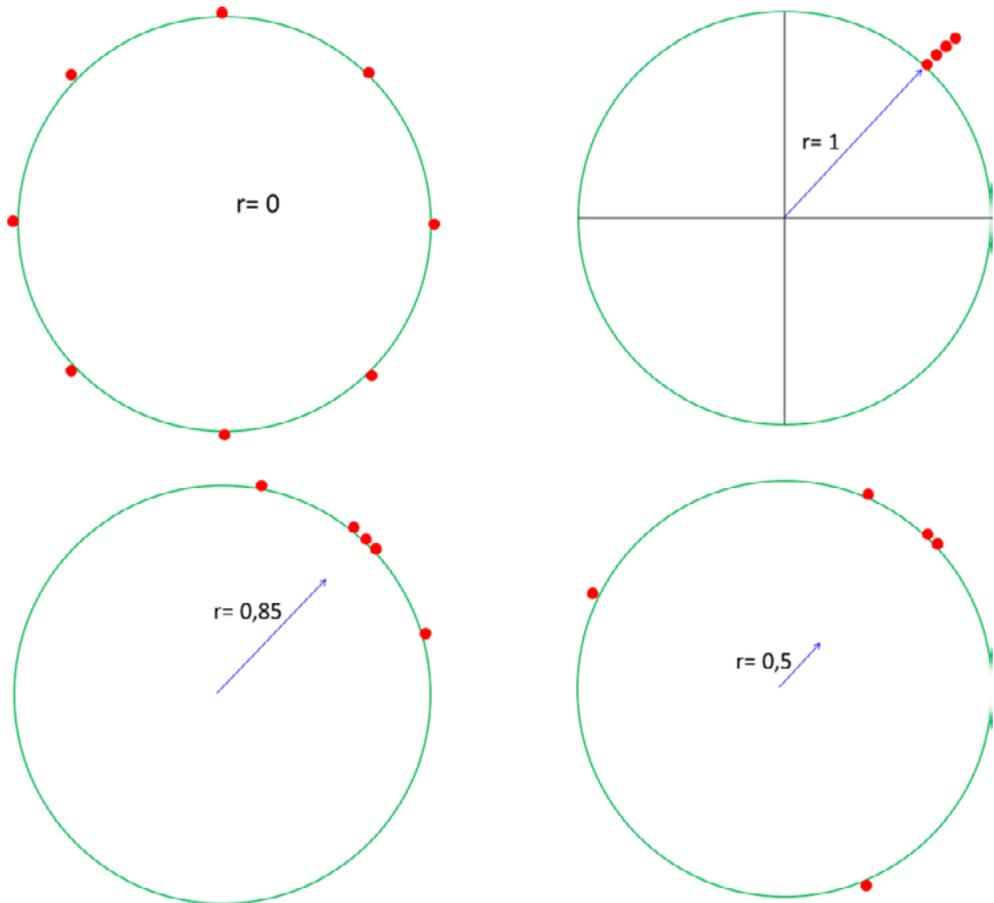
Figura 30. El arco delimitado en la zona superior de la línea azul constituye el rango circular.

### 3.3. MEDIDAS DE DISPERSIÓN

#### 3.3.1. Longitud del vector medio como medida de concentración

En muestras unimodales la longitud del vector medio ( $r$ ) sirve como medida de concentración. Una de las propiedades de  $r$  es que varía directamente con la dispersión de los datos: a mayor  $r$  mayor concentración de datos y viceversa (Fig. 31).

Por tanto un valor de  $r = 1$  implica que los datos están concentrados alrededor de un ángulo en concreto y un valor de  $r = 0$  implica que no hay concentración de los datos, aunque ello no siempre implique que estemos ante unos datos distribuidos uniformemente (ver figura 48).



**Figura 31.** La longitud del vector medio oscila entre 0 y 1, disminuyendo a medida que la concentración de puntos es menor.

### 3.3.2. Varianza angular y desviación angular

Si en la sección anterior se ha señalado que  $r$  es una medida de la concentración de los datos angulares, entonces  $1-r$  es una medida de dispersión. Este valor se utiliza para el cálculo de las principales medidas de dispersión angular, que son:

a) La varianza angular:

$$s^2 = 2 \cdot (1 - r)$$

La varianza circular por su parte tiene la siguiente expresión (Mardia y Jupp, 2000):

$$S^2 = (1 - r)$$

b) La desviación angular media (Batschelet, 1981) es:

$$s = \frac{180^\circ}{\pi} [2 \cdot (1 - r)]^{1/2}$$

c) Desviación circular estándar:

$$D(\bar{\theta}) = \frac{1}{n} \sum_{i=1}^n (\pi - |\pi - |\theta_i - \bar{\theta}||)$$

### 3.3.3. Distancia angular

Zar (1996) define *distancia angular* como la distancia más corta, en ángulos, entre dos puntos situados sobre un círculo. De este modo, la distancia angular entre 95° y 120° son 25°. Nos referiremos a la distancia angular entre los ángulos  $\alpha_1$  y  $\alpha_2$  como  $d_{\alpha_1-\alpha_2}$ . Así,  $d_{95^\circ-120^\circ}$ . La distancia angular es muy útil cuando deseamos obtener, por ejemplo, la dispersión de los datos respecto a una dirección determinada, como sucede cuando queremos saber la desviación de una serie de animales entrenados para volver siguiendo una determinada dirección (*homing*).

Mardia y Jupp (2000) proponen que la distancia entre dos ángulos  $\theta$  y  $\xi$  es igual a:

$$1 - \cos(\theta - \xi)$$

De modo que un modo de medir la dispersión de una serie de ángulos  $\theta_1, \theta_2, \dots, \theta_n$  respecto a un ángulo dado  $\alpha$  es:

$$D(\alpha) = \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \alpha)\}$$

### 3.3.4. Límites de confianza para la media

Los límites de confianza para la media ( $L_1, L_2$ ) pueden expresarse como:

$$\bar{\theta} \pm d$$

Donde  $d$  varía en función de  $r$ . Para  $r \leq 0,9$ :

$$d = \arccos \left[ \frac{\sqrt{\frac{2n(2R^2 - n\chi_{\alpha,1}^2)}{4n - \chi_{\alpha,1}^2}}}{R} \right]$$

Donde  $\chi_{\alpha,1}^2 = \chi_{0,05,1}^2 = 3,481$ . Se suele tomar  $\alpha = 0,05$ .

Para  $r \geq 0,9$ :

$$d = \arccos \left[ \frac{\sqrt{n^2 - (n^2 - R^2)e^{\chi_{\alpha,1}^2/n}}}{R} \right]$$

Siendo  $R = n \cdot r$ .

**Ejercicio 8** (tomado de Zar, 1996). Tenemos una muestra de 8 datos angulares en la que:  $\bar{\phi} = 104^\circ$ ,  $r = 0,8791$ , por tanto:

$$R = n \cdot r = 8 \cdot 0,87919 = 7,03352$$

De modo que:

$$d = \arccos \left[ \frac{\sqrt{\frac{2n(2R^2 - n\chi_{\alpha,1}^2)}{4n - \chi_{\alpha,1}^2}}}{R} \right] = \arccos \left[ \frac{\sqrt{\frac{2 \cdot 8(2 \cdot (7,03352)^2 - 8 \cdot 3,481)}{4 \cdot 8 - 3,481}}}{7,03352} \right] = 28^\circ$$

De modo que el intervalo de confianza al 95% es:  $104^\circ \pm 28^\circ$ . Por consiguiente,  $L_1 = 76^\circ$  y  $L_2 = 132^\circ$ .

### 3.4. MEDIDAS DE FORMA

#### 3.4.1. Asimetría

La asimetría mide el sesgo de la distribución, es decir, si existe desplazamiento de los datos hacia uno de los lados del valor medio. Se calcula mediante la siguiente expresión:

$$\hat{r} = \frac{r_2 \cdot \text{sen}(\bar{\theta}_2 - 2\bar{\theta})}{(1-r)^{3/2}}$$

Donde  $r_2$  es el la longitud del vector medio de la muestra una vez se han duplicado los ángulos;  $\bar{\theta}_2$  es el ángulo medio de la muestra con los ángulos duplicados y  $r$  y  $\theta$  son, respectivamente, los vectores medio y el ángulo medio de la muestra original.

#### 3.4.2. Curtosis

La curtosis expresa el grado de aplastamiento de una distribución simétrica respecto al que correspondería a una distribución normal. La curtosis tiene la siguiente expresión:

$$\hat{K} = \frac{r_2 \cdot \cos(\bar{\theta}_2 - 2\bar{\theta}) - r^4}{(1-r)^2}$$

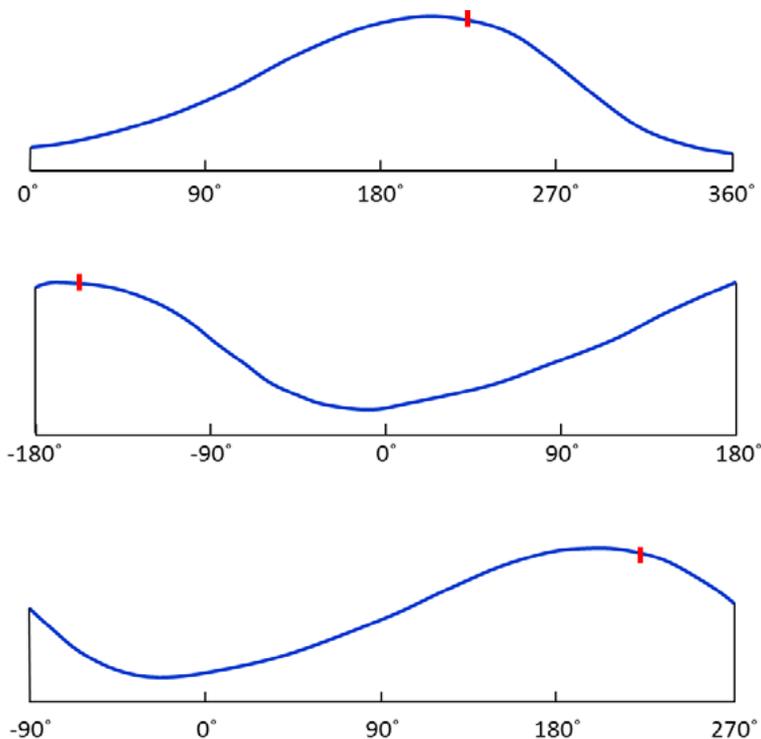
### 3.5. REPRESENTACIÓN DE LOS DATOS

#### 3.5.1. Introducción

En Estadística es muy recomendable realizar una representación gráfica de los datos que conforman una muestra antes de realizar cualquier tratamiento estadístico. Las ventajas que ofrece una representación gráfica son varias:

1. Nos da una idea inicial sobre algunas características importantes de la muestra. Por ejemplo, si sigue una distribución uniforme, se ajusta a una distribución de von Mises o es polimodal.
2. Nos sugiere un modelo al cual se pueden ajustar nuestros datos.
3. Permite detectar valores anómalos.

Una idea intuitiva de una distribución circular es un gráfico circular. También es frecuente emplear gráficos lineales, pero este tipo de representaciones tiene el inconveniente de que el punto de inicio es arbitrario y ello puede generar confusión (Fig. 32).



**Figura 32.** Tres gráficos lineales de la misma distribución. La distribución es unimodal, con moda en  $200^\circ$  (o de forma equivalente  $-160^\circ$ ). Los puntos de inicio de los gráficos son  $0^\circ$  (superior),  $-180^\circ$  (centro) y  $-90^\circ$  (inferior).

Disponemos de varias opciones para representar los datos, pudiendo elegir la que más nos interese en función de la naturaleza de los datos que deseamos representar.

### 3.5.2. Gráficos de distribución

Este tipo de gráfico nos permite determinar si nuestros datos se ajustan a una determinada distribución. El fundamento es sencillo: cada dato es representado frente al valor esperado bajo una distribución teórica dada (uniforme o von Mises). Si los datos proceden de una distribución dada deben aparecer alineados en una línea de  $45^\circ$ , que representan los valores teóricos. La desviación de los puntos respecto de esta línea supondría que los datos observados no se ajustarían a la distribución elegida. En la figura 33 se representa la orientación de los nidos de un insecto y deseamos saber si los datos se ajustan a una distribución de von Mises, que representa la distribución "normal" en estadística circular. Como se puede apreciar hay un buen ajuste de los datos a la línea de  $45^\circ$ , sugiriendo que los nidos se ajustan a una distribución de von Mises.

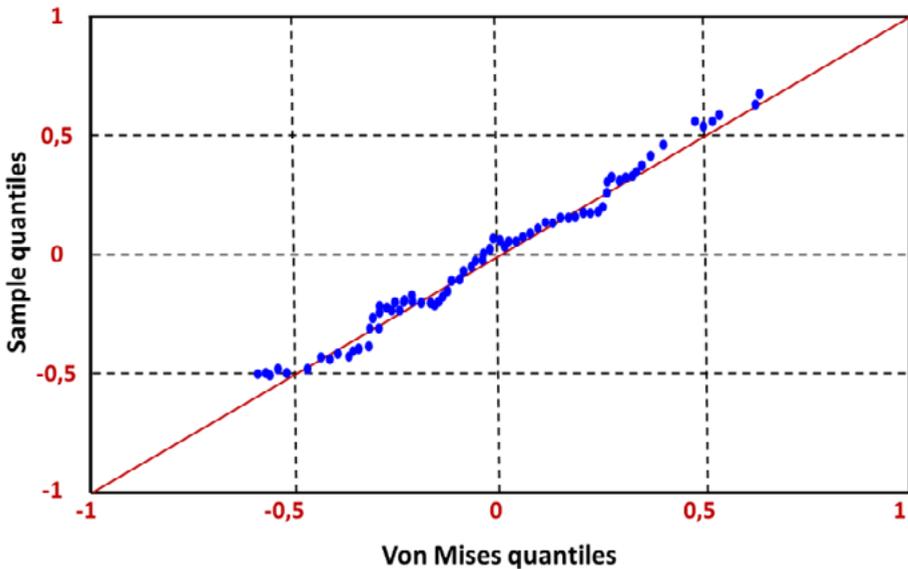


Figura 33. Gráfico de distribución de nidos de insectos.

En la figura 34 podemos comprobar como, por el contrario, la orientación de los nidos de insectos no se ajustan a una distribución uniforme (los puntos no se alinean con la línea de referencia).

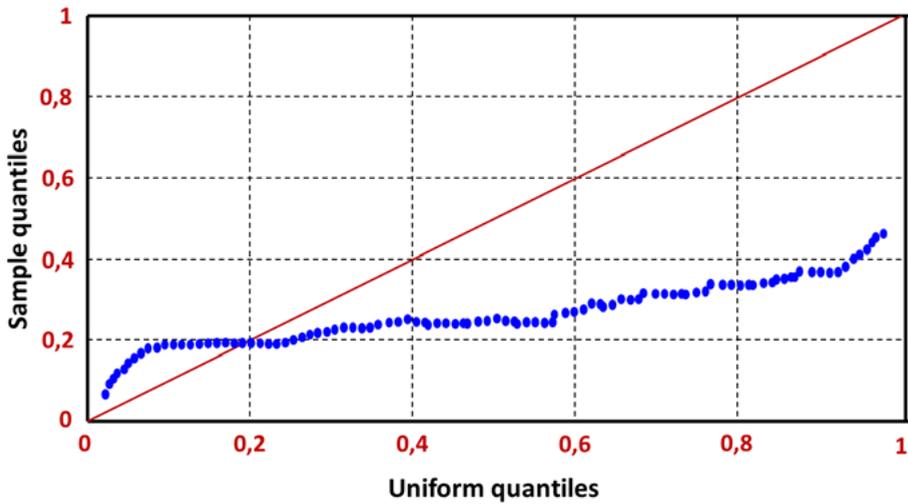


Figura 34. Gráfico de distribución de nidos de insectos (realizado con Oriana).

El gráfico Q-Q nos permite comparar la distribución de dos muestras. Si poseen distribuciones similares los puntos estarán próximos a la línea de referencia de 45°. En la figura 35 comparamos la orientación de nidos de insectos en dos años consecutivos. Podemos comprobar que los puntos se ajustan a la línea de referencia, por lo que se concluye que las dos muestras tienen distribuciones similares.

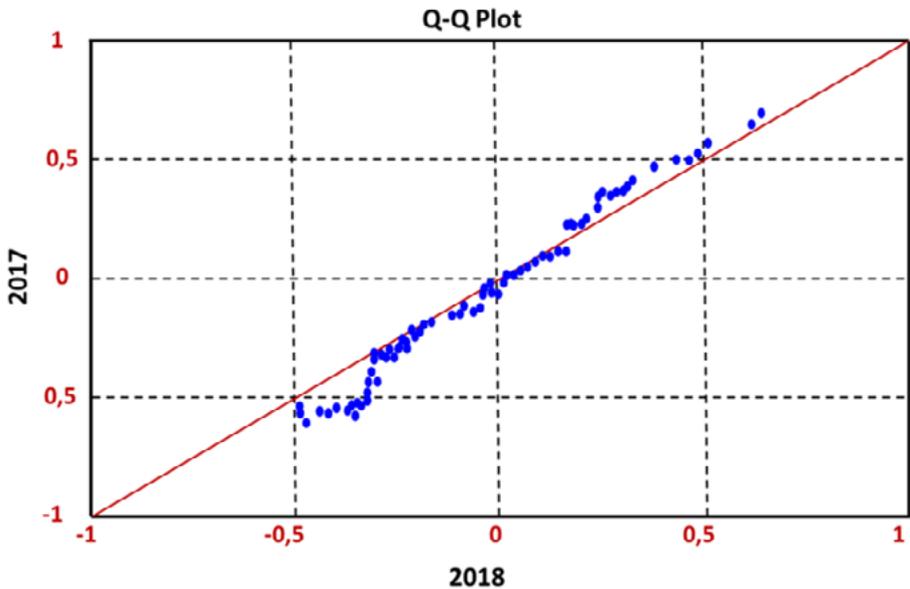


Figura 35. Gráfico Q-Q para datos de distribución de nidos de insectos en dos años consecutivos. Se comprueba que en los dos años los datos tienen distribuciones similares.

### 3.5.3. Gráfico de dispersión

Como sucede en estadística lineal este gráfico permite representar puntos (Fig. 36). En este caso se han representado la dirección del viento (en radianes) en Olivenza (Badajoz) durante 150 días.

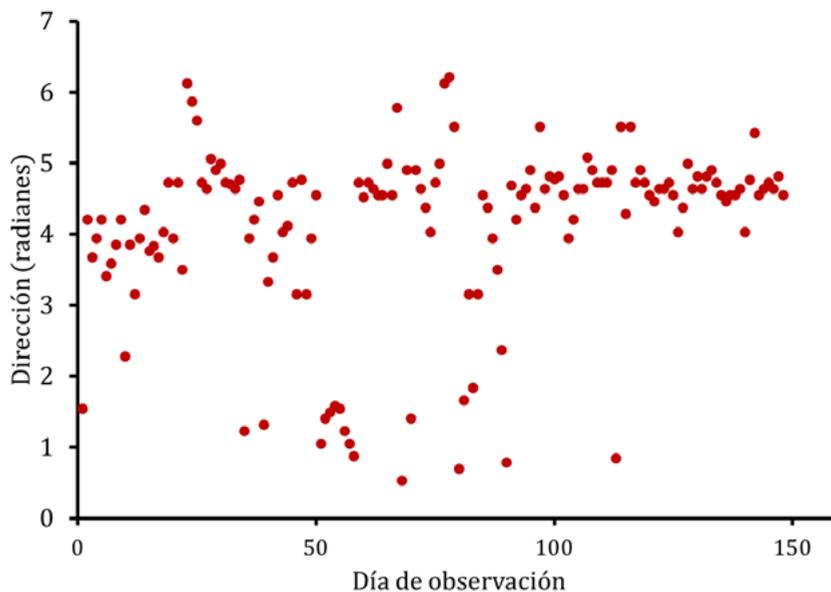


Figura 36. Dirección del viento en Olivenza (Badajoz).

Se puede comprobar como los datos se concentran en torno a los 3-5 radianes, que corresponden a una dirección este. A pesar de ser muy ilustrativos, este tipo de gráficos no es una buena opción para resumir los datos, ya que no refleja la periodicidad de los mismos, un aspecto muy importante en estadística circular para determinados tipos de datos.

### 3.5.4. Histogramas lineales

Se construyen considerando los datos como si fuesen lineales y no angulares, por lo que se grafican en un sistema de dos ejes (x, y) (Fig. 37). El punto de partida es  $0^\circ$  si utilizamos una escala de  $0^\circ$  a  $360^\circ$  o  $180^\circ$  si utilizamos una escala de  $-180^\circ$  a  $+180^\circ$ . A continuación elegimos un intervalo para el agrupamiento ( $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ) y procedemos de igual manera que en los histogramas lineales. Se pueden realizar en Excel.

Cuando generamos un histograma lineal de datos circulares lo que en realidad hacemos es “cortar” el círculo en un punto determinado y asociar los extremos con dos puntos del histograma lineal. La interpretación del histograma, en este caso, se ve dificultada porque perdemos la periodicidad de los datos. Además, la “zona de corte”

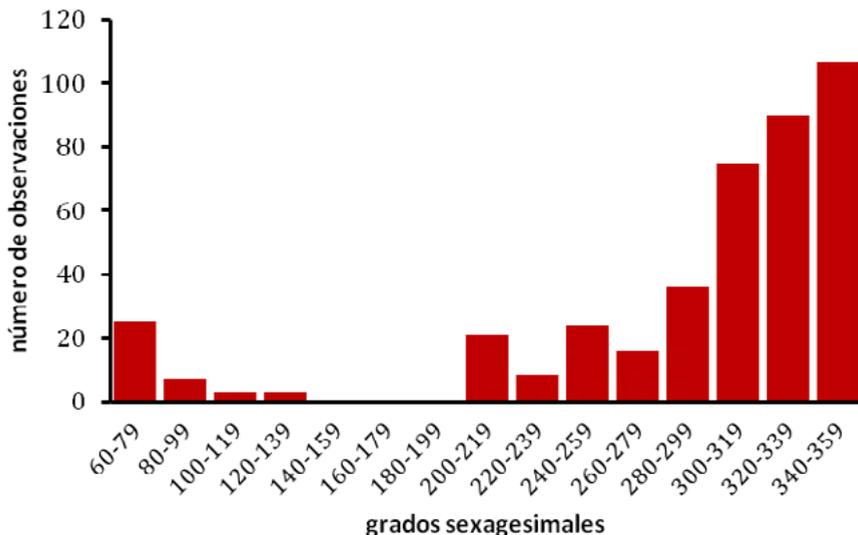


Figura 37. Histograma lineal de datos circulares. En el eje de abscisas se representan intervalos de ángulos.

sobre el círculo también puede generar problemas. Pewsey *et al.* (2013) ilustran este caso de forma muy elegante. Los dos histogramas de la figura 38 representan 310 direcciones de viento sobre un mismo lugar, pero los puntos de “corte” son diferentes. Del gráfico de la izquierda podríamos deducir que existen dos modas, mientras que en el de la derecha solo se apreciaría una moda, por tanto, es muy importante precisar bien el punto de corte para no obtener conclusiones erróneas. En el caso que nos ocupa la opción correcta es la que refleja el segundo histograma.

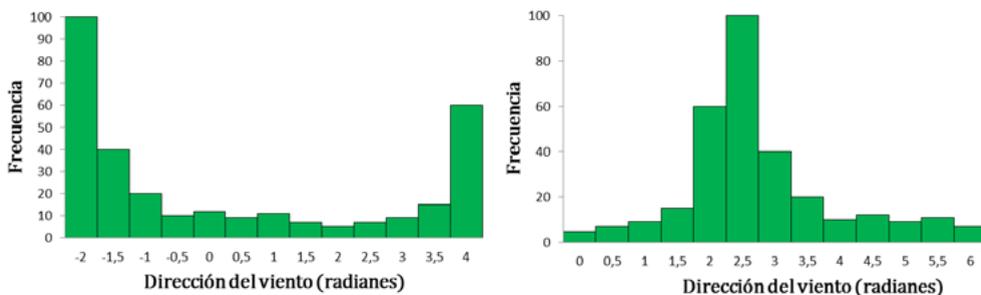


Figura 38. Dos histogramas realizados con los mismos datos pero con diferentes puntos de corte.

Para evitar los errores derivados de la elección del punto de corte muchos autores recomiendan duplicar el histograma (Fig. 39), pero esto puede generar confusión si no se advierte que el histograma ha sido duplicado.

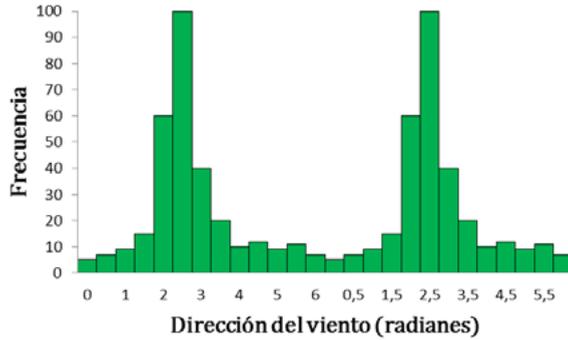


Figura 39. Histograma duplicado. Se comprueba cómo, efectivamente, solo hay una moda.

### 3.5.5. Gráfico de datos originales

Es la forma más sencilla de presentar los datos. Los datos son representados mediante círculos, triángulos o líneas que muestren la orientación de los mismos (desde el origen hasta el dato angular) (Fig. 40). Estos diagramas pueden realizarse sin necesidad de utilizar programas informáticos complejos. Una forma sencilla de realizarlos es utilizando powerpoint.

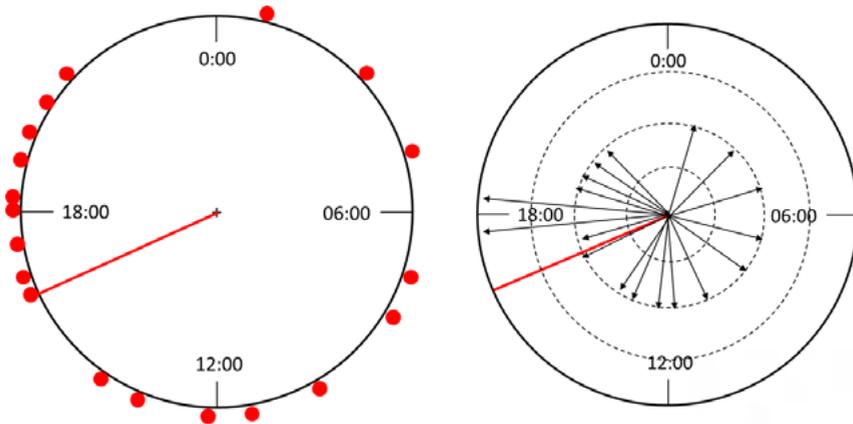


Figura 40. Representación de datos circulares por puntos y por radios.

Este tipo de gráficos nos permite tener una primera impresión sobre la distribución de los datos. En muestras grandes cada punto puede representar más de una observación. Se puede elegir entre varias opciones para realizarlos (Fig. 41).

Otra forma de representar los datos es la que se presenta en la figura 42. A la izquierda podemos ver un diagrama circular en la que el radio del círculo representa el mayor número de unidades de actividad. Podemos suponer, por ejemplo, que una serie de animales toman una dirección dada, siendo 47 el mayor grupo que

sigue una determinada dirección (izquierda). A la derecha se representa la misma información que a la izquierda, pero en forma de polígono relleno.

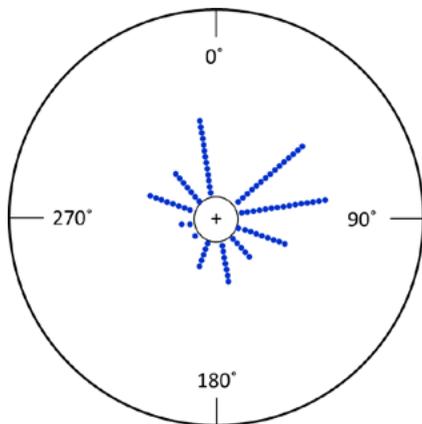


Figura 41. Gráfico de datos originales con origen en el centro.

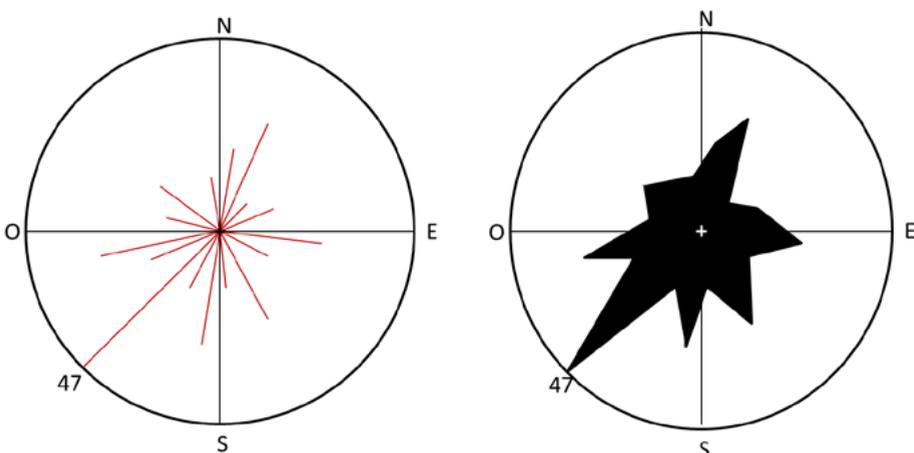


Figura 42. Dos gráficos circulares para representar direcciones. En ambos casos el radio del círculo lo determina el grupo de mayor actividad.

### 3.5.6. Histogramas circulares (o angulares)

Se construyen con barras paralelas que muestran el número de observaciones dentro de cada clase (Fig. 43). Cada barra de un histograma circular está centrada en el punto medio del correspondiente grupo de ángulos y su área es proporcional a la frecuencia de cada grupo. La anchura de las barras representa un intervalo de grados y deben definirse al principio, por ejemplo 10°, 20°, etc. Para construirlos es necesario utilizar programas específicos (R, Oriana, PAST, etc.).

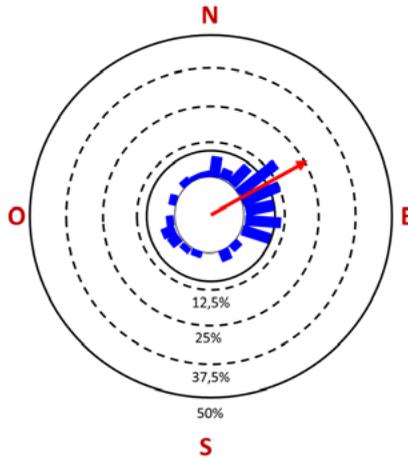


Figura 43. Histograma circular. Cada barra representan 20°.

### 3.5.7. Diagramas de rosa

Básicamente es un histograma circular en el que las barras han sido sustituidas por sectores (Fig. 44). Cada sección del histograma, que en este caso tienen forma de cuña, muestra el número de observaciones que aparecen dentro de ese rango. El área de cada sector es proporcional a la frecuencia del grupo correspondiente, mientras que el radio del sector es proporcional a la raíz cuadrada de la frecuencia relativa de cada grupo. Para construirlos es necesario utilizar programas específicos (R, Oriana, PAST, etc.). Es importante destacar que la elección del número de segmentos es crucial para una buena interpretación del diagrama de rosas. Como regla general podemos señalar que la raíz cuadrada del tamaño de la muestra es el mejor indicador para elegir el número de segmentos.

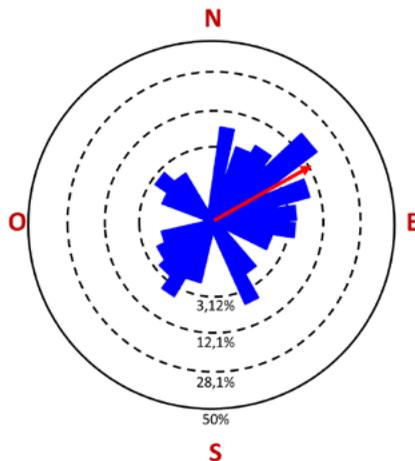


Figura 44. Diagrama de rosa

### 3.5.8. Densidad de Kernel

La densidad de Kernel es una forma cada vez más popular de representar gráficamente datos circulares. La densidad de Kernel es, en definitiva, una estimación no paramétrica de la densidad muestral, determinando la influencia de cada punto es su entorno próximo. Según Fisher (1993) el estimador de Kernel se ajusta a la siguiente expresión:

$$\hat{f}(\theta) = \frac{1}{n h} \sum_{i=1}^n k\left(\frac{\theta - \theta_i}{h}\right)$$

Siendo  $n$  el tamaño poblacional,  $h$  la ventana o banda, que representa el área de influencia que se pretende dar a cada valor de la muestra y  $k$  el parámetro de Kernel (es una función). Su elección se suele hacer depender de la dispersión de los valores: la desviación estándar en caso de datos lineales o el factor de concentración ( $\kappa$ ) para datos circulares. Uno de los principales problemas a la hora de estimar la densidad de Kernel con datos circulares es, precisamente, establecer al ancho de banda, es decir, el área de influencia que damos a cada observación. Queda lejos del objetivo de esta obra explicar los procedimientos de selección del ancho de banda por lo que se recomienda consultar a Oliveira *et al.* (2012) para profundiza en este tema. De igual modo la estima de la densidad de Kernel requiera de programas informáticos especializados (R, PAST, etc.).

En la figura 45 se muestra un gráfico circular en la que se refleja la densidad de Kernel (trazo negro). Como se puede apreciar donde los puntos están más concentrados suele formarse un “pico” mayor.

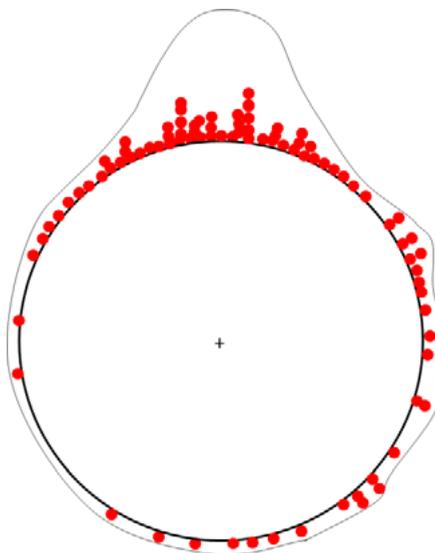


Figura 45. Densidad de Kernel (línea continua).

### 3.5.9. Diagramas de tallo y hoja

Básicamente se trata de un histograma de fácil construcción, donde cada tallo correspondería a cada intervalo (por ejemplo  $10^\circ$ ) y las hojas a cada uno de los datos. Se construyen de forma lineal y luego se “pega” a una circunferencia. Se pueden realizar en Excel y luego “ajustarlo” a una circunferencia en powerpoint. De la figura 46 podemos deducir, por ejemplo, que en el intervalo de  $30^\circ$  hay cuatro datos:  $30^\circ$ ,  $34^\circ$ ,  $34^\circ$  y  $38^\circ$ . De igual modo en el intervalo de  $110^\circ$  hay dos datos: un valor de  $113^\circ$  y otro de  $118^\circ$ .

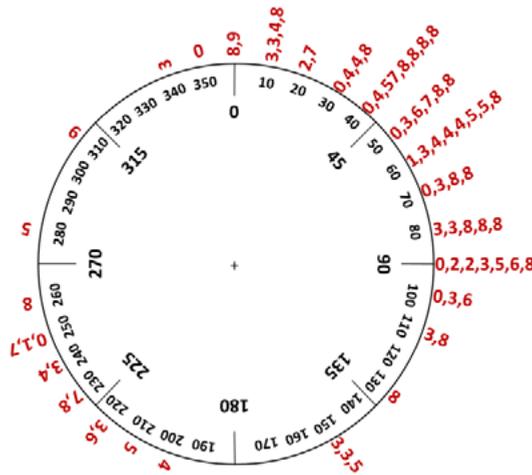


Figura 46. Diagrama de tallos y hojas de 100 datos angulares.

## 3.6. DISTRIBUCIONES CIRCULARES

### 3.6.1. Introducción

En estadística circular, al igual que ocurren en estadística lineal, es complicado trabajar con grandes cantidades de datos. Por ello, antes de nada, debemos establecer algún sistema que nos permita organizarlos u ordenarlos de algún modo. Lo más básico suele ser ordenarlos y plasmarlos en una tabla de recuento. Básicamente con este proceso lo que obtenemos es su *distribución de frecuencias*. Por tanto las distribuciones nos dan una idea de cómo se van a comportar nuestros datos o, por decirlo de otro modo, la “forma” que tendrán si hacemos una representación de los mismos: en forma de línea recta, en forma de campana, etc.

A modo de introducción podemos decir que en estadística circular vamos a encontrar dos tipos de distribuciones básicas: la uniforme y la de von Mises. Una distribución uniforme sería aquella en la que los datos se distribuyen de forma más o menos equidistante a lo largo del círculo, como lo hacen los radios de la rueda de una bicicleta, por ejemplo. La distribución de von Mises, por el contrario, tiene forma acampanada siendo muy similar a la distribución normal que conocemos en

estadística lineal, es decir, los datos suelen estar concentrados en algún punto de la circunferencia. Muchos autores consideran a la distribución de von Mises como la “distribución normal” de la estadística circular, de ahí su importancia. Esto supone que los datos están “orientados” hacia algún lugar o se concentran hacia un determinado punto. La pregunta que deberíamos plantearnos ahora es porque los datos se concentran hacia un punto determinado o que causa provoca la concentración. Explicaremos esto con un sencillo experimento. En la figura 47A hemos representado la orientación que adquieren 12 ramas de un árbol de una determinada especie que crece a nivel del mar, y en la figura 47B se representa el crecimiento de 15 ramas del mismo árbol desarrollándose a 3000 m de altitud en la costa norte (se ha marcado la dirección de los vientos predominantes en la zona).

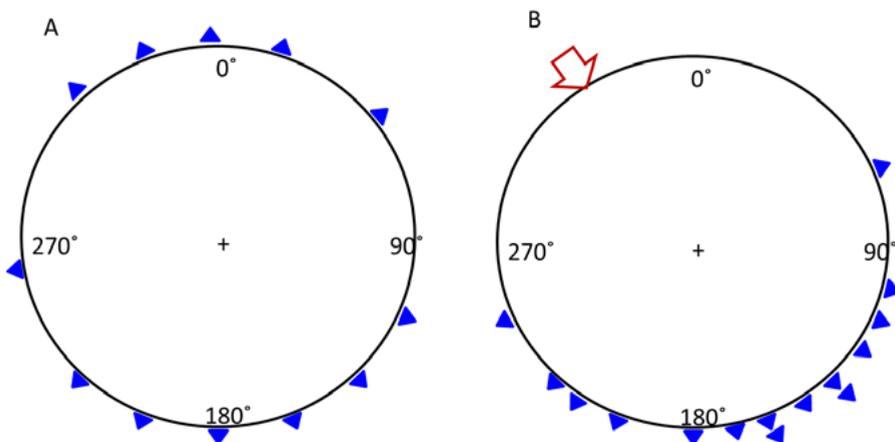


Figura 47. Distribución uniforme (A) y “orientada” (B).

En el primer caso podemos concluir que las ramas se distribuyen de forma más o menos uniforme alrededor del tronco, mientras que en el segundo caso se deduce que hay mayor número de ramas creciendo hacia el sureste: hay direccionalidad. El paso siguiente sería explicar a que se debe esa direccionalidad ¿a la dirección del viento?, ¿a la fuerza del viento?, ¿a ambos?

Una distribución circular es una distribución de probabilidad en la que la probabilidad total está concentrada en una circunferencia de radio 1. Dado que cada punto de la circunferencia representa una dirección, la distribución circular es una forma de asignar probabilidades a las diferentes direcciones o definir una distribución direccional. El rango de una variable circular aleatoria, medido en radianes, queda comprendido, por tanto, entre  $[0, 2\pi]$  o  $[-\pi, \pi]$ .

Ya se ha visto que existen varias formas de representar los datos circulares. Sin embargo, una forma de especificar una distribución en el círculo unidad es por medio de su función de distribución y por su función de densidad de probabilidad. La función de densidad de probabilidad es una función que permite determinar la probabilidad de que una variable aleatoria continua tome un valor entre cierto intervalo.

Así, para el caso de una variable circular continua la función de probabilidad de densidad  $f(\theta)$  cumple los siguientes requisitos:

- (i)  $f(\theta) \geq 0$ . Es siempre positiva.
- (ii)  $\int_0^{2\pi} f(\theta) d\theta = 1$
- (iii)  $f(\theta) = f(\theta + k \cdot 2\pi)$  para cualquier valor de  $k$ , por tanto,  $f$  es periódica.

La función de distribución  $F(\theta)$  correspondiente a  $f(\theta)$  puede definirse sobre cualquier intervalo  $(\theta_1 - \theta_2)$  de la siguiente manera:

$$F(\theta_2) - F(\theta_1) = \int_{\theta_1}^{\theta_2} f(\theta) d(\theta)$$

En la mayoría de los casos utilizaremos distribuciones que se inician en un punto convenido como  $0^\circ$  (o  $-180^\circ$ ) y definir  $F(\theta)$  del siguiente modo:

$$F(\theta) = \int_0^{\theta} f(\theta) d(\theta)$$

Cumpléndose que:  $F(2\pi) = 1$ .

La función de distribución se utiliza para estimar los momentos trigonométricos, tales como la dirección media, la varianza circular o los cuartiles, que ya han sido descritos.

En estadística lineal conocemos una serie de distribuciones cuyo rango puede ser finito (distribución binomial) o infinito (distribución normal, Poisson). Las distribuciones circulares son siempre finitas y continuas. A la hora de representarlas se suele utilizar representaciones lineales o circulares.

Vimos en el capítulo anterior algunos conceptos básicos sobre vectores, que recordaremos ahora para introducir las distribuciones circulares. Vamos a volver a expresar los momentos circulares pero, en este caso, utilizaremos la nomenclatura que podremos encontrar en la mayoría de los textos abordan la estadística circular.

Sabemos que el mejor modo de obtener la dirección media de una serie de direcciones es tratar los datos como vectores unitarios. Supongamos que tenemos una muestra de  $n$  datos  $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ . Entonces:

$$C = \sum_{i=1}^n \cos \theta_i, S = \sum_{i=1}^n \sen \theta_i, R^2 = C^2 + S^2, (R \geq 0)$$

La dirección media del vector resultante,  $\bar{\theta}$  de  $\theta_1, \theta_2, \theta_3, \dots, \theta_n$  viene dada por:

$$\cos \bar{\theta} = \frac{C}{R}; \quad \sen \bar{\theta} = \frac{S}{R}$$

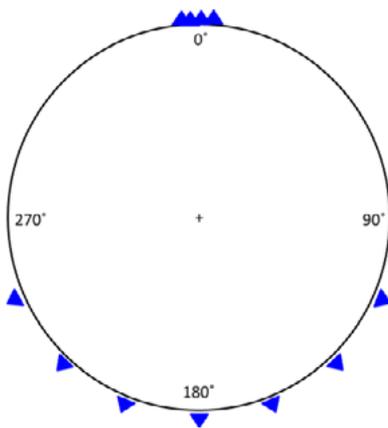
O

$$\bar{\theta} = \begin{cases} \arctan \frac{S}{C}, & \text{si } S > 0, C > 0 \\ \arctan \frac{S}{C} + \pi, & \text{si } C < 0 \\ \arctan \frac{S}{C} + 2\pi, & \text{si } C < 0, S < 0 \end{cases}$$

Donde  $\bar{\theta}$  es la *dirección media*.  $R$  es la *longitud del vector resultante* y oscila entre 0 y  $n$ . La *longitud resultante media*,  $\bar{R}$  (también es frecuente que se designe con la letra griega  $\rho$ ,  $\rho$ ), se calcula del siguiente modo:

$$\bar{R} = \frac{R}{n}$$

y oscila entre 0 y 1.  $\bar{R} = 1$  significa que todos los puntos coinciden, pero  $\bar{R} = 0$  no implica que exista una distribución uniforme, como ocurre en la figura 48.



**Figura 48.** Para esta serie de datos la longitud resultante media toma un valor de cero, pero como se puede observar la distribución no es uniforme, por eso este parámetro no es un buen estimador para expresar la concentración o falta de concentración de los datos.

La *varianza circular* ( $V$ ) se ajusta a la expresión:

$$V = 1 - \bar{R}$$

Al igual que ocurre en la estadística lineal, un elevado valor de la varianza circular implica una elevada concentración de los datos, sin embargo y, a diferencia de los que ocurre en estadística lineal, en estadística circular se cumple que  $0 \leq V \leq 1$ . Al igual que ocurre con  $\bar{R}$ , un valor de  $V = 0$  no implica necesariamente una distribución dispersada al máximo.

La *desviación circular estándar* ( $v$ ) se define como:

$$v = \{-2 \log(1 - V)\}^{1/2} = (-2 \log \bar{R})^{1/2}, \quad \in [0, \infty]$$

La *dispersión circular* se ajusta a la siguiente expresión:

$$\hat{\delta} = \frac{(1 - \bar{R}_2)}{2\bar{R}^2}$$

Donde  $\bar{R}_2$  es la *longitud resultante media* de los ángulos doblados.

Finalmente la *asimetría* ( $\hat{s}$ ) y la *curtosis* ( $\hat{k}$ ) adoptan las siguientes expresiones:

$$\hat{s} = \frac{\bar{R}_2 \text{sen}(\bar{\theta}_2 - 2\bar{\theta})}{(1 - \bar{R})^{3/2}}$$

$$\hat{k} = \frac{\bar{R}_2 \text{cos}(\bar{\theta}_2 - 2\bar{\theta}) - \bar{R}^4}{(1 - \bar{R})^2}$$

### 3.6.2. Distribución uniforme

Una distribución circular uniforme es aquella en la que todos los valores tienen la misma probabilidad de ser elegidos, o lo que es lo mismo, ningún sector del círculo es preferido sobre otros. La distribución de probabilidad sobre un círculo con estas propiedades es lo que se denomina distribución circular uniforme o, simplemente, distribución uniforme. En esta distribución no existe una dirección media y  $r = 0$  (Fig. 49). Es el modelo más básico de distribución sobre el círculo y a menudo es utilizada como modelo nulo.

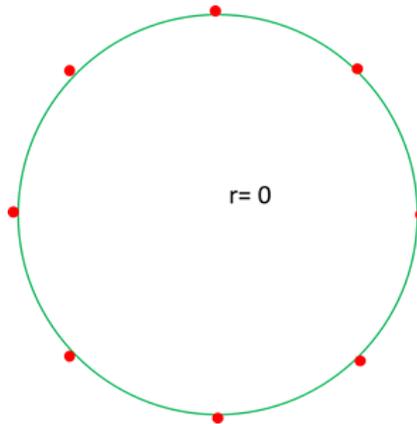


Figura 49. Puntos distribuidos uniformemente en el círculo.

La función de densidad para esta distribución es (Fig. 50):

$$f(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta < 2\pi$$

Integrando obtenemos su función de distribución:

$$F(\theta) = \frac{\theta}{2\pi}, \quad 0 \leq \theta < 2\pi$$

Momentos:

Dirección media:  $\mu$ , indefinida.

Longitud resultante media,  $\bar{R} = 0$ .

Dispersión circular,  $\delta = \infty$ .

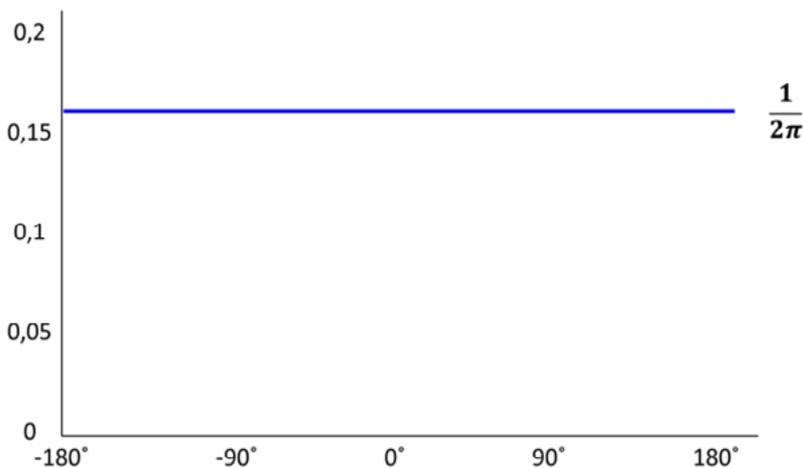


Figura 50. Función de densidad de la distribución uniforme.

### 3.6.3. Distribución de von Mises

La distribución de von Mises equivaldría, en estadística lineal, a la distribución normal por eso es tan importante en estadística circular. Se trata de una distribución unimodal y simétrica con función de probabilidad:

$$f(\varphi) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\varphi - \theta_1)]$$

Con:

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos \theta} d\theta$$

Donde  $\kappa$  es el *parámetro de concentración* y  $\theta_1$  el *ángulo promedio*. La función alcanza su valor máximo cuando  $\varphi = \theta_1$ , por tanto  $\theta_1$  es la moda.

Como ya se ha señalado esta distribución es unimodal y simétrica con respecto a  $\theta = \mu$  (Fig. 51). La moda se encuentra en  $\theta = \mu$  y la antimoda en  $\theta = \mu + \pi$ . El cociente entre el valor de la densidad en la moda y el valor de la densidad en la antimoda es  $e^{2\kappa}$ , así que cuanto mayor sea el valor de  $\kappa$ , mayor será la concentración alrededor de la moda.

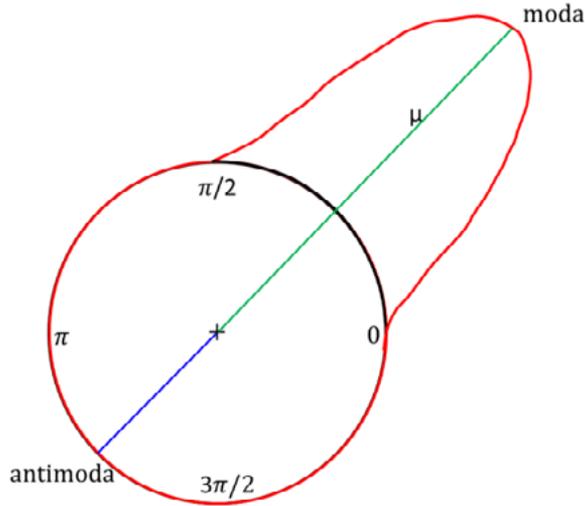


Figura 51. Distribución de von Mises.

Para  $\kappa = 0$  la distribución de von Mises degenera en una distribución uniforme. Cuanto mayor sea  $\kappa$  mayor concentración de datos hay alrededor del ángulo medio (Fig. 52).

El cálculo del parámetro de concentración resulta normalmente complicado, por ello se han propuesto varias aproximaciones para poder estimarlo:

$$\bar{R} < 0,53 \quad \kappa \approx 2\bar{R} + \bar{R}^3 + \frac{5}{6}\bar{R}^5$$

$$0,53 \leq \bar{R} < 0,85 \quad \kappa \approx -0,4 + 1,39\bar{R} + \frac{0,43}{1-\bar{R}}$$

$$\bar{R} > 0,85 \quad \kappa = \frac{1}{2(1-\bar{R}) + (1-\bar{R})^2 - (1-\bar{R})^3}$$

$$\bar{R} > 0,9 \quad \kappa = \frac{1}{2(1-\bar{R})}$$

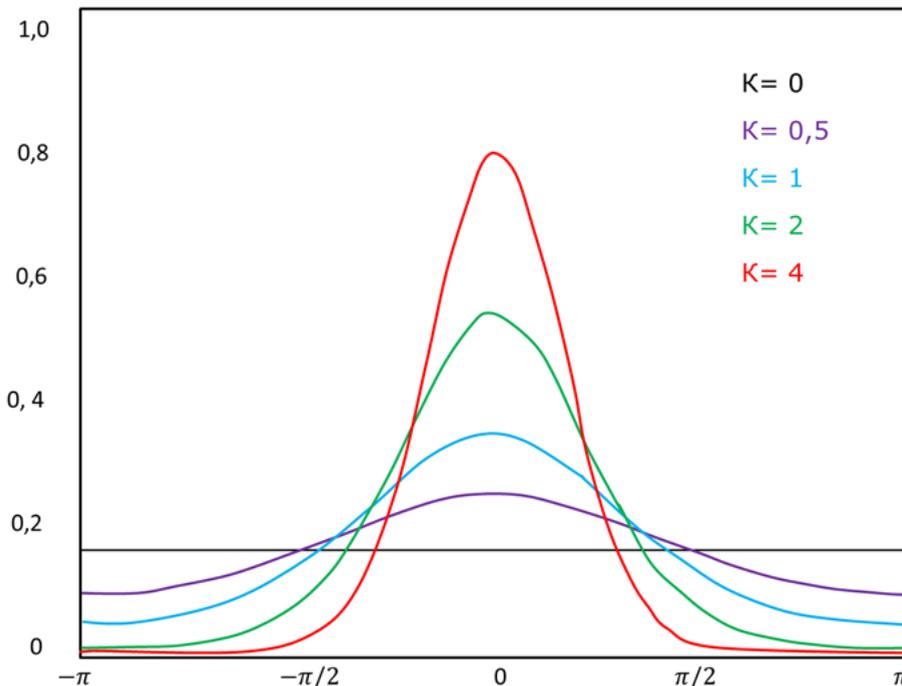


Figura 52. Distribuciones de von Mises con distintos parámetros de concentración.

En la práctica, podemos encontrar la distribución de von Mises que se ajusta a nuestros datos. En proceso no es complicado, ya que basta con una hoja de cálculo para poder completarlo. Vamos a encontrar la distribución de von Mises que corresponde a una muestra.

**Ejercicio 9** (tomado de Batschelet, 1981). Partimos de una muestra ( $n = 714$ ) de direcciones angulares que corresponden a las que siguen 714 aves.

Sabemos que la función de densidad de la distribución de von Mises es igual a:

$$f(\varphi) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\varphi - \theta_1)]$$

Siendo  $\theta_1$  el ángulo promedio de la población. Con los datos de la muestra calculamos:

$$\begin{aligned} \bar{x} &= 0,5021; & \bar{y} &= 0,5125; \\ r &= 0,7176; & \bar{\varphi} &= 45,6^\circ; \end{aligned}$$

Asumimos que:

$$\bar{\varphi} = \theta_1 = 45,6^\circ$$

Calcularemos ahora el valor de  $\tilde{\kappa}$ , para lo cual utilizaremos la tabla L (Anexo I), donde comprobamos que para un valor de  $r = 0,7176$  y  $n = 714$  obtenemos (por aproximación) un valor de  $\tilde{\kappa} = 2,129$ . Con estos datos tenemos que:

$$\varphi = \frac{1}{2\pi I_0(\kappa)} \exp^{[\kappa \cos(\varphi - \theta_1)]} = \frac{1}{2\pi I_0(\kappa)} \exp^{[2,129 \cos(\varphi - 45,6^\circ)]}$$

Ahora calcularemos el valor de  $I_0(\kappa)$  a partir de la siguiente tabla (Batschelet, 1981):

$\kappa$	$I_0(\kappa)$
0,0	1,00000
0,1	1,00250
0,2	1,01003
0,3	1,02263
0,4	1,04040
0,5	1,06348
0,6	1,09204
0,7	1,12630
0,8	1,16652
0,9	1,21299
1,0	1,26607
1,5	1,64672
2,0	2,27959
2,5	3,28984
3,0	4,88079
3,5	7,37820
4,0	11,30192
4,5	17,48117
5,0	27,23987

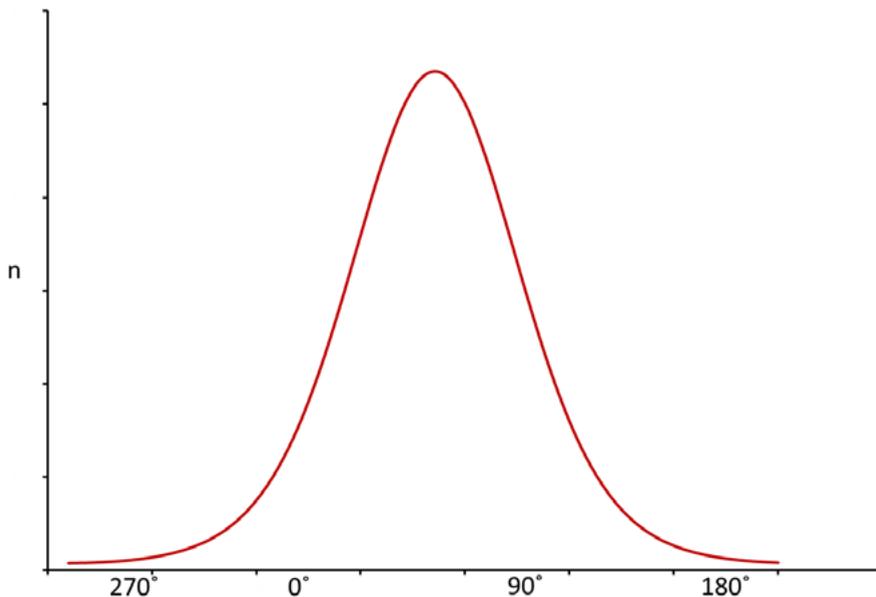
Así:

$$I_0(\kappa) = 2,42$$

Por tanto:

$$(\varphi) = 0,065 \exp^{[2,129 \cos(\varphi - 45,6^\circ)]}$$

Ahora en una hoja de cálculo dando valores a  $\varphi$  obtenemos el siguiente gráfico:



Como se puede apreciar la curva se ajusta perfectamente a una distribución de von Mises. Podemos afirmar, por tanto que los datos de nuestro ejemplo se ajustan a una distribución de von Mises, o lo que es lo mismo, nuestra muestra se ha obtenido de una población que se ajusta al modelo de von Mises (normal).

### 3.6.4. Distribución cardioide

Se trata de una función simétrica cuya función de distribución es:

$$f(\theta) = \frac{1}{2\pi} \{1 + 2\bar{R} \cos(\theta - \mu)\}, \quad 0 \leq \theta < 2\pi, \quad 0 \leq \bar{R} \leq 1/2$$

Siendo su función de distribución:

$$F(\theta) = \frac{\bar{R}}{\pi} \text{sen}(\theta - \mu) + \theta / 2\pi, \quad 0 \leq \theta < 2\pi$$

Momentos:

Dirección media:  $\mu$

Longitud resultante media,  $\bar{R} < 1/2$

Dispersión circular,  $\delta = \frac{1}{2\bar{R}^2}$

La distribución es simétrica y unimodal en  $\mu$  (si  $\bar{R} > 0$ ). Si  $\bar{R} = 0$  la distribución se reduce a la distribución uniforme.

### 3.6.5. Distribución wrapped

Dada una distribución en la recta real, es posible *envolverla* alrededor de la circunferencia del círculo de radio uno. Si  $X$  es una variable aleatoria en la línea, la variable aleatoria correspondiente  $X_w$  de la distribución wrapped viene dada por:

$$X_w = X(\text{mod } 2\pi)$$

Si el círculo es identificado con el conjunto de números complejos de módulo 1, entonces la función que envuelve se puede escribir como:

$$X \rightarrow e^{iX}$$

Si  $X$  tiene función de distribución  $F$ , la función de distribución  $F_w$  de  $X_w$  viene dada por:

$$F_w(\theta) = \sum_{k=-\infty}^{\infty} [F(\theta + 2\pi k) - F(2\pi k)], \quad 0 \leq \theta \leq 2\pi$$

En particular, si la variable aleatoria  $X$  tiene función de densidad  $f$ , la correspondiente función de densidad  $f_w$  de  $X_w$  es:

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} [f(\theta + 2\pi k)]$$

### 3.6.6. Distribución wrapped normal

La distribución wrapped normal  $WN(\mu, \bar{R})$  se obtiene *envolviendo* en el círculo una distribución  $N(\mu, \sigma^2)$ , donde:

$$\sigma^2 = -2 \log \bar{R}$$

Es decir:

$$\bar{R} = e^{-\sigma^2/2}$$

Resulta entonces, que la función de densidad de una distribución  $WN(\mu, \bar{R})$  es:

$$\varphi_w(\theta; \mu, \bar{R}) = \frac{1}{\sigma \sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{-\frac{(\theta - \mu + 2\pi k)^2}{2\sigma^2}}$$

# CAPÍTULO 4

## INFERENCIA ESTADÍSTICA

### 4.1. INTRODUCCIÓN

La Inferencia Estadística se ocupa de estudiar los métodos necesarios para extraer, o inferir, conclusiones válidas e información sobre una *población* a partir del estudio experimental de una muestra de dicha población. Los métodos utilizados en la Inferencia Estadística dependen de la información previa que se tenga de la población a estudiar. Cuando se conoce la forma de la distribución de probabilidad que sigue la variable aleatoria a estudiar en la población, el problema consiste en determinar los diferentes parámetros de dicha distribución (ej. media y varianza para la distribución normal). Para ello se utilizan los métodos paramétricos, consistentes en procedimientos óptimos para encontrar dichos parámetros. Cuando la distribución de la población es desconocida, el problema principal es encontrar la forma y características de la distribución, lo cual se hace mediante los llamados métodos no paramétricos (Gorgas García *et al.*, 2011). En este capítulo nos limitaremos a estudiar los principales métodos de inferencia en estadística circular.

Antes de continuar recordaremos algunos conceptos útiles para este tema. Definiremos *población* como el conjunto de individuos sobre el cual estamos interesados en extraer conclusiones. Una *muestra*, por tanto, es un subconjunto de la población al que tenemos acceso y sobre el que realizamos las observaciones. Un *parámetro* es la cantidad numérica calculada sobre la población, mientras que un *estadístico* es una cantidad numérica calculada sobre la muestra. Cuando un estadístico conduce a un valor aproximado de un parámetro se denomina estimador del parámetro:

	<i>Estadístico (muestra)</i>	<i>Parámetro (población)</i>
Vector medio	$m$	$\mu_1$
Longitud del vector medio	$r$	$\rho_1$
Ángulo medio	$\phi$	$\theta_1$

### 4.2. UNA MUESTRA

En la figura 53 se representa la dirección que tomaron 9 tortugas bobas cuando emergieron de la arena en la playa que su madre depositó los huevos. La pregunta que podríamos plantearlos a partir de estas observaciones es la siguiente ¿se dispersaron

las tortugas bobas de forma uniforme o, por el contrario, siguieron una dirección determinada?, y ¿la población de la que procede la muestra, sigue también una dirección determinada o no? A simple vista no podemos tomar una decisión exacta sobre la situación que se presenta: necesitamos aplicar un test estadístico para poder llegar a una conclusión significativa. Veremos a continuación varios test que nos permitirán determinar si una serie de datos (una muestra) se distribuye uniformemente alrededor del círculo o, por el contrario, muestran direccionalidad (test de uniformidad). También nos puede interesar determinar si una muestra se ajusta a una determinada distribución. Para ello aplicaremos los denominados test de bondad del ajuste.

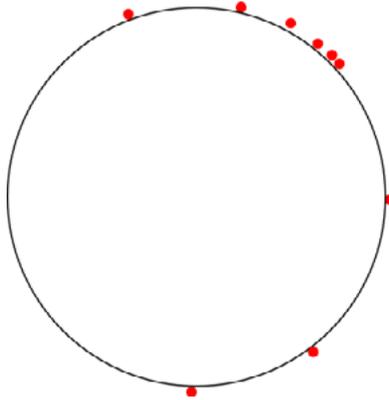


Figura 53. Dirección que toman las 9 tortugas tras emerger de la arena.

Todo test estadístico requiere un *nivel de significación*, que proporciona una idea de cuál es la fortaleza de los datos para contradecir la hipótesis nula que establezcamos. Normalmente es predefinido, siendo el más frecuente  $\alpha = 0,05$ . Como norma, los científicos también desean conocer el menor nivel de significancia que permite rechazar la hipótesis nula ( $H_0$ ), el familiar valor de  $P$  (o  $p$ ). Hay que tener cuidado de no confundir  $\alpha$  con  $P$ . El valor de  $\alpha$  es predeterminado, es decir, es un valor fijado antes de realizar los experimentos, mientras que  $P$  es determinado por la muestra. En todos los casos la hipótesis nula,  $H_0$ , es rechazada si  $P \leq \alpha$  (o  $p \leq \alpha$ ). Si  $P > \alpha$  (o  $p > \alpha$ ) concluimos que no existe evidencia de que  $H_0$  sea falsa.

Llegados a este punto, por tanto, tendremos una muestra a la que habremos estimado sus estadísticos descriptivos y de la que deseamos saber si sigue una dirección determinada o si se ajusta a cierto tipos de distribución. Para ello podemos aplicar los siguientes test, considerando que unos son paramétricos y otros no. Además, para poder aplicarlos debemos considerar si los datos se ajustan a los criterios que exige el test para poder ser aplicado.

#### 4.2.1. Test de Rayleigh

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si la población de la que se ha obtenido la muestra se distribuye uniformemente o, por el contrario, hay direccionalidad en los datos.

**Consideraciones sobre los datos:**

- a) La muestra debe ser unimodal.
- b) Si los datos son axiales, deben doblarse.
- c) Si los datos están agrupados deben corregirse (por la agrupación).

**Ecuaciones.** La longitud del vector medio,  $r$ , proporciona información sobre si la muestra se distribuye uniformemente o, si por el contrario, no lo hace. Si  $r$  es muy pequeño los datos están poco concentrados, eso significa que están distribuidos uniformemente. Por el contrario, si  $r$  es muy elevado los datos están concentrados con lo cual hay direccionalidad.

Los estadísticos utilizados son: la  $R$  de Rayleigh denotada por  $R$  o  $\bar{R}$ :

$$R = n \cdot r$$

y  $z$ , que es utilizado para testar la hipótesis nula:

$$z = R^2/n \quad \text{o} \quad z = n \cdot r^2$$

**Hipótesis nula:**

$H_0$ : Los datos están uniformemente distribuidos.

Que los datos se distribuyan de forma uniforme implica que no existe una “dirección media”. Algunos autores (Zar, 1996), para reflejar que no existe dirección media, utilizan otra notación:

$H_0: \rho = 0$  (es decir, la muestra está uniformemente distribuida alrededor del círculo).

$H_A: \rho \neq 0$  (es decir, la muestra no está uniformemente distribuida alrededor del círculo).

**Regla de decisión:**

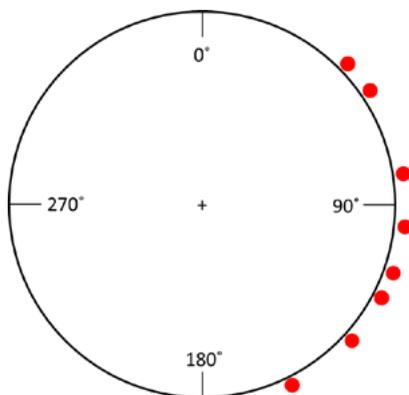
Si el valor crítico de  $z$  obtenido en la tabla A (Anexo I) es menor que el nivel de significancia elegido,  $\alpha$  ( $P < \alpha$ ) la hipótesis nula se rechaza. En este caso la longitud del vector medio difiere significativamente de 0. Si  $P > \alpha$  los datos se distribuyen uniformemente.

**Ejercicio 10** (tomado de Zar, 1996). Se han observado ocho árboles inclinados en las siguientes direcciones: 45°, 55°, 81°, 96°, 110°, 117°, 132° y 154°. Queremos testar la siguiente hipótesis:

$H_0: \rho = 0$  (las inclinaciones de los árboles se distribuyen uniformemente a lo largo del círculo).

$H_A: \rho \neq 0$  (las inclinaciones de los árboles NO se distribuyen uniformemente a lo largo del círculo).

Primero podemos hacer una representación gráfica para tener una primera aproximación de la distribución que siguen los datos (Fig. 54).



**Figura 54.** Representación de los datos del ejercicio 10. Parece que hay cierta direccionalidad.

Ahora podemos proceder a realizar los cálculos para obtener los estadísticos:

Ángulo	Coseno	Seno
45°	0,7071	0,7071
55°	0,5736	0,8192
81°	0,1564	0,9877
96°	-0,1045	0,9945
110°	-0,3420	0,9397
117°	-0,4540	0,8910
132°	-0,6691	0,7431
154°	-0,8988	0,4384
Suma	-1,0313	6,5207

Ahora estimamos las medias:

$$\bar{x} = \frac{\sum_i^n \cos \phi_i}{n} = \frac{-1,0313}{8} = -0,1289; \quad \bar{y} = \frac{\sum_i^n \sen \phi_i}{n} = \frac{6,527}{8} = 0,8151$$

A partir de aquí obtenemos  $r$ :

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} = 0,82522$$

Por lo que:

$$R = n \cdot r = 8 \cdot 0,82522 = 6,60176$$

$$z = R^2/n = (6,60176)^2/8 = 5,448$$

Comprobamos ahora que para un nivel de significación de  $\alpha = 0,05$  y  $n = 8$ :  $z_{0,05,8} = 2,899$  (tabla A). Al ser  $z_{0,05} < z$  rechazamos  $H_0$ , es decir, la muestra no se distribuye uniformemente a lo largo del círculo, por tanto hay cierta direccionalidad.

#### 4.2.2. Test de Hodges-Ajne

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2 \dots \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si la población de la que se ha obtenido la muestra se distribuye uniformemente o, por el contrario, hay direccionalidad en los datos.

##### Consideraciones sobre los datos:

- a) La muestra puede ser unimodal, bimodal o multimodal. No tiene las limitaciones del test de Rayleigh en cuanto a que la muestra tiene que ser unimodal.

**Ecuaciones.** El procedimiento consiste en encontrar el menor número de datos en un rango de  $180^\circ$ . Para ello debemos dibujar una línea que pase por el centro de la circunferencia y “moverla” hasta encontrar una posición en la que a un lado de la línea quede el menor número posible de observaciones.

El estadístico  $m$ , que representa el menor número de observaciones a un lado de la línea, tiene una probabilidad de:

$$P = 2^{1-n} (n - 2m) \binom{n}{m}$$

siendo  $n$  es el tamaño muestral.

##### Regla de decisión:

Si el valor crítico de  $m$  obtenido es mayor que el nivel de significancia elegido en la tabla B (Anexo I) deberemos aceptar  $H_0$ .

**Ejercicio 11** (tomado de Zar, 1996). Disponemos de una serie de datos distribuidos según se muestra en la figura 55 y queremos testar la siguiente hipótesis:

$H_0$ : Los datos están uniformemente distribuidos.

$H_A$ : Los datos no están uniformemente distribuidos.

La línea marcada nos deja 3 observaciones a la derecha, por tanto  $m = 3$ . En la tabla B (Anexo II) obtenemos que  $m_{0,05,24} = 4$ , por tanto rechazamos  $H_0$ . La probabilidad exacta es:

$$P = 2^{1-n} (n - 2m) \binom{n}{m} = 2^{1-23} (24 - 2 \cdot 3) \binom{24}{3} = 2^{-23} \frac{24!}{3!(24-3)!} = 0,00024$$

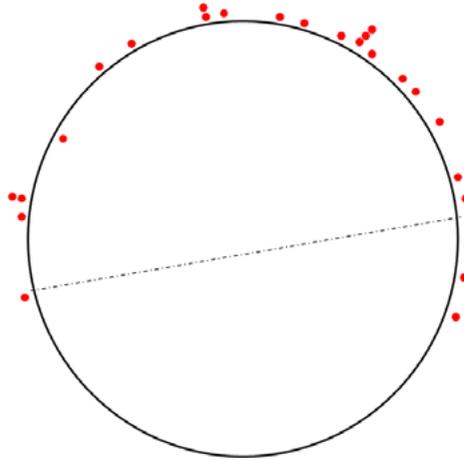


Figura 55. Representación de los datos del ejercicio 11. La línea de puntos deja a su derecha 3 valores.

#### 4.2.3. Test V

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico. Además, se proporciona una dirección teórica determinada por el ángulo  $\theta_0$ .

**Propósito.** Determinar si las direcciones observadas tienen tendencia a concentrarse alrededor del ángulo  $\theta_0$ , además de determinar si la distribución de ángulos difiere de la uniformidad.

##### Consideraciones sobre los datos:

- La muestra debe ser unimodal.
- Si los datos son axiales, deben doblarse.
- Si los datos están agrupados deben corregirse.

**Ecuaciones.** En algunas ocasiones nos puede interesar determinar si los datos se agrupan entorno a un valor dado. Si conocemos la longitud del vector medio,  $r$ , y el ángulo medio,  $\bar{\phi}$  de una serie de datos y además contamos con el ángulo que determina una dirección hipotética,  $\theta_0$ , entonces:

$$v = r \cdot \cos(\bar{\phi} - \theta_0)$$

Si los ángulos observados,  $\phi_i$ , no difieren mucho de  $\theta_0$  entonces  $v$  estará próximo a 1.

El estadístico utilizado es:

$$u = (2 \cdot n)^{1/2} \cdot v$$

**Hipótesis nula:**

$H_0$ : Los datos están uniformemente distribuidos y se concentran alrededor de algún punto.

$H_A$ : Los datos no están uniformemente distribuidos y no se concentran alrededor de algún punto.

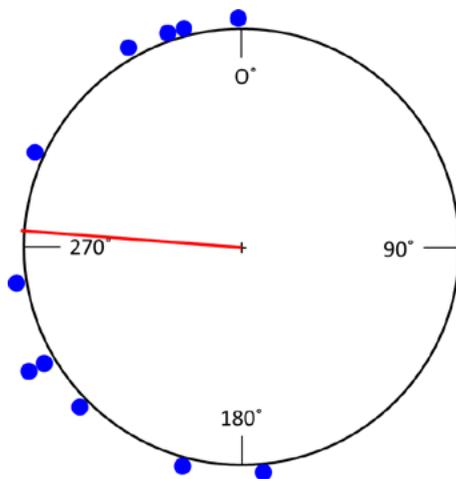
**Regla de decisión:**

Si  $u \geq u(\alpha)$  (tabla C, Anexo I)  $H_0$  se rechaza, en este caso la longitud del vector medio difiere de cero y, por tanto, los datos se distribuyen alrededor de la dirección hipotética.

El test de Rayleigh testa la hipótesis nula de uniformidad frente a una alternativa no conocida, mientras que el test V testa la hipótesis nula de uniformidad frente a una alternativa conocida (el que marca la dirección hipotética). Batschelet (1981) considera que este test es más potente que el de Rayleigh si las direcciones tienden a concentrarse frente a una dirección hipotética dada.

**Ejercicio 12** (tomado de Batschelet, 1989). En un experimento con palomas se quería determinar si las palomas volvían a la zona donde habían sido criadas tras ser liberadas a varios kilómetros de distancia. Tras la liberación se comprobó que las palomas habían tomado las siguientes direcciones: 0°, 175°, 195°, 225°, 240°, 240°, 260°, 295°, 330°, 340° y 345°. La dirección hipotética hacía la que debían dirigirse era 279°. Deseamos comprobar si hay diferencias entre la dirección hipotética y la seguida por el grupo de palomas.

Procedemos, en primer lugar a hacer una representación de los datos (Fig. 56):



**Figura 56.** Representación de los datos del ejercicio 12. En rojo el ángulo medio.

Realizando los cálculos obtenemos que:

$$r = 0,5104; \quad \bar{\phi} = 274^\circ$$

Ahora calculamos  $v$ :

$$v = r \cdot \cos(\bar{\phi} - \theta_0) = 0,5104 \cdot \cos(274^\circ - 279^\circ) = 0,508$$

y

$$u = (2 \cdot n)^{1/2} \cdot v = (2 \cdot 11)^{1/2} \cdot 0,508 = 2,38$$

En la tabla C comprobamos que  $u_{(11, 00.5)} = 1,6479$ . Como  $u \geq u_{(11, 00.5)}$  rechazamos  $H_0$ . Por tanto los datos no se distribuyen uniformemente y se concentran alrededor de  $279^\circ$ .

#### 4.2.4. Test de Rao (Test del espaciado de Rao)

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2 \dots \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si la distribución de ángulos difiere de la uniformidad. Se puede aplicar a distribuciones bimodales o multimodales.

**Consideraciones sobre los datos:**

a) Los datos no deben estar agrupados.

**Ecuaciones.** Se supone que en una distribución uniforme los datos deben estar situados equidistantemente. Si esto es así, el arco entre dos puntos adyacentes debe tener una longitud de  $360^\circ/n$ , de modo que pequeñas desviaciones pueden ser atribuidas al azar, pero grandes desviaciones indican que la distribución no es uniforme.

Partimos de una serie de ángulos ordenados de forma creciente:

$$\phi_1 \leq \phi_2 \leq \phi_3 \dots \leq \phi_n$$

La longitud de  $n$  arcos ( $T$ ) entre puntos consecutivos será:

$$T_1 = \phi_2 - \phi_1$$

$$T_2 = \phi_3 - \phi_2$$

$$T_{n-1} = \phi_n - \phi_{n-1}$$

$$T_n = 360^\circ + \phi_1 - \phi_n$$

Sabemos que:

$$\sum T_i = 360^\circ$$

La desviación de  $T_i$  de su valor esperado es:

$$|T_i - (360^\circ/n)|, \text{ siendo } i = 1, 2, 3 \dots n$$

El estadístico empleado es:

$$U = 1/2 \sum_{i=1}^n |T_i - (360^\circ / n)|$$

o

$$U = \sum |T_i - (360^\circ / n)|$$

donde solo se contabilizan los valores positivos.

**Hipótesis nula:**

$H_0$ : Los datos están uniformemente distribuidos.

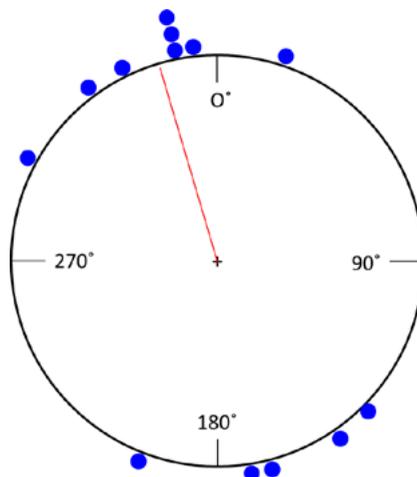
$H_A$ : Los datos no están uniformemente distribuidos.

**Regla de decisión:**

Si  $U$  excede el valor crítico  $U(\alpha)$  (tabla D, Anexo I)  $H_0$  se rechaza, los datos no se distribuyen uniformemente.

**Ejercicio 13** (tomado de Batschelet, 1989). Disponemos de las direcciones que tomaron una serie de aves al ser liberadas en un mismo punto ( $n = 13$ ):  $20^\circ, 135^\circ, 145^\circ, 165^\circ, 170^\circ, 200^\circ, 300^\circ, 325^\circ, 335^\circ, 350^\circ, 350^\circ, 350^\circ$  y  $355^\circ$ . Deseamos saber si las aves se dispersaron de forma uniforme o, por el contrario, tomaron una determinada dirección.

Si observamos la figura 57, en la que hemos representado la dirección que toman las aves, parece muestran una distribución bimodal, pero ¿puede esto excluir la uniformidad? Para determinarlo aplicaremos el test de Rao.



**Figura 57.** Representación de los datos del ejercicio 9. En rojo el ángulo medio.

Calcularemos, en primer lugar las diferencias, en valores absolutos de  $a_n - a_{n-1}$ , recordando que el último valor se calcula:

$$T_n = 360^\circ + \phi_1 - \phi_n = 360^\circ + 20^\circ - 355^\circ = 25^\circ$$

$i$	$T_i =  a_n - a_{n-1} $	$T_i - (360^\circ/13)$
20°	115°	87,3°
135°	10°	-17,7°
145°	20°	-7,7°
165°	5°	-22,7°
170°	30°	2,3°
200°	100°	72,3°
300°	25°	-2,7°
325°	10°	-17,7°
335°	15°	-12,7°
350°	0°	-27,7°
350°	0°	-27,7°
350°	5°	-22,7°
355°	25°	-2,7°

Ahora sumamos los valores positivos y negativos por separado:

<i>Desviaciones negativas</i>	<i>Desviaciones positivas</i>
17,7°	87,3°
7,7°	2,3°
22,7°	72,3°
2,7°	Suma = 161,9°
17,7°	
12,7°	
27,7°	
27,7°	
22,7°	
2,7°	
Suma = 162°	

Obviamente ambos valores deben coincidir o ser muy próximos. El valor del estadístico lo podemos calcular a través de:

$$U = \sum |T_i - (360^\circ / 13)| = 161,9^\circ$$

En la tabla D (Anexo I) comprobamos que  $U_{0,05} = 167,8^\circ$ . Como  $U < U_{0,05}$  no podemos rechazar  $H_0$ , por tanto la distribución es uniforme, aunque de manera bimodal.

### 4.2.5. Test de Kuiper

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico. Además, debe considerarse que los datos se ajustan a una distribución circular determinada (uniforme, de von Mises).

**Propósito.** Determinar si los datos se ajustan a una distribución determinada. El procedimiento consiste en graficar la función teórica y la función derivada de los datos y compararlas. Así obtendremos un valor máximo de desviación de la función de los datos por encima de la función teórica ( $D^+$ ) y un valor mínimo por debajo de la función teórica ( $D^-$ ). Si obtenemos:

$$V_n = D^+ + D^-$$

$V_n$ ,  $D^+$  y  $D^-$  oscilan entre 0 y 1. Un valor bajo de  $V_n$  implicaría un buen ajuste.

**Consideraciones sobre los datos:**

- a) Los datos no deben estar agrupados.

**Ecuaciones.** De acuerdo con Mardia y Jupp (2000) el estadístico que emplearemos es:

$$V_n^* = n^{1/2} V_n \cdot \left(1 + \frac{0,155}{\sqrt{n}} + \frac{0,24}{n}\right)$$

**Hipótesis nula:**

$H_0$ : La distribución de los datos coincide con la distribución dada.

$H_A$ : La distribución de los datos no coincide con la distribución dada.

**Regla de decisión:**

$H_0$  se rechaza si el valor crítico obtenido en las tabla E (Anexo I) es menor que el valor obtenido en el test.

**Ejercicio 14** (tomado de Mardia y Jupp, 2000). Queremos determinar si una serie de aves ( $i$ ) se distribuyen uniformemente a partir de su liberación en un punto dado. Los ángulos de dispersión fueron: 55°, 60°, 65°, 90°, 100°, 110°, 260°, 275°, 285° y 290°.

Hacemos, en primer lugar una representación de los datos (Fig. 58).

Para estimar las desviaciones primero debemos calcular:

$$U_i = \frac{\phi_i}{2\pi}$$

y

$$\frac{U_i - n}{n}$$

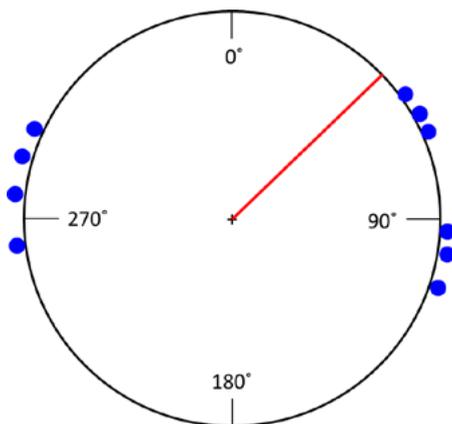


Figura 58. Representación de los datos del ejercicio 14. En rojo el ángulo medio.

De modo que obtenemos la siguiente tabla:

$i$	$\theta_i$	$U_i$	$i/n$	$U_i - i/n$
1	55°	0,153	0,1	<b>0,053</b>
2	60°	0,167	0,2	-0,033
3	65°	0,181	0,3	-0,119
4	95°	0,264	0,4	-0,136
5	100°	0,278	0,5	-0,222
6	110°	0,306	0,6	<b>-0,294</b>
7	260°	0,722	0,7	0,022
8	275°	0,764	0,8	-0,036
9	285°	0,792	0,9	-0,108
10	295°	0,819	1	-0,181

Vemos, por tanto, que  $D^+ = 0,053$  y  $D^- = -0,294$ . Entonces:

$$V_n = D^+ + D^- = 0,053 + 0,294 + 0,1^* = 0,447$$

De modo que:

$$V_n^* = n^{1/2} V_n \cdot \left(1 + \frac{0,155}{\sqrt{n}} + \frac{0,24}{n}\right) = 10^{1/2} 0,447 \cdot \left(1 + \frac{0,155}{\sqrt{10}} + \frac{0,24}{10}\right) = 1,517$$

Como  $V_{0,05,10}^* = 1,625$  (tabla E, Anexo V) aceptamos  $H_0$ , los datos se distribuyen de forma uniforme.

\* Ajuste propuesto por Mardia y Jupp (2000).

#### 4.2.6. Test de la chi-cuadrado

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico. Además, debe considerarse una distribución circular (uniforme, von Mises).

**Propósito.** Determinar si los datos se ajustan a una distribución determinada. El fundamento de este test es dividir el círculo en una serie de arcos, establecer el número de datos que hay en cada arco y calcular las frecuencias esperadas en cada arco según la distribución circular establecida. Los datos se ajustarán a la distribución dada si las frecuencias esperadas no difieren mucho de las frecuencias observadas. El test de chi-cuadrado también es un test de uniformidad si la distribución dada es la uniforme. En este sentido el test de la chi-cuadrado es útil para distribuciones bi y multimodales.

##### Consideraciones sobre los datos:

- Los datos deben estar agrupados.
- Los intervalos deben ser equidistantes.
- Las frecuencias esperadas en cada intervalo deben ser, como mínimo, iguales o mayores a 4. Para cumplir este requisito normalmente son necesarios tamaños muestrales muy grandes, como mínimo del orden de cinco veces  $k$  (ver más adelante).

**Ecuaciones.** Partimos de una serie de datos divididos en  $k$  grupos, siendo  $n_i$  y  $e_i$  las frecuencias observadas y esperadas del grupo  $i$ . Entonces, el estadístico a aplicar es:

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - e_i)^2}{e_i}$$

##### Hipótesis nula:

$H_0$ : La distribución de los datos coincide con la distribución dada.

$H_0$ : La distribución de los datos no coincide con la distribución dada.

##### Regla de decisión:

$H_0$  se rechaza si el valor crítico determinado en la tabla F (Anexo I) es menor que el valor obtenido en el test.

**Ejercicio 15** (tomado de Zar, 1996). Un grupo de hormigas ( $n = 105$ ) fueron sometidas a un estímulo sonoro y se observó la dirección que tomaron ( $k = 12$ ). Queremos testar la hipótesis nula de que las hormigas se dispersaron de forma uniforme.

En primer lugar calculamos las frecuencias esperadas para todas las categorías:

$$e_i = \frac{n}{k} = \frac{105}{12} = 8,75$$

Ahora podemos calcular el estadístico, una vez que tenemos las dos frecuencias:

Dirección	Frecuencia observada ( $n_i$ )	Frecuencia esperada ( $e_i$ )
0°-30°	0	8,75
30°-60°	6	8,75
60°-90°	9	8,75
90°-120°	13	8,75
120°-150°	15	8,75
150°-180°	22	8,75
180°-210°	17	8,75
210°-240°	12	8,75
240°-270°	8	8,75
270°-300°	3	8,75
300°-330°	0	8,75
330°-360°	0	8,75
Suma =	105	

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - e_i)^2}{e_i} = \frac{(0 - 8,75)^2}{8,75} + \frac{(6 - 8,75)^2}{8,75} + \dots + \frac{(0 - 8,75)^2}{8,75} = 66,543$$

Ahora debemos consultar en las tablas de chi-cuadrado (Anexo F) el valor crítico para  $\alpha = 0,05$  y  $k-1$  grados de libertad, obteniendo  $\chi_{0,05,11}^2 = 19,675$ . Como  $19,675 < 66,543$  rechazamos  $H_0$ , por lo que concluimos que el sonido afecta la orientación de las hormigas. Si observamos la figura 59 podemos comprobar que hay cierta direccionalidad.

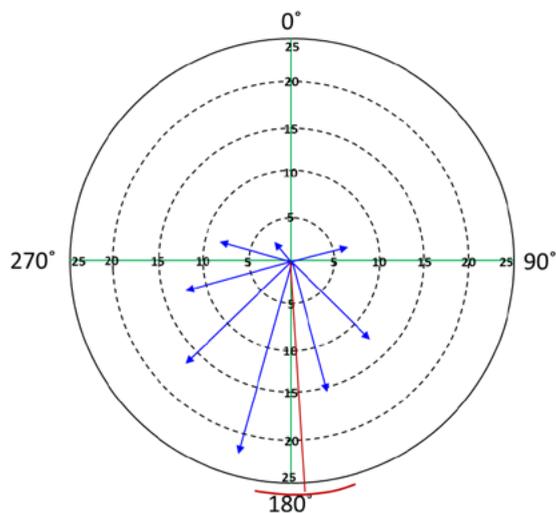


Figura 59. Datos del ejercicio 15.

#### 4.2.7. Test $U^2$ de Watson

**Datos.** Una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$ , que representan direcciones en un plano, tiempo o un fenómeno cíclico. Además, debe considerarse una distribución circular (uniforme, von Mises).

**Propósito.** Determinar si los datos se ajustan a una distribución determinada. En el test de Kuiper vimos que obteníamos unas desviaciones máxima ( $D^+$ ) y mínimas ( $D^-$ ) respecto a la distribución teórica. Aquí el procedimiento es el mismo, salvo porque se utilizan los cuadrados de las desviaciones para realizar el test. Se puede aplicar tanto a distribuciones unimodales como multimodales y es muy potente cuando los tamaños muestrales son pequeños.

##### Consideraciones sobre los datos:

a) Los datos deben estar ordenados en forma ascendente.

**Ecuaciones.** Para cada dato ( $i$ ) debemos calcular:

$$u_i = \frac{\phi_i}{360^\circ}$$

$$\sum u_i^2; \quad \sum i \cdot u_i;$$

Y a partir de aquí:

$$U^2 = \sum u_i^2 - \frac{(\sum u_i)^2}{n} - \frac{2}{n} \sum i \cdot u_i + (n+1) \cdot \bar{u} + \frac{n}{12}$$

##### Hipótesis nula:

$H_0$ : La distribución de los datos coincide con la distribución dada.

$H_A$ : La distribución de los datos no coincide con la distribución dada.

##### Regla de decisión:

$H_0$  se rechaza si el valor obtenido en el test es menor que el obtenido en las tablas de la  $U^2$  de Watson (tabla G, Anexo I).

**Ejercicio 16** (tomado de Zar, 1996). Deseamos saber si un grupo de árboles orientados en las siguientes direcciones:  $45^\circ, 55^\circ, 81^\circ, 96^\circ, 110^\circ, 117^\circ, 132^\circ$  y  $154^\circ$  (Fig. 60) se distribuyen uniformemente.

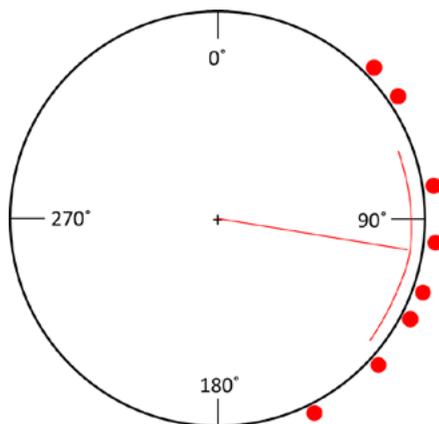


Figura 60. Representación gráfica de los datos del ejercicio 16.

En la siguiente tabla se recogen los cálculos señalados anteriormente:

$i$	$\theta_i$	$u_i$	$u_i^2$	$u_i \cdot i$
1	45°	0,1250	0,0156	0,1250
2	55°	0,1528	0,0233	0,3056
3	81°	0,2250	0,0506	0,6750
4	96°	0,2667	0,0711	1,0667
5	110°	0,3056	0,0934	1,5278
6	117°	0,3250	0,1056	1,9500
7	132°	0,3667	0,1344	2,5667
8	154°	0,4278	0,1830	3,4222
Suma:		2,1944	0,6771	11,6389

A partir de aquí calculamos el estadístico  $U^2$ :

$$U^2 = \sum u_i^2 - \frac{(\sum u_i)^2}{n} - \frac{2}{n} \sum i \cdot u_i + (n+1) \cdot \bar{u} + \frac{n}{12} =$$

$$0,6771 - (2,1944)^2/8 - 2/8 (11,6389) + 9 (2,1944/8) + 8/12 = 0,2989$$

En la tabla G (Anexo VII) obtenemos que:  $U^2_{0,05,7,7} = 0,1986$ , por tanto rechazamos  $H_0$ , los árboles no se distribuyen uniformemente.

#### 4.2.8. Test de Hotelling para una muestra

**Datos.** Sabemos que a una muestra de  $n$  valores angulares  $\phi_1, \phi_2, \dots, \phi_n$  le podemos calcular su ángulo medio ( $\bar{\phi}$ ) y su vector medio asociado,  $r$ . Imaginemos que tenemos ahora  $k$  grupos de  $n$  valores angulares.

**Propósito.** Determinar si existen diferencias en los ángulos medios.

**Ecuaciones.** Sabemos que cuando tenemos una muestra con  $k$  ángulos medios podemos calcular:

$$\bar{X} = \frac{\sum_{j=1}^k r_j \cdot \cos \bar{\phi}_j}{k}; \quad \bar{Y} = \frac{\sum_{j=1}^k r_j \cdot \sen \bar{\phi}_j}{k}$$

A partir de aquí obtenemos:

$$\begin{aligned} \sum x^2 &= \sum X_j^2 - \frac{(\sum X_j)^2}{k}, \\ \sum y^2 &= \sum Y_j^2 - \frac{(\sum Y_j)^2}{k}, \\ \sum x \cdot y &= \sum X_j \cdot Y_j - \frac{\sum X_j \cdot \sum Y_j}{k} \end{aligned}$$

Entonces obtenemos el estadístico  $F$ :

$$F = \frac{k \cdot (k-2)}{2} \left[ \frac{\bar{X}^2 \sum y^2 - 2\bar{X} \cdot \bar{Y} \sum x \cdot y + \bar{Y}^2 \sum x^2}{\sum x^2 \sum y^2 - (\sum x \cdot y)^2} \right]$$

con 2 y  $k-2$  grados de libertad.

**Consideraciones sobre los datos:**

a) Los datos deben ser normales (von Mises).

**Hipótesis nula:**

$H_0$ : No hay una dirección poblacional media, por tanto hay una distribución uniforme de los datos.

$H_A$ : Hay una dirección poblacional media, por tanto no hay una distribución uniforme de los datos.

**Regla de decisión:**

Si el valor crítico en la tabla H (Anexo I) es menor que el nivel de significancia elegido la hipótesis nula se rechaza.

**Ejercicio 17.** Retomamos los cálculos del ejercicio 5 donde vimos que:

Población ( $j$ )	$\phi_j$	$r_j$	$X_j = r_j \cdot \cos \phi_j$	$Y_j = r_j \cdot \sen \phi_j$
1	160°	0,8954	-0,8414	0,3062
2	169°	0,7747	-0,7605	0,1478
3	117°	0,4696	-0,2132	0,4184
4	140°	0,8794	-0,6737	0,5653
5	186°	0,3922	-0,3901	-0,0410
6	134°	0,6952	-0,4829	0,5001
7	171°	0,3338	-0,3297	0,0522
		Suma =	-3,6914	1,9491

Entonces:

$$\bar{X} = \frac{\sum_{j=1}^7 r_j \cdot \cos \bar{\phi}_j}{7} = \frac{-3,6914}{7} = -0,5273$$

$$\bar{Y} = \frac{\sum_{j=1}^7 r_j \cdot \sen \bar{\phi}_j}{7} = \frac{1,9491}{7} = 0,2784$$

Con estos datos podemos calcular:

Población (j)	$\phi_j$	$r_j$	$X_j = r_j \cos \phi_j$	$Y_j = r_j \sen \phi_j$	$X_j^2$	$Y_j^2$	$X_j \cdot Y_j$
1	160°	0,8954	-0,8414	0,3062	0,7080	0,0938	-0,2577
2	169°	0,7747	-0,7605	0,1478	0,5783	0,0219	-0,1124
3	117°	0,4696	-0,2132	0,4184	0,0455	0,1751	-0,0892
4	140°	0,8794	-0,6737	0,5653	0,4538	0,3195	-0,3808
5	186°	0,3922	-0,3901	-0,0410	0,1521	0,0017	0,0160
6	134°	0,6952	-0,4829	0,5001	0,2332	0,2501	-0,2415
7	171°	0,3338	-0,3297	0,0522	0,1087	0,0027	-0,0172
Suma =			-3,6914	1,9491	2,2796	0,8647	-1,0828

Por tanto tenemos que:

$$k = 7; \quad \bar{X} = -0,5273; \quad \bar{Y} = 0,2784;$$

$$\sum X_j = -3,6914; \quad \sum X_j^2 = 2,2796; \quad \sum x^2 = 0,3329;$$

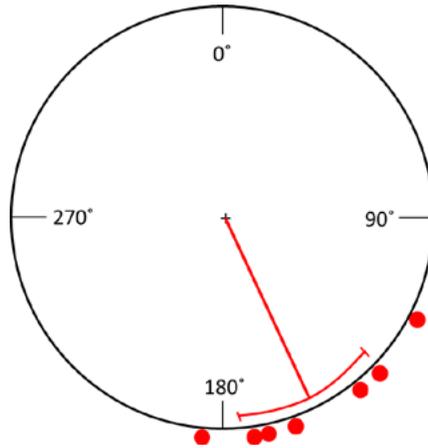
$$\sum Y_j = 1,9490; \quad \sum Y_j^2 = 0,8647; \quad \sum y^2 = 0,3220;$$

$$\sum X_j Y_j = -1,0828, \quad \sum x \cdot y = -0,0550$$

Finalmente:

$$F = \frac{7 \cdot (7-2)}{2} \left[ \frac{(-0,5273)^2 \cdot (0,3220) - 2 \cdot (-0,5273) \cdot (0,2784) \cdot (-0,0550) + (0,2784)^2 \cdot (0,3329)}{(0,3329)(0,3220) - (-0,0550)^2} \right] = 16,66$$

Como  $F_{0,05,2,5} = 5,79$  rechazamos  $H_0$ , por tanto hay una dirección poblacional media (Fig. 61).



**Figura 61.** Representación gráfica de los datos del ejercicio 17. Se comprueba que hay una dirección poblacional media.

#### 4.2.9. Test de Moore para una muestra

**Datos.**  $k$  grupos de  $n$  valores angulares.

**Propósito.** Determinar si la población de la que se han obtenido los ángulos medios se distribuye uniformemente. A diferencia del test de Hotelling este test es no paramétrico.

**Ecuaciones.** Debemos calcular los siguientes estadísticos:

$$X = \frac{\sum_{i=1}^k i \cdot \cos \bar{\phi}_i}{k}; \quad Y = \frac{\sum_{i=1}^k i \cdot \sen \bar{\phi}_i}{k};$$

$$R' = \sqrt{\frac{X^2 + Y^2}{k}}$$

Donde  $i$  son los rangos asignados.

**Consideraciones sobre los datos:**

- a) La población no es normal (Von Mises).

**Hipótesis nula:**

$H_0$ : La población de la que se han obtenido los ángulos medios se distribuye uniformemente.

$H_A$ : La población de la que se han obtenido los ángulos medios no se distribuye uniformemente.

### Regla de decisión:

Si el valor crítico obtenido en la tabla I (Anexo I) es menor que el obtenido en el test la hipótesis nula se rechaza. En este caso la longitud del vector medio difiere significativamente de 0. Si  $P > \alpha$  los datos se distribuyen aleatoriamente.

**Ejercicio 18** (tomado de Zar, 1996). Tenemos  $k = 7$  muestras formadas por 10 medidas cada una de las que conocemos su ángulo medio y su vector longitud según se muestra en la tabla siguiente. Nos preguntamos si la población de la que proceden las muestras está distribuida uniformemente a lo largo del círculo.

Rango ( $i$ )	$r_i$	$\Phi_i$	$\cos \Phi_i$	$i \cdot \cos \Phi_i$	$\text{seno } \Phi_i$	$i \cdot \text{seno } \Phi_i$
1	0,3338	171°	-0,9877	-0,9877	0,1564	0,1564
2	0,3922	186°	-0,9945	-1,9890	-0,1045	-0,2091
3	0,4649	117°	-0,4540	-1,3620	0,8910	2,6730
4	0,6962	134°	-0,6947	-2,7786	0,7193	2,8774
5	0,7747	169°	-0,9816	-4,9081	0,1908	0,9540
6	0,8794	140°	-0,7660	-4,5963	0,6428	3,8567
7	0,8954	160°	-0,9397	-6,5778	0,3420	2,3941
Suma:				-23,1996		12,7027

Partiendo de estos datos podemos calcular:

$$X = \frac{\sum_{i=1}^k i \cdot \cos \bar{\phi}_i}{k} = \frac{-23,196}{7} = -3,3142$$

$$Y = \frac{\sum_{i=1}^k i \cdot \text{sen } \bar{\phi}_i}{k} = \frac{12,7027}{7} = 1,8146$$

$$R' = \sqrt{\frac{(-3,3142)^2 + (1,8143)^2}{7}} = 1,428$$

Como  $R'_{0,05,7} = 1,061$  rechazamos  $H_0$ ; es decir la población de la que proceden las muestras no se distribuye uniformemente a lo largo del círculo.

#### 4.2.10. Bootstrapping

El uso de ordenadores nos permite analizar grandes cantidades de datos en poco tiempo. El bootstrap es la técnica más versátil y conocida de los llamados métodos de *remuestreo*. El fundamento de esta técnica es la repetición, ya que consiste en generar un elevado número de muestras como base para estudiar el comportamiento de determinados estadísticos, como los intervalos de confianza para la media, por ejemplo. La técnica de bootstrap solo puede realizarse en un ordenador.

En el caso de la estimación de parámetros, las técnicas bootstrap implican básicamente desarrollar un proceso en diferentes pasos (Flores *et al.*, 2007):

1. A partir de la muestra original, se extrae una nueva muestra, por medio de muestreo con reposición. Es decir, tras la extracción de un primer elemento, este se repone en la muestra original de tal forma que podría ser elegido de nuevo como segundo elemento de la muestra extraída. De este modo, cada observación individual tiene una probabilidad  $1/n$  de ser elegida cada vez, como si el muestreo se realizara sin reposición en un universo infinitamente grande construido a partir de la información que provee la muestra.
2. Para la muestra obtenida se calcula el valor de un determinado estadístico  $\hat{A}$  que se utiliza como estimador del parámetro poblacional  $A$ , en cuyo estudio estamos interesados.
3. Repetimos los dos pasos anteriores, hasta obtener un elevado número de estimaciones  $\hat{A}$ .
4. Se construye una distribución empírica del estadístico  $\hat{A}$ , que representa una buena aproximación a la verdadera distribución de probabilidad para ese estadístico. Es decir, se determina de este modo la distribución muestral de un estadístico sin haber hecho suposiciones sobre la distribución teórica a la que esta se ajusta y sin manejar fórmulas analíticas para determinar los correspondientes parámetros de esa distribución.

De acuerdo con la idea central en que se basa el método bootstrap, el procedimiento supone utilizar la muestra considerando que en sí misma contiene la información básica sobre la población. Por tanto, la adecuación de este método será tanto mayor cuanto más información aporte la muestra sobre la población. Una consecuencia directa es que a medida que aumenta el tamaño de la muestra mejor será la estimación que podemos hacer sobre la distribución muestral de un estadístico. No obstante, incluso con muestras pequeñas, entre 10 y 20 casos, el método bootstrap puede ofrecer resultados correctos.

**Ejemplo 19.** Con los datos del ejercicio 5 y con el programa PAST obtenemos los siguientes resultados para el intervalo de confianza para la media:

Media circular: 98,988

Bootstrapped 95%: (73,67, 123,7)

Que no son muy diferentes de los obtenidos en el ejercicio 5: (76°, 132°)

### 4.3. DOS O MÁS MUESTRAS

Cuando se analizan dos muestras debemos asegurarnos:

1. Que las muestras son elegidas aleatoriamente de la misma población y
2. Que las muestras son independientes una de la otra.

### 4.3.1. Test de Watson-Williams

**Datos.** Dos muestras  $n_1$  y  $n_2$  con valores angulares  $\phi_1, \phi_2 \dots \phi_n$  y  $\gamma_1, \gamma_2 \dots \gamma_n$  que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si los ángulos medios de las dos muestras difieren significativamente uno del otro.

**Ecuaciones.** Para cada muestra calculamos el vector medio:

$$x_i = \cos \phi_i; \quad y_i = \text{sen} \phi_i$$

Si denominamos  $\bar{x}$  e  $\bar{y}$  a la media de  $x_i$  e  $y_i$  respectivamente se cumplirá que:

$$\bar{x}_1 = \frac{\sum_i^n \cos \phi_i}{n}$$

$$\bar{y}_1 = \frac{\sum_i^n \text{sen} \phi_i}{n}$$

$$\bar{x}_2 = \frac{\sum_i^n \cos \gamma_i}{n}$$

$$\bar{y}_2 = \frac{\sum_i^n \text{sen} \gamma_i}{n}$$

Con estas coordenadas se obtiene la longitud del vector medio,  $r$ :

$$r_1 = \sqrt{\bar{x}_1^2 + \bar{y}_1^2}; \quad r_2 = \sqrt{\bar{x}_2^2 + \bar{y}_2^2}$$

Ahora debemos realizar los mismos cálculos pero considerando las dos muestras combinadas, de modo que:

$$n = n_1 + n_2$$

Para los datos combinados los componentes del vector resultante son:

$$x = (\bar{x}_1 + \bar{x}_2) / n; \quad y = (\bar{y}_1 + \bar{y}_2) / n$$

y

$$r = \sqrt{\bar{x}^2 + \bar{y}^2}$$

#### Consideraciones sobre los datos:

- Las dos muestras deben proceder de una población que se ajusta a la distribución de von Mises (normal).

b) El parámetro de concentración debe ser igual en las dos poblaciones:

$$\kappa_1 = \kappa_2 = \kappa_n$$

c) El parámetro de concentración debe ser suficientemente grande:

$$\kappa > 2$$

### Hipótesis nula:

Si  $\theta_1$  y  $\theta_2$  son los ángulos medios de las dos muestras:

$H_0: \theta_1 = \theta_2$ ; por consiguiente  $H_A: \theta_1 \neq \theta_2$ .

### Estadístico:

$$F = g(n-2) \frac{r_1 + r_2 - \bar{r}}{n - (r_1 + r_2)}$$

Para calcular la constante  $g$  debemos estimar:

$$\bar{r} = \frac{r_1 + r_2}{n}$$

y

$$\bar{n} = \frac{n_1 + n_2}{2}$$

Conocido  $\bar{r}$  y  $\bar{n}$  empleando la tabla K podemos calcular el valor de  $\bar{\kappa}$ . Entonces:

$$g = 1 + \frac{3}{8\bar{\kappa}}$$

### Regla de decisión:

Si el estadístico  $F$  es mayor que  $F_{1,n-2}$  rechazamos la hipótesis nula, por tanto concluimos que las direcciones medias de las dos muestras se desvían significativamente una de la otra.

También existe una versión de este test para  $k$  muestras. En este caso:

$$F = g \frac{(n-k) \cdot (\sum r_i - r)}{(k-1) \cdot (n - \sum r_i)}$$

Para calcular  $g$  usaremos las ecuaciones:

$$\bar{n} = \frac{n_1 + n_2 + \dots + n_k}{k}$$

$$\bar{r} = \sum r_i / n$$

y procederemos como en el caso anterior.

**Ejercicio 20** (tomado de Batschelet, 1981). Disponemos de 20 hormigas entrenadas para volver a sus nidos bajo diferentes condiciones. Se pide si existen diferencias en las rutas de que siguieron las hormigas control y las experimentales. Los datos se exponen a continuación:

<i>Desviación respecto a la dirección verdadera</i>	<i>Frecuencias observadas</i>	
	<i>Experimental</i>	<i>Control</i>
-20°	1	0
-10°	7	3
0°	2	3
10°	0	3
20°	0	1
	n = 10	n = 10

En primer lugar debemos transformar los ángulos negativos en positivos (-20° = 340°; -10° = 350°) y crear la tabla de frecuencias, para luego poder calcular los cosenos y senos de los dos grupos:

<i>Experimental</i>	<i>Coseno</i>	<i>Seno</i>	<i>Control</i>	<i>Seno</i>	<i>Coseno</i>
340°	0,940	-0,342	350°	-0,174	0,985
350°	0,985	-0,174	350°	-0,174	0,985
350°	0,985	-0,174	350°	-0,174	0,985
350°	0,985	-0,174	0°	0,000	1,000
350°	0,985	-0,174	0°	0,000	1,000
350°	0,985	-0,174	0°	0,000	1,000
350°	0,985	-0,174	10°	0,174	0,985
350°	0,985	-0,174	10°	0,174	0,985
0°	1,000	0,000	10°	0,174	0,985
0°	1,000	0,000	20°	0,342	0,940
Suma =	9,833	-1,558		0,342	9,849

A partir de aquí obtenemos:

$$r_1 = 9,956; \quad r_2 = 9,854; \quad r = 19,719$$

Por otro lado:

$$r_1 + r_2 = 19,810; \quad r_1 + r_2 - r = 0,091; \quad n - (r_1 + r_2) = 0,190;$$

$$\bar{r} = \frac{r_1 + r_2}{n} = 0,991$$

$$\bar{n} = \frac{n_1 + n_2}{2} = 10$$

Con  $\bar{r} = 0,991$  y  $\bar{n} = 10$  vamos a la tabla L, donde obtenemos que  $\tilde{\kappa} = 9,70$ . Ahora podemos calcular g:

$$g = 1 + \frac{3}{8\tilde{\kappa}} = 1 + 0,03 = 1,003$$

Por tanto:

$$F = g(n - 2) \frac{r_1 + r_2 - r}{n - (r_1 + r_2)} = 1,003 \cdot 18 \frac{0,091}{0,190} = 8,6469$$

Con un valor de  $\alpha = 0,01$  en la tabla H comprobamos que  $F_{1,18} = 8,29$ . Como  $F > F_{1,18}$  rechazamos  $H_0$ : hay diferencias en la ruta que siguieron ambos grupos de hormigas.

Zar (1996) propone una modificación para es test. El estadístico a utilizar es:

$$F = K \frac{(N - 2) \cdot (R_1 + R_2 - R)}{N - R_1 - R_2}$$

Donde  $N = n_1 + n_2$ ,  $R$  es la R de Rayleigh de las dos muestras en conjunto;  $R_1$  y  $R_2$  son la R de Rayleigh de cada una de las dos muestras y  $K$  es un factor de corrección que se estima en la tabla K (Anexo I). El valor crítico para este test es  $F_{\alpha(1),1,N-2}$ .

**Ejercicio 21** (tomado de Zar, 1996). Partimos de dos muestras y deseamos saber si las medias angulares son iguales ( $H_0: \mu_1 = \mu_2$ ).

Muestra 1			Muestra 2		
$\alpha_i$	seno $\alpha_i$	coseno $\alpha_i$	$\alpha_i$	seno $\alpha_i$	coseno $\alpha_i$
94°	0,9976	-0,0698	77 °	0,9744	0,2250
65°	0,9063	0,4226	70 °	0,9397	0,3420
45°	0,7071	0,7071	61 °	0,8746	0,4848
52°	0,7880	0,6157	45 °	0,7071	0,7071
38°	0,6157	0,7880	50 °	0,7660	0,6428
47°	0,7314	0,6820	35 °	0,5736	0,8192
73°	0,9563	0,2924	48 °	0,7431	0,6691
82°	0,9903	0,1392	65 °	0,9063	0,4226
90°	1,0000	0,0000	36 °	0,5878	0,8090
40°	0,6428	0,7660			
87°	0,9986	0,0523			
Suma:	9,3340	4,3956		7,0726	5,1216

Ahora debemos calcular los valores medios:

$$\bar{Y}_1 = \frac{9,3340}{11} = 0,8485, \quad \bar{X}_1 = \frac{4,39560}{11} = 0,3996$$

Ahora calculamos  $r_1$ :

$$r_1 = \sqrt{0,8485^2 + 0,3996^2} = 0,9379$$

A continuación procedemos a la estima del ángulo medio para la muestra 1:

$$\bar{\alpha}_1 = \arctan \frac{0,8485}{0,3996} = 64,78^\circ \approx 65^\circ$$

Y el seno y coseno de este ángulo:

$$\text{sen } \bar{\alpha}_1 = 0,9047; \quad \text{cos } \bar{\alpha}_1 = 0,42605$$

Finalmente calculamos  $R_1$ :

$$R_1 = n_1 \cdot r_1 = 11 \cdot 0,9379 = 10,317$$

Repetimos los cálculos para la muestra 2, obteniendo los siguientes resultados:

$$\bar{Y}_2 = \frac{7,07265}{9} = 0,7858; \quad \bar{X}_2 = \frac{5,1216}{9} = 0,5690$$

$$r_2 = 0,97026; \quad \bar{\alpha}_2 = 54^\circ$$

$$\text{sen } \bar{\alpha}_2 = 0,8099, \quad \text{cos } \bar{\alpha}_2 = 0,5865$$

$$R_2 = n_2 \cdot r_2 = 9 \cdot 0,97026 = 8,7323$$

Ahora procedemos a realizar los cálculos para las dos muestras combinadas:

$$N = 11 + 9 = 20$$

$$\sum \text{sen } \alpha_i = 9,334 + 7,0726 = 16,4066$$

$$\sum \text{cos } \alpha_i = 4,3956 + 5,1216 = 9,5171$$

Por tanto:

$$Y = \frac{16,4066}{20} = 0,82033; \quad X = \frac{9,5171}{20} = 0,4758$$

$$r = \sqrt{0,82033^2 + 0,4758^2} = 0,9483$$

$$R = 20 \cdot 0,9483 = 18,967$$

Para calcular  $K$  necesitamos estimar  $r_w$ :

$$r_w = \frac{n_1 \cdot r_1 + n_2 \cdot r_2}{N} = \frac{R_1 + R_2}{N}$$

En este caso:

$$r_w = \frac{R_1 + R_2}{N} = \frac{10,317 + 8,732}{20} = 0,952$$

Con este valor, en la tabla K (Anexo I), obtenemos  $K = 1,0351$ . Por consiguiente:

$$F = K \frac{(N - 2) \cdot (R_1 + R_2 - R)}{N - R_1 - R_2} = 1,0351 \frac{(20 - 2) \cdot (10,317 + 8,7323 - 18,967)}{20 - 10,317 - 8,7323} = 1,61$$

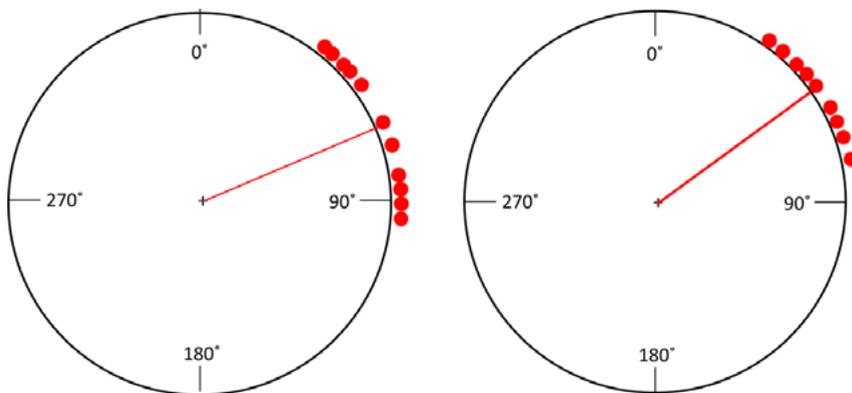
En la tabla H (Anexo I) comprobamos que  $F_{0,05(1),1,18} = 4,41$ , por tanto no rechazamos  $H_0$ , concluyendo que las medias de las dos muestras son iguales (Fig. 62).

Finalmente podemos calcular el ángulo medio para las dos poblaciones:

$$\text{sen } \bar{\alpha} = \frac{Y}{r} = 0,86500; \quad \text{cos } \bar{\alpha} = X / r = 0,50177$$

Por tanto

$$\bar{\alpha} = \arccos 0,50177 = 59,88^\circ$$



**Figura 62.** Representación gráfica de los datos del ejercicio 21 (a la izquierda muestra 2, a la derecha muestra 1). Se puede apreciar como las direcciones medias son casi iguales.

El test de Watson-Williams puede generalizarse para muestras múltiples (Zar, 1996). En este caso el estadístico que se aplica es:

$$F = K \frac{(N - k) \cdot (\sum_{j=1}^k R_j - R)}{(k - 1) - (N - \sum_{j=1}^k R_j)}$$

Donde  $k$  es el número de muestras,  $R$  es la  $R$  de Rayleigh para todas las muestras combinadas y  $N$  es la suma de los tamaños muestrales de cada una de las muestras analizadas. El valor crítico en este caso es  $F_{\alpha(1),k-1,N-k}$ .

**Ejercicio 22** (tomado de Zar, 1996). Disponemos de tres muestras angulares (Fig. 63) y deseamos saber si las medias angulares son iguales ( $H_0: \mu_1 = \mu_2 = \mu_3$ ).

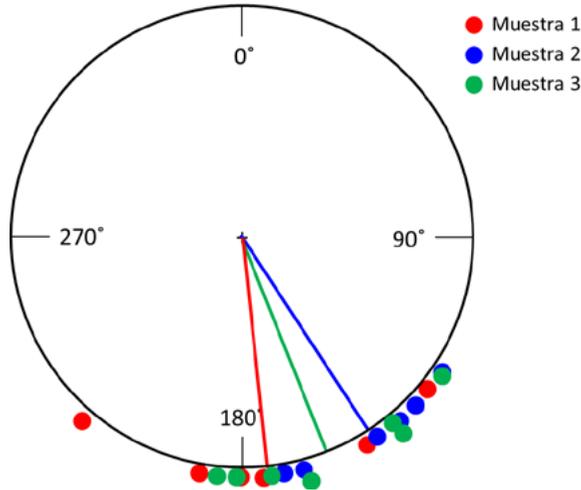


Figura 63. Representación de los datos del ejercicio 22.

Muestra 1			Muestra 2			Muestra 3		
$\alpha_i$	seno $\alpha_i$	coseno $\alpha_i$	$\alpha_i$	seno $\alpha_i$	coseno $\alpha_i$	$\alpha_i$	seno $\alpha_i$	coseno $\alpha_i$
135°	0,7071	-0,7071	150°	0,5000	-0,8660	140°	0,6428	-0,7660
145°	0,5736	-0,8192	130°	0,7660	-0,6428	165°	0,2588	-0,9659
125°	0,8192	-0,5736	175°	0,0872	-0,9962	185°	-0,0872	-0,9962
140°	0,6428	-0,7660	190°	-0,1736	-0,9848	180°	0,0000	-1,0000
165°	0,2588	-0,9659	180°	0,0000	-1,0000	125°	0,8192	-0,5736
170°	0,1736	-0,9848	220°	-0,6428	-0,7660	175°	0,0872	-0,9962
						140°	0,6428	-0,7660
Suma:	3,1751	-4,8166		0,5368	-5,2559		2,3635	-6,0640

Procediendo como en el ejemplo anterior obtenemos:

$$\bar{\alpha}_1 = 147^\circ; \quad r_1 = 0,9615; \quad R_1 = 5,7689$$

$$\bar{\alpha}_2 = 174^\circ; \quad r_2 = 0,8805; \quad R_2 = 5,2832$$

$$\bar{\alpha}_3 = 159^\circ; \quad r_3 = 0,9297; \quad R_3 = 6,5083$$

Para todos los datos combinados tendremos que:

$$k = 3$$

$$N = 6 + 6 + 7 = 19$$

De modo que:

$$\sum \text{sen } \alpha_i = 3,1751 + 0,5368 + 2,3635 = 6,0754$$

$$\sum \text{cos } \alpha_i = -4,8162 - 5,2558 - 6,064 = -16,1364$$

Por tanto:

$$Y = \frac{6,0754}{19} = 0,31976; \quad X = \frac{-16,1364}{19} = -0,8492$$

$$r = \sqrt{0,31976^2 + (-0,8492)^2} = 0,9075$$

$$R = 19 \cdot 0,9075 = 17,2425$$

Calculamos ahora el valor de  $r_w$ :

$$r_w = \frac{R_1 + R_2 + R_3}{N} = \frac{5,7689 + 5,2832 + 6,5083}{19} = 0,924$$

Comprobamos en la tabla K (Anexo I) que para  $r_w = 0,924$ ,  $K = 1,0546$ . Ahora:

$$F = K \frac{(N - k) \cdot (\sum_{j=1}^k R_j - R)}{(k - 1) - (N - \sum_{j=1}^k R_j)} =$$

$$F = 1,0546 \frac{(19 - 3) \cdot (5,7689 + 5,2832 + 6,5083 - 17,2423)}{(3 - 1) - (19 - 5,7689 + 5,2832 - 6,5083)} = 1,86$$

El valor crítico es  $F_{0,05(1),2,16} = 3,63$  (tabla H, Anexo I), por tanto no rechazamos  $H_0$  concluyendo que las tres muestras poseen la misma media, que podemos estimar a partir de:

$$\text{sen } \bar{\alpha} = \frac{Y}{r} = 0,3523; \quad \text{cos } \bar{\alpha} = \frac{X}{r} = -0,9358$$

Por tanto:

$$\bar{\alpha} = \arccos(-0,9358) = 159,35^\circ$$

Como se ha dicho anteriormente este test es paramétrico, por tanto se supone que las muestras proceden de una población que se ajusta a una distribución de von Mises.

Afortunadamente este test es muy robusto, pero si las condiciones de normalidad son severamente violadas es aconsejable utilizar un test no paramétrico.

#### 4.3.2. Test de Mardia-Watson-Wheeler

**Datos.** Dos muestras de  $n_1$  y  $n_2$  con valores angulares  $\phi_1, \phi_2 \dots \phi_n$  y  $\gamma_1, \gamma_2 \dots \gamma_n$  que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si las dos muestras difieren significativamente una de la otra. Las diferencias pueden ser en el ángulo medio, en la varianza angular o en los dos parámetros. Se trata de un test no paramétrico en el que lo importante es el orden en el que se disponen los ángulos. El proceso implica alterar la separación existente entre los ángulos para situarlos equidistantemente. Para calcular la separación entre ángulos recalculamos:

$$\delta = \frac{360^\circ}{n_1 + n_2}$$

**Ecuaciones.** Una vez tengamos los datos de la muestra combinada equidistantemente separados les asignamos un rango (1, 2, 3, 4...n) a partir de una dirección arbitraria. Si  $c_1, c_2 \dots c_n$  son los rangos de la primera muestra los ángulos correspondientes a cada punto de esta muestra son:

$$\beta_i = c_i \cdot \delta$$

El vector medio de la primera muestra tiene los siguientes componentes:

$$x_1 = \sum_i^n \cos \phi_i; \quad y_1 = \sum_i^n \sin \phi_i$$

Con estas coordenadas se obtiene la longitud del vector medio,  $r$ :

$$r_1 = \sqrt{x_1^2 + y_1^2}$$

Para la segunda muestra se procedería del mismo modo hasta obtener  $r_2$ .

#### Consideraciones sobre los datos:

- a) No debe haber datos empatados (valores que se repitan en las dos muestras).

#### Hipótesis nula:

$H_0$ : Las dos muestras proceden de la misma población.

$H_A$ : Las dos muestras no proceden de la misma población.

#### Estadístico:

$$B = r^2$$

Debe cumplirse que  $r_1 = r_2 = r$ .

**Regla de decisión:**

Debemos consultar en una tabla los valores de  $B(\alpha)$  (tabla J, Anexo I) cuando  $n = 8, 9, \dots, 17$ . Si  $n > 17$  calculamos  $P$  en la tabla de chi-cuadrado. Si  $B > B(\alpha)$  rechazamos la hipótesis nula y concluimos que las dos muestras difieren significativamente.

**Ejercicio 23** (tomado de Batschelet, 1981). Disponemos dos muestras y queremos saber si difieren o no.

Muestra 1	Muestra 2
20°	125°
40°	190°
80°	230°
95°	275°
140°	
155°	

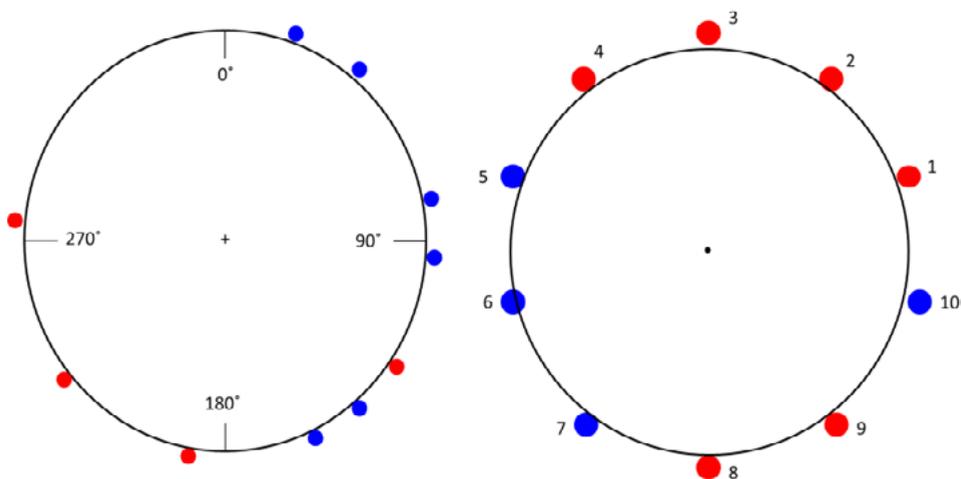
Ahora:

$$n = n_1 + n_2 = 6 + 4 = 10$$

Por tanto:

$$\delta = \frac{360^\circ}{n_1 + n_2} = \frac{360^\circ}{10} = 36^\circ$$

Ahora asignamos rangos y vemos que la muestra 1 tiene rangos 1, 2, 3, 4, 8 y 9 (Fig. 64):



**Figura 64.** Datos de la muestra 1 (azul) y 2 (rojo) en la imagen de la izquierda. A la derecha los rangos asignados a cada punto de la muestra.

Ahora calculamos los ángulos para cada punto empleando la expresión:

$$\beta_i = c_i \cdot \delta$$

Y tenemos que:

$$\beta_1 = 1 \cdot 36^\circ = 36^\circ, \beta_2 = 2 \cdot 36^\circ = 72^\circ, \beta_3 = 3 \cdot 36^\circ = 108^\circ, \beta_4 = 4 \cdot 36^\circ = 144^\circ, \\ \beta_5 = 8 \cdot 36^\circ = 288^\circ, \beta_6 = 9 \cdot 36^\circ = 324^\circ$$

Con estos nuevos ángulos calculamos la longitud del vector medio,  $r$ :

En primer lugar obtenemos la siguiente tabla:

<i>Muestra 1</i>	<i>Coseno</i>	<i>Seno</i>
36°	0,8090	0,5878
72°	0,3090	0,9511
108°	-0,3090	0,9511
144°	-0,8090	0,5878
288°	0,3090	-0,9511
324°	0,8090	-0,5878
Suma:	$x_1 = 1,1180$	$y_1 = 1,5388$

Ahora calculamos

$$r_1 = \sqrt{x_1^2 + y_1^2} = \sqrt{1,25 + 2,3680} = \sqrt{3,618} = 1,9021$$

Ya podemos calcular el estadístico para la muestra 1:

$$B = r_1^2 = (1,9021)^2 = 3,6180$$

La segunda muestra tiene rangos 5, 6 7 y 10. Por lo que sus ángulos son:

$$\beta_7 = 5 \cdot 36^\circ = 180^\circ, \beta_8 = 6 \cdot 36^\circ = 216^\circ, \beta_9 = 7 \cdot 36^\circ = 252^\circ, \beta_{10} = 10 \cdot 36^\circ = 360^\circ$$

Por tanto tendríamos:

<i>Muestra 2</i>	<i>Coseno</i>	<i>Seno</i>
180°	-1	1,2E-16
216°	-0,8090	-0,5878
252°	-0,3090	-0,9511
360°	1	-2,5E-16
Suma:	$x_2 = -1,1180$	$y_2 = -1,5388$

Y obtenemos que  $r_2 = 1,9021$ . Por tanto:

$$B = r_2^2 = 1,9021^2 = 3,6180$$

Vemos que  $r_1^2 = r_2^2$ . De las tablas obtenemos que  $B(0,05) = 9,47$ . Como  $B < B(0,05)$  aceptamos la hipótesis nula.

Zar (1996) recoge la modificación del test de test de Wheeler-Watson realizada por Mardia (que aquí hemos llamado Test de Mardia-Whatson-Wheeler), que implica asignar rangos a cada uno de los  $\alpha$  ángulos de la muestra, de modo que podemos calcular el rango circular ( $d$ ):

$$d = \frac{360^\circ \cdot (\text{rango de } \alpha)}{N}$$

Entonces:

$$C_i = \sum_{j=1}^{n_i} \cos d_j$$

y

$$S_i = \sum_{j=1}^{n_i} \text{sen } d_j$$

El estadístico a utilizar es:

$$W = 2 \cdot \left[ \frac{C_1^2 + S_1^2}{n_1} + \frac{C_2^2 + S_2^2}{n_2} \right]$$

que debe ser comparado con una chi-cuadrado de 2 grados de libertad.

Este test es aplicable si los tamaños muestrales son como mínimo de 10 valores y si no hay datos repetidos en las dos muestras. Si esto ocurre es recomendable utilizar el Test  $U^2$  de Watson.

**Ejemplo 24** (tomado de Zar, 1996). Tenemos dos series de datos (Fig. 65) listados en la tabla que se detalla a continuación y deseamos determinar:

$H_0$ : Las dos muestras proceden de la misma población o tienen la misma dirección.

$H_A$ : Las dos muestras no proceden de la misma población o no tienen la misma dirección.

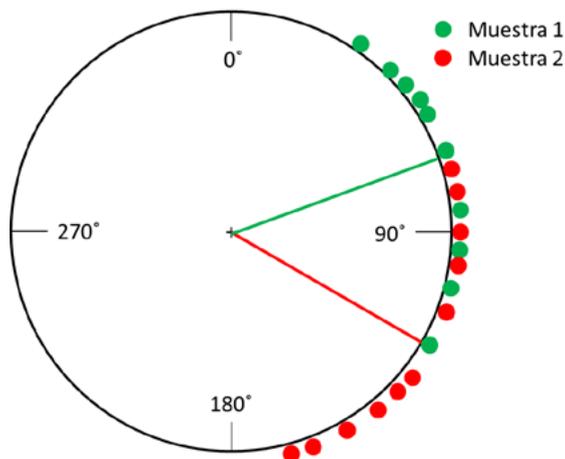


Figura 65. Representación de los datos del ejercicio 24.

Aplicando las fórmulas anteriores calculamos  $d$ ,  $C_i$  y  $S_i$ :

<i>Muestra 1</i>					<i>Muestra 2</i>				
<i>Dirección</i>	<i>Rango</i>	<i>Rango Circular</i>	$C_1$	$S_1$	<i>Dirección</i>	<i>Rango</i>	<i>Rango Circular</i>	$C_2$	$S_2$
35°	1	17,14°	0,96	0,29					
45°	2	34,29°	0,83	0,56					
50°	3	51,43°	0,62	0,78					
55°	4	68,57°	0,37	0,93					
60°	5	85,71°	0,07	1					
70°	6	102,86°	-0,22	0,97					
					75°	7	120°	-0,50	0,87
					80°	8	137,14°	-0,73	0,68
85°	9	154,29°	-0,90	0,43					
					90°	10	171,43°	-0,99	0,15
95°	11	188,57°	-0,99	-0,15					
					100°	12	205,71°	-0,90	-0,43
105°	13	222,86°	-0,73	-0,68					
					110°	14	240°	-0,50	-0,87
120°	15	257,14°	-0,22	-0,97					
					130°	16	274,29°	0,07	-1
					135°	17	291,43°	0,37	-0,93
					140°	18	308,57°	0,62	-0,78
					150°	19	325,71°	0,83	-0,56
					160°	20	342,86°	0,96	-0,29
					165°	21	360°	1	0
Suma:			-0,2225	3,1727				0,2225	-3,1727

Con estos datos podemos calcular  $W$ :

$$W = 2 \cdot \left[ \frac{C_1^2 + S_1^2}{n_1} + \frac{C_2^2 + S_2^2}{n_2} \right] =$$

$$W = 2 \cdot \left[ \frac{-0,2225^2 + 3,1717^2}{10} + \frac{0,2225^2 + -3,1727^2}{11} \right] = 3,862$$

Como  $\chi_{0,05,2}^2 = 5,991$ , aceptamos  $H_0$ .

También existe una versión de este test para más de dos muestras. En este caso el estadístico es:

$$W = 2 \sum_{i=1}^k \frac{C_i^2 + S_i^2}{n_i}$$

que debe ser comparado con una chi-cuadrado de 2 (k-1) grados de libertad. Tiene las mismas restricciones que para el caso de dos muestras.

### 4.3.3. Test $U^2$ de Watson

**Datos.** Dos muestras de  $m$  y  $n$  con valores angulares  $\phi_1, \phi_2 \dots \phi_n$  y  $\gamma_1, \gamma_2 \dots \gamma_n$  que representan direcciones en un plano, tiempo o un fenómeno cíclico.

**Propósito.** Determinar si las dos muestras difieren significativamente una de la otra en el ángulo medio, en la varianza angular u otro parámetro.

**Ecuaciones.** Este test también ser basa en las diferencias entre las dos funciones para cada par de datos de cada una de las distribuciones. El estadístico empleado es:

$$U^2 = \frac{n \cdot m}{N^2} \left[ \sum_{k=1}^N d_k^2 - \frac{1}{N} \left( \sum_{k=1}^N d_k \right)^2 \right]$$

Donde  $n$  y  $m$  son los tamaños muestrales de los dos grupos de datos que se quiere comparar,  $N = n + m$  y  $d$  son las diferencias entre las dos funciones para cada uno de los  $N$  puntos considerados.

#### Consideraciones sobre los datos:

- a) Los datos deben estar ordenados de forma ascendente. Es un test no paramétrico, por lo tanto debe emplearse cuando una de las muestras no es unimodal o no se cumple cualquier otro de los criterios que implican normalidad.
- b) Se puede emplear con datos agrupados siempre y cuando la agrupación no sea mayor de 5°.

**Hipótesis nula:**

$H_0$ : Las dos muestras proceden de la misma población o de dos poblaciones que tienen la misma dirección.

$H_A$ : Las dos muestras no proceden de la misma población o de dos poblaciones que tienen distinta dirección.

**Regla de decisión:**

$H_0$  se rechaza si el valor obtenido en las tablas  $U^2_{\alpha,m,n}$  es menor que el arrojado por el estadístico.

**Ejercicio 25** (modificado de Zar, 1996). Tenemos dos poblaciones de aves que se han dispersado según los datos presentados en la siguiente tabla, y queremos saber si lo han hecho en la misma dirección.

$i$ ( $m = 10$ )	$\varphi_i$	$i/m$	$j$ ( $n = 11$ )	$\omega_j$	$j/n$	$d_k = i/m - j/n$	$d_k^2$
1	35°	0,1			0	0,1	0,01
2	45°	0,2			0	0,2	0,04
3	50°	0,3			0	0,3	0,09
4	55°	0,4			0	0,4	0,16
5	60°	0,5			0	0,5	0,25
6	70°	0,6			0	0,6	0,36
		0,6	1	75°	0,0909	0,5091	0,2592
		0,6	2	80°	0,1818	0,4182	0,1749
7	85°	0,7			0,1818	0,5182	0,2685
		0,7	3	90°	0,2727	0,4273	0,1826
8	95°	0,8			0,2727	0,5273	0,2780
		0,8	4	100°	0,3636	0,4364	0,1904
9	105°	0,9			0,3636	0,5364	0,2877
		0,9	5	110°	0,4545	0,4455	0,1984
10	120°	1			0,4545	0,5455	0,2976
		1	6	130°	0,5455	0,4545	0,2066
		1	7	135°	0,6364	0,3636	0,1322
		1	8	140°	0,7273	0,2727	0,0744
		1	9	150°	0,8182	0,1818	0,0331
		1	10	155°	0,9091	0,0909	0,0083
		1	11	165°	1	0	0
					Suma:	7,8274	3,5019

Sabemos que  $N = m + n = 10 + 11 = 21$ . Por tanto:

$$U^2 = \frac{n \cdot m}{N^2} \left[ \sum_{k=1}^N d_k^2 - \frac{1}{N} \left( \sum_{k=1}^N d_k \right)^2 \right] = \frac{11 \cdot 10}{21^2} \left[ 3,5019 - \frac{1}{21} (7,8274)^2 \right] = 0,1558$$

Como  $U_{0,05,10,11}^2 = 0,1856$ , aceptamos  $H_0$ : los dos grupos de aves siguen la misma dirección (Fig. 66).

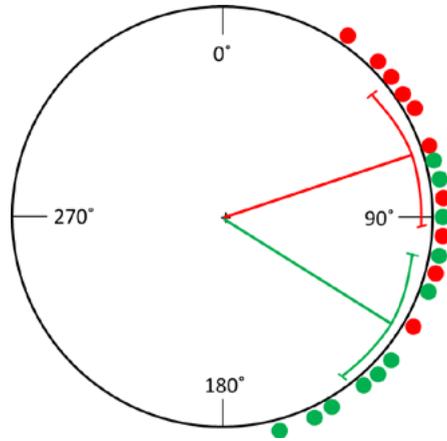


Figura 66. Representación gráfica de los datos del ejercicio 25 (en rojo datos de la muestra 1, en verde datos de la muestra 2).

#### 4.3.4. Test de la chi-cuadrado

**Datos.** Dos muestras independientes de observaciones circulares con datos agrupados.

**Propósito.** Determinar si existen diferencias entre las dos muestras. No se especifica qué tipo de diferencias. Es un test muy poderoso para tamaños muestrales grandes.

**Ecuaciones.** Partimos de los siguientes datos:

Muestra	Categorías (k)				Total
S1	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$	$n_{1,k}$	$M_1$
S2	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$	$n_{2,k}$	$M_2$
Total	$N_1$	$N_2$	$N_3$	$N_k$	$N$

Las frecuencias esperadas serán:

$$e_{ij} = \frac{N_j \cdot M_i}{N}$$

Con  $i = 1, 2, \dots$  y  $j = 1, 2, \dots, k$

El estadístico de la chi-cuadrado es:

$$\chi^2 = \sum_{ij} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

**Consideraciones sobre los datos:**

- a) Las frecuencias esperadas deben ser, como mínimo, 5.

**Hipótesis nula:**

$H_0$ : Las muestras son iguales o proceden de la misma población.

$H_A$ : Las muestras son no iguales o no proceden de la misma población.

**Regla de decisión:**

Si el valor crítico de  $P$  obtenido en las tablas es menor que el nivel de significancia elegido,  $\alpha$  y con  $n-1$  grados de libertad, la hipótesis nula se rechaza.

**Ejercicio 26** (tomado de Batschelet, 1989). Se tienen las direcciones que toman dos poblaciones de tortugas y se desea saber si se orientan de igual forma.

Intervalos	Frecuencia observada		Total	Frecuencia esperada		Fr. obs - Fr. esp		$(Fr. obs - Fr. esp)^2 / Fr. esp$	
	S1	S2		S1	S2	S1	S2	S1	S2
0°-45°	0	0	0	0	0	0	0		
45°-90°	4	15	19	10,2	8,8	-6,2	6,2	3,8	4,3
90°-135°	0	27	27	14,5	12,5	-14,5	14,5	14,5	16,7
135°-180°	17	21	38	20,4	17,6	-3,4	3,4	0,6	0,6
180°-225°	28	5	33	17,7	15,3	10,3	-10,3	6,0	6,9
225°-271°	11	0	11	5,9	5,1	5,1	-5,1	4,4	5,1
270°-315°	15	0	15	8,0	7,0	7,0	-7,0	6,0	7,0
315°-360°	7	3	10	5,4	4,6	1,6	-1,6	0,5	0,6
Total	82	71	153	82,0	71,0	0,0	0,0	35,7	41,3

Por tanto:

$$\chi_2^2 = 35,7 + 41,4 = 77,3$$

Como  $\chi_{0,05,7}^2 = 14,45$  rechazamos  $H_0$ : las orientaciones son diferentes.

**4.3.5. Test de Hotelling para dos muestras apareadas**

**Datos.** Dos muestras de datos apareados.

**Propósito.** Determinar si existen diferencias entre las dos muestras.

**Ecuaciones.** Se emplean los mismos procedimientos que en el test de Hotelling visto con anterioridad, ya que podemos obtener una sola muestra al poder calcular las diferencias entre datos apareados. Si  $\alpha_{ij}$  es el ángulo  $i$  del dato  $j$ , entonces  $\alpha_{1j}$  y  $\alpha_{2j}$  son datos apareados, entonces podemos obtener las coordenadas rectangulares de estos datos del siguiente modo:

$$X_j = \cos \alpha_{2j} - \cos \alpha_{1j}$$

$$Y_j = \sen \alpha_{2j} - \sen \alpha_{1j}$$

Ya podemos obtener el estadístico  $F$ :

$$F = \frac{k \cdot (k - 2)}{2} \left[ \frac{\bar{X}^2 \sum y^2 - 2\bar{X} \cdot \bar{Y} \sum x \cdot y + \bar{Y}^2 \sum x^2}{\sum x^2 \sum y^2 - (\sum x \cdot y)^2} \right]$$

**Consideraciones sobre los datos:**

a) Los datos deben estar apareados.

**Hipótesis nula:**

$H_0$ : No hay diferencias entre las dos muestras.

$H_A$ : Hay diferencias entre las dos muestras.

**Regla de decisión:**

Si el valor crítico obtenido en las tablas es menor que el nivel de significancia elegido la hipótesis nula se rechaza.

**Ejercicio 27** (tomado de Zar, 1996). Deseamos saber si existen diferencias en la zona donde descansan un grupo de aves por la mañana y por la tarde.

Hacemos los cálculos:

Ave (j)	Mañana			Tarde			Diferencia	
	Dirección ( $\alpha_{1j}$ )	sen $\alpha_{1j}$	cos $\alpha_{1j}$	Dirección ( $\alpha_{2j}$ )	sen $\alpha_{2j}$	cos $\alpha_{2j}$	$Y_j$	$X_j$
1	105°	0,9659	-0,2588	205°	-0,4226	-0,9063	-1,3885	-0,6475
2	120°	0,8660	-0,5000	210°	-0,5000	-0,8660	-1,3660	-0,3660
3	135°	0,7071	-0,7071	235°	-0,8192	-0,5736	-1,5263	0,1335
4	95°	0,9962	-0,0872	245°	-0,9063	-0,4226	-1,9025	-0,3355
5	155°	0,4226	-0,9063	260°	-0,9848	-0,1736	-1,4074	0,7327
6	170°	0,1736	-0,9848	255°	-0,9659	-0,2588	-1,1396	0,7260
7	160°	0,3420	-0,9397	240°	-0,8660	-0,5000	-1,2080	0,4397
8	155°	0,4226	-0,9063	245°	-0,9063	-0,4226	-1,3289	0,4837
9	120°	0,8660	-0,5000	210°	-0,5000	-0,8660	-1,3660	-0,3660
10	115°	0,9063	-0,4226	200°	-0,3420	-0,9397	-1,2483	-0,5171

Tenemos que:

$$k = 10; \quad \bar{X} = 0,0284; \quad \bar{Y} = -1,3981;$$

Con esos datos y siguiendo el mismo procedimiento que en test de Hotelling para una muestra llegamos a que:

$F = 217$  y como  $F_{0,05,2,8} = 4,46$  rechazamos  $H_0$ . Hay diferencias en las zonas de descanso.

#### 4.3.6. Test de Moore para dos muestras

**Datos.** Dos muestras de  $n_1$  y  $n_2$  valores angulares  $\phi_1, \phi_2, \dots, \phi_{n_r}$  que representan direcciones en un plano, tiempo o un fenómeno cíclico. Cada valor se ha medido sobre el mismo individuo.

**Propósito.** Determinar si existen diferencias entre ambas muestras. Se trata de un test no paramétrico.

**Ecuaciones.** Si tenemos dos muestras apareadas, por tanto con  $k$  pares, calculamos para cada dato su seno y su coseno para obtener las coordenadas rectangulares. Así las coordenadas rectangulares del par  $j$  serán:

$$X_j = \cos \alpha_{2j} - \cos \alpha_{1j}$$

$$Y_j = \sin \alpha_{2j} - \sin \alpha_{1j}$$

A continuación para cada diferencia calculamos:

$$r_j = \sqrt{X_j^2 + Y_j^2}$$

$$\cos \alpha_j = \frac{X_j}{r_j}; \quad \sin \alpha_j = \frac{Y_j}{r_j}$$

Seguidamente asignamos rangos ( $i$ ) a  $r_j$ , desde 1 hasta  $n$  y calculamos:

$$X = \frac{\sum_{i=1}^k i \cdot \cos \bar{\alpha}_i}{k}$$

$$Y = \frac{\sum_{i=1}^k i \cdot \sin \bar{\alpha}_i}{k}$$

y

$$R' = \frac{\sqrt{X^2 + Y^2}}{k}$$

#### Consideraciones sobre los datos:

a) Los datos deben estar apareados.

**Hipótesis nula:**

$H_0$ : Los ángulos son iguales.

$H_A$ : Los ángulos son diferentes.

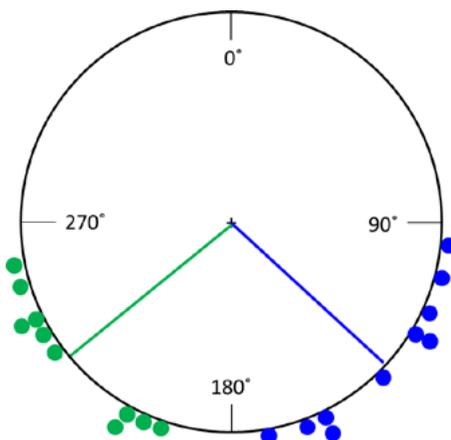
**Regla de decisión:**

Si el valor crítico es menor que el nivel de significancia elegido la hipótesis nula se rechaza.

**Ejercicio 28** (tomado de Zar, 1996). Se marcaron una serie de aves y se determinó la orientación a que descansaban por la mañana y por la tarde. Se pretende determinar si hay diferencias entre la mañana y la tarde a la hora de elegir la zona de descanso.

<i>Ave</i>	<i>Mañana</i>	<i>Tarde</i>
1	105°	205°
2	120°	210°
3	135°	235°
4	95°	245°
5	155°	260°
6	170°	255°
7	160°	240°
8	155°	245°
9	120°	210°
10	115°	200°

En la figura 67 se representan los datos correspondientes a los dos momentos del día.



**Figura 67.** Datos del ejercicio 28. En azul posiciones ocupadas por las aves durante la mañana. En verde posiciones ocupadas durante la tarde.

En primer lugar calculamos las coordenadas rectangulares de cada par:

Ave (j)	Dirección ( $\alpha_{1j}$ )	Mañana		Dirección ( $\alpha_{2j}$ )	Tarde		Diferencia	
		sen $\alpha_{1j}$	cos $\alpha_{1j}$		sen $\alpha_{2j}$	cos $\alpha_{2j}$	$Y_j$	$X_j$
1	105°	0,9659	-0,2588	205°	-0,4226	-0,9063	-1,3885	-0,6475
2	120°	0,8660	-0,5000	210°	-0,5000	-0,8660	-1,3660	-0,3660
3	135°	0,7071	-0,7071	235°	-0,8192	-0,5736	-1,5263	0,1335
4	95°	0,9962	-0,0872	245°	-0,9063	-0,4226	-1,9025	-0,3355
5	155°	0,4226	-0,9063	260°	-0,9848	-0,1736	-1,4074	0,7327
6	170°	0,1736	-0,9848	255°	-0,9659	-0,2588	-1,1396	0,7260
7	160°	0,3420	-0,9397	240°	-0,8660	-0,5000	-1,2080	0,4397
8	155°	0,4226	-0,9063	245°	-0,9063	-0,4226	-1,3289	0,4837
9	120°	0,8660	-0,5000	210°	-0,5000	-0,8660	-1,3660	-0,3660
10	115°	0,9063	-0,4226	200°	-0,3420	-0,9397	-1,2483	-0,5171

Continuando con los cálculos obtenemos:

$Y_{2j}$	$X_{2j}$	$Y_{2j} + X_{2j}$	$r$	sen $\alpha_j$	cos $\alpha_j$	rango (i)	cos $\alpha_j \cdot i$	sen $\alpha_j \cdot i$
1,9281	0,4192	2,3473	1,5321	-0,9063	-0,4226	7,5	-3,170	-6,797
1,8660	0,1340	2,0000	1,4142	-0,9659	-0,2588	4,5	-1,165	-4,347
2,3295	0,0178	2,3473	1,5321	-0,9962	0,0872	7,5	0,654	-7,471
3,6195	0,1125	3,7321	1,9319	-0,9848	-0,1736	10,0	-1,736	-9,848
1,9808	0,5368	2,5176	1,5867	-0,8870	0,4617	9,0	4,156	-7,983
1,2986	0,5271	1,8257	1,3512	-0,8434	0,5373	2,0	1,075	-1,687
1,4594	0,1933	1,6527	1,2856	-0,9397	0,3420	1,0	0,342	-0,940
1,7660	0,2340	2,0000	1,4142	-0,9397	0,3420	6,0	2,052	-5,638
1,8660	0,1340	2,0000	1,4142	-0,9659	-0,2588	4,5	-1,165	-4,347
1,5583	0,2674	1,8257	1,3512	-0,9239	-0,3827	3	-1,148	-2,772

Así pues tendremos que:

$$X = \frac{\sum_{i=1}^k i \cdot \cos \bar{\alpha}_i}{k} = \frac{-0,106}{10} = -0,011$$

$$Y = \frac{\sum_{i=1}^k i \cdot \sin \bar{\alpha}_i}{k} = \frac{-51,83}{10} = -5,183$$

Por tanto:

$$R' = \frac{\sqrt{X^2 + Y^2}}{k} = \frac{\sqrt{(-0,011)^2 + (-5,183)^2}}{10} = 1,639$$

Como  $R'_{0,05,10} = 1,144$  rechazamos  $H_0$ , por tanto concluimos que la dirección en la que descansan las aves durante la mañana y durante la tarde es diferente.

#### 4.3.7. Test no paramétrico de Wallraff

Este test basado en Wallraff (1979) permite testar la homocedasticidad entre dos muestras, es decir, si las muestras tienen la misma concentración (o dispersión). El procedimiento consiste en calcular la dirección media de cada muestra y las distancias angulares entre los datos de cada muestra y su dirección media. A continuación se asignan rangos a los resultados y se aplica el test de Mann-Whitney (dos muestras) o el test de Kruskal-Wallis (más de dos muestras).

**Ejercicio 29** (tomado de Zar, 1996). En la siguiente tabla se presentan datos relativos a la hora de nacimiento de niños y niñas durante un determinado día. También se ha calculado el ángulo medio: niños,  $\alpha_1 = 7:55$  AM; niñas,  $\alpha_{12} = 8:15$  AM. Con estos datos la hipótesis que se plantea es la siguiente:

$H_0$ : las horas de nacimientos de los niños son tan variables como las de las niñas.

$H_A$ : las horas de nacimiento de los niños no son tan variables como las de las niñas.

Para cada muestra calculamos la distancia angular y asignamos rango a cada uno de los valores.

Niños			Niñas		
Hora	Distancia angular	Rango	Hora	Distancia angular	Rango
5:10:00	2:45:00	11	8:15:00	0:00:00	1
6:30:00	1:25:00	4	10:20:00	2:05:00	8,5
9:40:00	1:45:00	6	9:45:00	1:30:00	5
10:20:00	2:25:00	10	6:10:00	2:05:00	8,5
4:20:00	3:35:00	13	4:05:00	4:10:00	14
11:15:00	3:20:00	12	7:50:00	0:25:00	2
			9:00:00	0:45:00	3
			10:10:00	1:55:00	7

De modo que tenemos:  $n_1 = 6$ ,  $R_1 = 56$ ,  $n_2 = 8$ ,  $R_2 = 49$ , con lo que podemos aplicar el test de Mann-Whitney:

$$U = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 6 \cdot 8 + \frac{6(6 + 1)}{2} - 56 = 13$$

Con este resultado  $P = 0,17$ , por tanto no rechazamos  $H_0$ .

#### 4.4. ANOVA CIRCULAR

Una ANOVA Circular es un análisis de la varianza de una vía que permite testar la igualdad de direcciones medias en una investigación con datos circulares. Para que pueda realizarse una ANOVA circular es necesario que las  $n$ -muestras independientes objeto de análisis se ajusten a una distribución de von Mises y que posean medias direccionales y parámetros de concentración iguales (Tasdan y Yenisey, 2014). En caso de que alguno o todos estos condicionantes no se cumplan nos veremos obligados a utilizar test no paramétricos para poder testar la igualdad de medias.

Vamos a construir la tabla para la ANOVA circular siguiendo el ejemplo propuesto por Firdu y Başpınar (2017). Esta tabla es similar a la utilizada en el caso de la ANOVA lineal. Como sabemos para poder aplicar una ANOVA lineal la muestra debe ajustarse a la distribución uniforme. En el caso de la ANOVA circular hemos señalado que la muestra debe ajustarse a una distribución de von Mises. Ya vimos cómo podemos comprobar si nuestras muestras se ajustan a una distribución de von Mises.

Firdu y Başpınar (2017) estudian la incidencia mensual de la fiebre hemorrágica Crimeas-Congo en el periodo 2008-2014 en Turquía. Los datos aportados, por tanto, son frecuencias mensuales (de enero a diciembre) medidos en una escala circular.

La tabla de ANOVA que utilizaremos para la realización de los cálculos es la siguiente:

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Suma de los cuadrados</i>	<i>Media cuadrática</i>
Entre muestras	$p-1$	SSB	MSB
Dentro de muestras	$\sum n - p$	SSW	MSW
Total	$\sum n - 1$	SST	

Siendo  $p$  el número total de muestras o grupos y  $n$  los tamaños muestrales de cada grupo. Para calcular el estadístico debemos obtener previamente los siguientes resultados:

$$df = p - 1$$

$$SSB = \sum_{i=1}^p (R_i - \bar{R})$$

$$MSB = \frac{\sum_{i=1}^p (R_i - \bar{R})}{p - 1}$$

$$SSW = \sum_{i=1}^p (n_i - \bar{R})$$

$$SST = n - \bar{R}$$

$$MSW = \frac{\sum_{i=1}^p (n_i - R_i)}{\sum n_i - p}$$

El valor de F se calcula de acuerdo a la siguiente expresión:

$$F = \frac{\sum_{i=1}^p (R_i - \bar{R}) / (p - i)}{\sum_{i=1}^p (n_i - R_i) / \sum n_i - p}$$

Los datos que aportan Firdu y Başpınar (2017) son los siguientes:

Año	$n_i$	$\tilde{\phi}$	$\bar{R}_i$	$\kappa$	Varianza circular	Desviación estándar circular
2008	1273	2,7329	0,8567	0,3325	0,6675	0,9780
2009	1277	2,8633	0,8402	0,3265	0,6734	0,9859
2010	835	2,6477	0,8476	0,2488	0,7511	1,0992
2011	1053	2,9268	0,8447	0,2903	0,7097	1,0336
2012	777	2,7217	0,8206	0,2287	0,7713	1,1320
2013	863	2,6789	0,8305	0,2493	0,7507	1,0985
2014	931	2,1658	0,8001	0,2530	0,7469	1,0926
2015	684	2,9139	0,8285	0,2099	0,7900	1,1640

Siendo  $\bar{R} = 0,580445$  y  $n$  total igual a 7693. Utilizando una tabla de Excel podemos hacer los cálculos fácilmente:

Año	$n_i$	$\bar{R}_i$	$\bar{R}$	$\bar{R}_i - \bar{R}$	$n_i - \bar{R}$	$n_i - \bar{R}_i$	$n_i - p$
2008	1273	0,8567	0,580445	0,2762	1272,42	1272,1433	1272
2009	1277	0,8402	0,580445	0,2597	1276,42	1276,1598	1275
2010	835	0,8476	0,580445	0,2671	834,42	834,1524	832
2011	1053	0,8447	0,580445	0,2642	1052,42	1052,1553	1049
2012	777	0,8206	0,580445	0,2401	776,42	776,1794	772
2013	863	0,8305	0,580445	0,2500	862,42	862,1695	857
2014	931	0,8001	0,580445	0,2196	930,42	930,1999	924
2015	684	0,8285	0,580445	0,24805	683,42	683,1715	676
Total:	7693			2,02534	7688,36	7686,3311	7657

A partir de estos datos podemos calcular parámetros de la ANOVA circular:

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Suma de los cuadrados</i>	<i>Media cuadrática</i>
Entre muestras	$p - 1 = 8 - 1 = 7$	$SSB = \sum_{i=1}^p (R_i - \bar{R}) = 2,02534$	$MSB = \frac{\sum_{i=1}^p (R_i - \bar{R})}{p - 1} = \frac{2,02354}{7} = 0,289334$
Dentro de muestras	$\sum n - p = 7693 - 8 = 7685$	$SSW = \sum_{i=1}^p (n_i - \bar{R}) = 7688,36$	$MSW = \frac{\sum_{i=1}^p (n_i - R_i)}{\sum n_i - p} = \frac{7688,36}{7657} = 1,00038$
Total	$\sum n - 1 = 7693 - 1 = 7692$	$SST = n - \bar{R} = 7693 - 0,580445 = 7692,42$	

Finalmente:

$$F = \frac{\sum_{i=1}^p (R_i - \bar{R}) / (p - i)}{\sum_{i=1}^p (n_i - R_i) / \sum n_i - p} = \frac{0,289334}{1,00038} = 0,2882$$

Utilizando Excel (=DISTR.F(valor de F obtenido; p-1; n-1)) podemos calcular el valor teórico para  $F_{0,2882;7;7692} = 0,9588$ . Como  $0,9588 > 0,2882$  concluimos que no existen diferencias en la incidencia de la enfermedad en los diferentes años estudiados.

# CAPÍTULO 5

## CORRELACIÓN

### 5.1. INTRODUCCIÓN

En este capítulo se analizan las relaciones entre dos variables circulares aleatorias y entre una variable circular aleatoria y una lineal. Por ejemplo nos podría interesar estudiar si existe alguna relación entre la dirección del viento (circular) y las direcciones que toman las ramas de un árbol (circular) expuesto al citado viento. Otro caso interesante resultaría de investigar la relación que existe entre la dirección del viento (circular) y la cantidad de lluvia caída (lineal).

### 5.2. CORRELACIÓN CIRCULAR-CIRCULAR

Partimos de dos variables circulares donde denotaremos a cada par  $i$  de datos como  $a_i$  y  $b_i$ , donde  $a$  y  $b$  son ángulos. Entonces, el coeficiente de correlación es:

$$r_{aa} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}(a_i - a_j) \text{sen}(b_i - b_j)}{\sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}^2(a_i - a_j) \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}^2(b_i - b_j)}}$$

Para determinar si  $r_{aa}$  es significativo se estima  $n \cdot r_{aa}$  adicionales, eliminando cada vez uno de los  $n$  pares de  $a$  y  $b$ , técnica conocida como *jackknife*. A continuación se calcula la media y la varianza de estos  $r_{aa}$  para estimar los límites de confianza para  $\rho_{aa}$  que son:

$$L_1 = n r_{aa} - (n-r) \bar{r}_{aa} - Z_{\alpha(2)} \sqrt{\frac{S_{r_{aa}}^2}{n}}$$

$$L_2 = n r_{aa} - (n-r) \bar{r}_{aa} - Z_{\alpha(2)} \sqrt{\frac{S_{r_{aa}}^2}{n}}$$

Si el intervalo de confianza, es decir, el intervalo entre  $L_1$  y  $L_2$  no incluye al cero, entonces  $H_0 = \rho_{aa} = 0$  es rechazada en favor de  $H_A = \rho_{aa} \neq 0$ .

**Ejercicio 30** (tomado de Zar, 1996). Deseamos saber la relación entre la orientación de los insectos y la dirección de la luz, según la siguiente tabla:

$i$	Insecto $a_i$	luz $b_i$
1	145°	120°
2	190°	180°
3	310°	330°
4	210°	225°
5	80°	55°

Cuya representación gráfica se presenta en la figura 68:

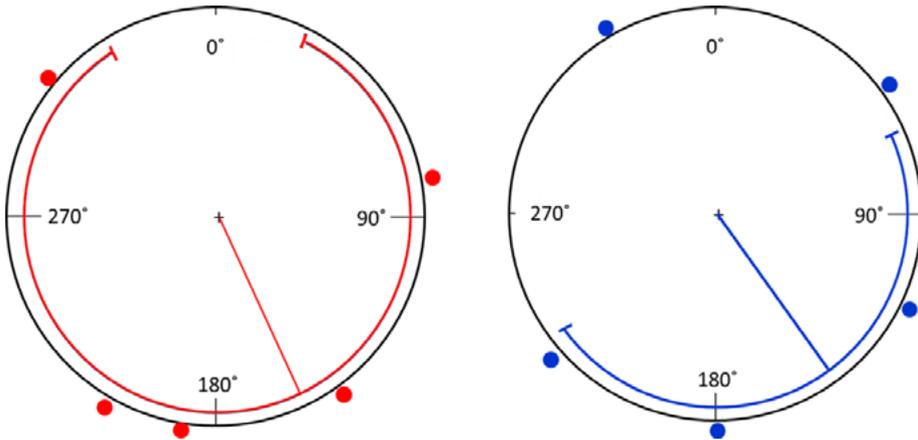


Figura 68. Dirección que toman los insectos (izquierda) y de la luz (derecha).

Por tanto:  $H_0 = \rho_{aa} = 0$  y  $H_A = \rho_{aa} \neq 0$ .

El procedimiento a seguir es el siguiente:

$i$	$j$	$a_i - a_j$	$b_i - b_j$	$\text{sen}(a_i - a_j)$	$\text{sen}(b_i - b_j)$	$\text{sen}(a_i - a_j) \cdot \text{sen}(b_i - b_j)$	$\text{sen}^2(a_i - a_j)$	$\text{sen}^2(b_i - b_j)$
1	2	-45	-60	-0,7071	-0,8660	0,6124	0,5000	0,7500
1	3	-165	-210	-0,2588	0,5000	-0,1294	0,0670	0,2500
1	4	-65	-105	-0,9063	-0,9659	0,8754	0,8214	0,9330
1	5	65	65	0,9063	0,9063	0,8214	0,8214	0,8214
2	3	-120	-150	-0,8660	-0,5000	0,4330	0,7500	0,2500
2	4	-20	-45	-0,3420	-0,7071	0,2418	0,1170	0,5000
2	5	110	125	0,9397	0,8192	0,7698	0,8830	0,6710
3	4	100	105	0,9848	0,9659	0,9513	0,9698	0,9330
3	5	230	275	-0,7660	-0,9962	0,7631	0,5868	0,9924
4	4	130	170	0,7660	0,1736	0,1330	0,5868	0,0302
Suma:						5,4718	6,1033	6,1310

Las diferencias  $a_i - a_j$  para el primer par (1,2) se calcularía:

$$145^\circ - 190^\circ = -45^\circ$$

Para el caso  $b_i - b_j$  del primer par (1,2) se calcularía:

$$120^\circ - 180^\circ = -60^\circ$$

y así sucesivamente.

Podemos calcular  $r_{aa}$ :

$$r_{aa} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}(a_i - a_j) \text{sen}(b_i - b_j)}{\sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}^2(a_i - a_j) \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sen}^2(b_i - b_j)}} = \frac{5,4718}{\sqrt{6,1033 \cdot 6,1310}} = 0,8945$$

Como se ha señalado el procedimiento continua eliminando cada uno de los pares y calculando de nuevo  $r_{aa}$ . Repetimos los cálculos eliminando  $i = 1$  y tenemos:

$i$	$j$	$a_i - a_j$	$b_i - b_j$	$\text{sen}(a_i - a_j)$	$\text{sen}(b_i - b_j)$	$\text{sen}(a_i - a_j) \cdot \text{sen}(b_i - b_j)$	$\text{sen}^2(a_i - a_j)$	$\text{sen}^2(b_i - b_j)$
2	3	-120	-150	-0,8660	-0,5000	0,4330	0,7500	0,2500
2	4	-20	-45	-0,3420	-0,7071	0,2418	0,1170	0,5000
2	5	110	125	0,9397	0,8192	0,7698	0,8830	0,6710
3	4	100	105	0,9848	0,9659	0,9513	0,9698	0,9330
3	5	230	275	-0,7660	-0,9962	0,7631	0,5868	0,9924
4	4	130	170	0,7660	0,1736	0,1330	0,5868	0,0302
Suma:						3,2920	3,8935	3,3766

Entonces:

$$r_{aa} = \frac{3,2920}{\sqrt{3,8935 \cdot 3,3766}} = 0,9079$$

Si eliminamos  $i = 2$ , tenemos que  $r_{aa} = 0,87419$ . De modo que:

$i$ eliminado	1	2	3	4	5
$r_{aa}$	0,908	0,874	0,929	0,891	0,874

Ahora calculamos la media y la varianza para determinar los intervalos de confianza:

$$\bar{r}_{aa} = 0,8951; \quad s_{aa}^2 = 0,0005552$$

$$n \cdot r_{aa} - (n - 1) \cdot \bar{r}_{aa} = 5 \cdot 0,8941 - (5 - 1) \cdot 0,8951 = 0,8917$$

$$z_{0,05(2)} \sqrt{\frac{s_{r_{aa}}^2}{n}} = 1,96 \sqrt{\frac{0,0005552}{5}} = 0,0206$$

Por tanto:

$$L_1 = 0,8917 - 0,0206 = 0,8711 \text{ y } L_2 = 0,8917 + 0,0206 = 0,9123.$$

Como el intervalo de confianza no incluye al cero rechazamos  $H_0$ : hay correlación.

### 5.3. CORRELACIÓN CIRCULAR-LINEAL

En muchos casos es necesario poder correlacionar variables lineales y circulares, por ejemplo entre la dirección y la distancia a la que se desplazan algunos animales. El coeficiente de correlación circular-lineal es el siguiente:

$$r_{al} = \sqrt{\frac{r_{XC}^2 + r_{XS}^2 - 2 \cdot r_{XC} \cdot r_{XS} \cdot r_{CS}}{1 - r_{CS}^2}}$$

El coeficiente de correlación oscila entre 0 y 1, por tanto no existen correlaciones negativas. Si  $n$  es grande el nivel de significación puede obtenerse comparando  $n \cdot r_{al}^2$  con una  $\chi_{\alpha,2}^2$ .

**Ejercicio 31** (tomado de Zar, 1996). Deseamos saber si existe correlación entre la dirección tomada por 7 aves y la distancia viajada. En este caso:

$$H_0 = \rho_{al} = 0 \text{ y } H_A = \rho_{al} \neq 0, \text{ con } \alpha = 0,05.$$

Distancia (km) $x$	Dirección ( $\alpha$ )	seno	coseno	sen <sup>2</sup>	cos <sup>2</sup>	$x^2$	$x \cdot \cos$	$x \cdot \text{sen}$	sen $\cdot$ cos
48	190°	-0,1736	-0,9848	0,0302	0,9699	2304	-47,2709	-8,3347	0,1710
55	160°	0,3420	-0,9397	0,1170	0,8830	3025	-51,6830	18,8111	-0,3214
26	210°	-0,5000	-0,8660	0,2500	0,7500	676	-22,5168	-13,0000	0,4330
23	225°	-0,7071	-0,7071	0,5000	0,5000	529	-16,2635	-16,2635	0,5000
22	220°	-0,6428	-0,7660	0,4132	0,5868	484	-16,8529	-14,1414	0,4924
62	140°	0,6428	-0,7660	0,4132	0,5868	3844	-47,4945	39,8530	-0,4924
64	120°	0,8660	-0,5000	0,7500	0,2500	4096	-32,0000	55,4259	-0,4330
Suma = 300		-0,1727	-5,52972	2,4735	4,5265	14958	-234,0815	62,3504	0,3496

Por tanto disponemos de los siguientes resultados:

$$\sum x_i = 300 \text{ km}; \quad \sum x_i^2 = 14952 \text{ km}^2; \quad \sum \text{sen}\alpha = -0,1727^\circ;$$

$$\sum \text{sen}^2\alpha = 2,4735^\circ; \quad \sum \text{cos}\alpha = -5,52972^\circ; \quad \sum \text{cos}^2\alpha = 4,5265^\circ;$$

Para la suma de los cuadrados aplicamos la siguiente ecuación:

$$\sum x^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

De modo que:

$$\text{Suma de los cuadrados de } x = 14952 - \frac{300^2}{7} = 2100,86$$

$$\text{Suma de los cuadrados de seno } \alpha = 2,4735 - \frac{(-0,1727)^2}{7} = 2,4692$$

$$\text{Suma de los cuadrados de coseno } \alpha = 4,5265 - \frac{(-5,5297)^2}{7} = 0,1582$$

$$\text{Suma de los productos de } x \text{ por } \cos \alpha = 48 \cdot (-0,98481) + 55 \cdot (-0,96969) + \dots + 64 \cdot (-0,5) - 300 \cdot (-5,52972)/7 = 2,908$$

$$\text{Suma de los productos de } x \text{ por } \sin \alpha = 48 \cdot (-0,17364) + 55 \cdot (0,34202) + \dots + 64 \cdot (0,86603) - 300 \cdot (-0,1727)/7 = 67,90$$

$$\text{Suma de los productos } \cos \alpha \text{ por } \sin \alpha = (-0,98481) \cdot (-0,17364) + (-0,96969) \cdot (0,34202) + \dots + (-0,5) \cdot (0,86603) - (-5,52972) \cdot (-0,1727) / 7 = 0,2139$$

Ya podemos obtener los coeficientes de correlación, aplicando la siguiente fórmula:

$$r = \frac{\sum x \cdot y}{\sqrt{\sum x^2 \cdot \sum y^2}}$$

De modo que:

$$r_{x,\cos\alpha} = \frac{2,908}{\sqrt{2100,86 \cdot 0,1582}} = 0,1594, \quad \text{por tanto: } r_{x,\cos\alpha}^2 = 0,02543$$

$$r_{x,\sin\alpha} = \frac{67,70}{\sqrt{2100,86 \cdot 2,4692}} = 0,9677, \quad \text{por tanto: } r_{x,\sin\alpha}^2 = 0,9366$$

$$r_{\cos\alpha,\sin\alpha} = \frac{0,2139}{\sqrt{2,4692 \cdot 0,1582}} = 0,3421, \quad \text{por tanto: } r_{\cos\alpha,\sin\alpha}^2 = 0,1170$$

De aquí obtenemos aplicando la siguiente ecuación, que:

$$r_{al} = \sqrt{\frac{r_{x,\cos\alpha}^2 + r_{x,\sin\alpha}^2 - 2 \cdot r_{x,\cos\alpha} \cdot r_{x,\sin\alpha} \cdot r_{\cos\alpha,\sin\alpha}}{1 - r_{\cos\alpha,\sin\alpha}^2}}$$

$$r_{al} = \sqrt{\frac{0,2543 + 0,9366 - 2 \cdot 0,1594 \cdot 0,9677 \cdot 0,3421}{1 - 0,1170}} = 0,9854$$

Ahora calculamos:

$$n \cdot r_{ad}^n = 7 \cdot 0,9854 = 6,8978$$

Como  $\chi_{0,05,2}^2 = 5,991$  no rechazamos  $H_0$ . No hay correlación entre la dirección tomada y la distancia en las aves estudiadas.

## CAPÍTULO 6

# BIBLIOGRAFÍA

- Batschelet, E. (1981). *Circular Statistics in Biology*. London: Academic Press.
- Borycka, K. y Kasprzyk, I. (2018). Hourly pattern of allergenic alder and birch pollen concentrations in the air: Spatial differentiation and the effect of meteorological conditions. *Atmospheric Environment*, 182: 179-192.
- Cabrera, J.; Schmidt-Koenig, K. y Watson, G.S. (1991). The statistical analysis of circular data. En: P.P.G. Bateson y P.H. Kopler (eds): *Human Understanding and Animal Awareness*, pp: 285-306. New York: Plenum Press.
- Firdu, D. y Başpınar, E. (2017). Application of circular ANOVA on biological data: Case study on Crimea-Congo hemorrhage fever cases in Turkey. *International Annals of Medicine*. 2017; 1(3). <<https://doi.org/10.24087/IAM.2017.1.3.67>>.
- Fisher, N.I. (1953). *Dispersion on a sphere*. Proceedings of the Royal Society, London, 125: 54-59.
- Fisher, N.I. (1993). *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press.
- Fisher, N.I.; Lewis, T. y Embleton, B.J.J. (1987). *Statistical Analysis of Spherical Data*. Cambridge: Cambridge University Press.
- Flores Hernández, D.; Ramos Miranda, J. y Sosa López, A. 2007. *Estadística Descriptiva, Probabilidad y Pruebas de Hipótesis I*. Universalidad Autónoma de Campeche. Facultad de Ciencias Químico Biológicas.
- Gill, J. y Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, 18: 316-336.
- Gorgas García, J.; Cardiel López, N. y Zamorano Calvo, J. 2011. *Estadística básica para estudiantes de Ciencias*. Madrid: Universidad Complutense.
- Hammer, O.; Harper, D.A.T. y Ryan, P.O.D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Paleontologica Electronica*, 4 (1): 9 pp.
- Jammalamadaka, S.R. y SenGupta, A. (2001): *Topics in circular statistics*. Singapore: World Scientific Publishing.
- Karoly, Ph.J.; Goldenholz, D.M.; Freestone, D.R.; Mass, R.E.; Grayden, D.B.; Theodore, W.T. y Cook, M.J. (2018). Circadian and circaseptan rhythms in human epilepsy: a retrospective cohort study. *Lancet Neurology*, 17: 977-988.
- Kovach, W.L. (2011). *Oriana-Circular Statistics for Windows*, ver. 4. Pentraeth, Wales: Kovach Computing Services.
- Ley, Ch. y Verdebout, Th. (2017). *Modern Directional Statistics*. London: CRC Press.
- Mardia, K.V. (1972). *Statistics of directional data*. London: Academic Press.
- Mardia, K.V. y Jupp, P.E. (2000). *Directional statistics*. Chichester: Wiley.
- Oliveira, M.; Crujeiras, R.M. y Rodríguez-Casal, A. (2012). A plug-in rule for bandwidth selection in circular density estimation. *Computational Statistics and Data Analysis*, 56: 3898-3908.

- Patricia, L.; Morellato, C.; Alberti, L.F. y Hudson, I.L., (2010). Applications of circular statistics in plant phenology: a case studies approach. En: I.L. Hudson y M.R. Keatley (eds.), *Phenological Research*: 339-359. Springer Science + Business Media.
- Pewsey, A.; Neuhausser, M. y Ruxton, G.D. (2013). *Circular Statistics in R*. New York: Oxford University Press.
- Schaaf, AA.; García, C.G.; Puechagut, P.B.; Silvetti, L.E. y Quaglia, A.I.E., (2018). Effect of geographical latitude and sun exposure on Rufous Hornero (*Furnarius rufus*) nest orientation. *Journal of Ornithology*, 159: 967-974.
- Sokal, R.T. y Rohlf, F.J., (1979). *Biometría. Principios y métodos estadísticos en la investigación biológica*. Madrid: Blume.
- Sopini, F.; Rossano, C.; Marchetti, G.M. y Morgan, E. (2005). The role of the biological clock in the sun compass orientation of free-running individuals of *Talitrus saltator*. *Animal Behaviour*, 69: 835-843.
- Tasdan, F. y Yenisay, O. (2014). Power study of circular ANOVA test against nonparametric alternatives. *Hacettepe Journal of Mathematics and Statistics*, 43: 97-115.
- Upton, G.J.G. y Fingleton, B. (1989). *Spatial data analysis by example, Vol. 2: categorical and directional data*. New York: John Wiley.
- Wallraff, H.G. (1979). Goal-orientated and compass-oriented movements of displayed homing pigeons after confinement in differentially shielded aviaries. *Behavioral Ecology and Sociobiology*, 5: 201-225.
- Zar, J.H. (1984). *Biostatistical analysis*. New Jersey: Prentice Hall.
- Zar, J.H. (1996). *Biostatistical analysis*. (3<sup>rd</sup> edition). New Jersey: Prentice Hall.

ANEXO I  
**TABLAS**

Tabla A. Valores críticos para la  $z$  de Rayleigh (tomado de Zar, 1996)

$n$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
6	0,734	1,639	2,274	2,865	3,576	4,058	4,491	4,985	5,297
7	0,727	1,634	2,278	2,885	3,627	4,143	4,617	5,181	5,556
8	0,723	1,631	2,281	2,899	3,665	4,205	4,710	5,322	5,743
9	0,719	1,628	2,283	2,910	3,694	4,252	4,780	5,430	5,885
10	0,717	1,626	2,285	2,919	3,716	4,289	4,835	5,514	5,996
11	0,715	1,625	2,287	2,926	3,735	4,319	4,879	5,582	6,085
12	0,713	1,623	2,288	2,932	3,750	4,344	4,916	5,638	6,158
13	0,711	1,622	2,289	2,937	3,763	4,365	4,947	5,685	6,219
14	0,710	1,621	2,290	2,941	3,774	4,383	4,973	5,725	6,271
15	0,709	1,620	2,291	2,945	3,784	4,398	4,996	5,759	6,316
16	0,708	1,620	2,292	2,948	3,792	4,412	5,015	5,789	6,354
17	0,707	1,619	2,292	2,951	3,799	4,423	5,033	5,815	6,388
18	0,706	1,619	2,293	2,954	3,806	4,434	5,048	5,838	6,418
19	0,705	1,618	2,293	2,956	3,811	4,443	5,061	5,858	6,445
20	0,705	1,618	2,294	2,958	3,816	4,451	5,074	5,877	6,466
21	0,704	1,617	2,294	2,960	3,821	4,459	5,085	5,893	6,491
22	0,704	1,617	2,295	2,961	3,825	4,466	5,095	5,908	6,510
23	0,703	1,616	2,295	2,963	3,829	4,472	5,104	5,922	6,528
24	0,703	1,616	2,295	2,964	3,833	4,478	5,112	5,935	6,544
25	0,702	1,616	2,296	2,966	3,836	4,483	5,120	5,946	6,559
26	0,702	1,616	2,296	2,967	3,839	4,448	5,127	5,957	6,573
27	0,702	1,615	2,296	2,968	3,842	4,492	5,133	5,966	6,586
28	0,701	1,615	2,296	2,969	3,844	4,496	5,139	5,975	6,598
29	0,701	1,615	2,297	2,970	3,847	4,500	5,145	5,948	6,609
30	0,701	1,615	2,297	2,971	3,849	4,504	5,150	5,992	6,619
32	0,700	1,614	2,297	2,972	3,853	4,510	5,159	6,006	6,637
34	0,700	1,614	2,297	2,974	3,856	4,516	5,168	6,018	6,654
36	0,700	1,614	2,298	2,975	3,859	4,521	5,175	6,030	6,668
38	0,699	1,614	2,298	2,976	3,862	4,525	5,182	6,039	6,681
40	0,699	1,613	2,298	2,977	3,865	4,529	5,188	6,048	6,692
42	0,699	1,613	2,298	2,978	3,867	4,533	5,193	6,056	6,703
44	0,698	1,613	2,299	2,979	3,869	4,536	5,198	6,064	6,712
46	0,698	1,613	2,299	2,979	3,871	4,539	5,202	6,070	6,721
48	0,698	1,613	2,299	2,980	3,873	4,542	5,206	6,076	6,729
50	0,698	1,613	2,299	2,981	3,874	4,545	5,210	6,082	6,736
55	0,697	1,612	2,299	2,982	3,878	4,550	5,218	6,094	6,752
60	0,697	1,612	2,300	2,983	3,881	4,555	5,225	6,104	6,765
65	0,697	1,612	2,300	2,984	3,883	4,559	5,231	6,113	6,776
70	0,696	1,612	2,300	2,985	3,885	4,562	5,235	6,120	6,786
75	0,696	1,612	2,300	2,986	3,887	4,565	5,240	6,127	6,794
80	0,696	1,611	2,300	2,986	3,889	4,567	5,243	6,132	6,801
90	0,696	1,611	2,301	2,987	3,891	4,572	5,249	6,141	6,813
100	0,695	1,611	2,301	2,988	3,893	4,575	5,254	6,149	6,822

**Tabla B.** Valores críticos para la  $m$  del test de Hodges-Ajne (tomado de Zar, 1996)

$n$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
4	0								
5	0								
6	0	0							
7	0	0							
8	1	0	0						
9	1	0	0	0					
10	1	1	0	0	0				
11	2	1	1	0	0				
12	2	1	1	0	0	0			
13	3	2	1	1	0	0	0		
14	3	2	1	1	0	0	0	0	
15	3	2	2	1	1	0	0	0	0
16	4	3	2	2	1	1	0	0	0
17	4	3	2	2	1	1	1	0	0
18	4	3	3	2	2	1	1	0	0
19	5	4	3	3	2	2	1	1	0
20	5	4	3	3	2	2	1	1	1
21	6	4	4	3	3	2	2	1	1
22	6	5	4	4	3	2	2	2	1
23	6	5	4	4	3	3	2	2	1
24	7	6	5	4	3	3	3	2	2
25	7	6	5	5	4	3	3	2	2
26	8	6	6	5	4	4	3	3	2
27	8	7	6	5	4	4	4	3	3
28	8	7	6	6	5	4	4	3	3
29	9	7	7	6	5	5	4	4	3
30	9	8	7	6	6	5	4	4	3
31	10	8	7	7	6	5	5	4	4
32	10	9	8	7	6	6	5	4	4
33	11	9	8	7	7	6	5	5	4
34	11	9	9	8	7	6	6	5	5
35	11	10	9	8	7	7	6	5	5
36	12	10	9	9	8	7	6	6	5
37	12	11	10	9	8	7	7	6	6
38	13	11	10	9	8	8	7	6	6
39	13	11	10	10	9	8	7	7	6
40	13	12	11	10	9	8	8	7	7
41	14	12	11	10	9	9	8	7	7
42	14	13	12	11	10	9	9	8	7
43	15	13	12	11	10	9	9	8	8
44	15	13	12	12	11	10	9	8	8
45	16	14	13	12	11	10	10	9	8
46	16	14	13	12	11	11	10	9	9
47	16	15	14	13	12	11	10	10	9
48	17	15	14	13	12	11	11	10	9
49	17	15	14	13	12	12	11	10	10
50	18	16	15	14	13	12	11	11	10

**Tabla C.** Valores críticos para la  $u$  del test de la  $V$  (tomado de Bastchelet, 1981)

$n$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,001$	$\alpha = 0,0001$
5	1,3051	1,6524	2,2505	2,4459	2,7938	3,0825
6	1,3009	1,6509	2,2640	2,4695	2,8502	3,2114
7	1,2980	1,6499	2,2734	2,4858	2,8886	3,2970
8	1,2958	1,6492	2,2803	2,4978	2,9164	3,3358
9	1,2942	1,6484	2,2856	2,5070	2,9375	3,4034
10	1,2929	1,6482	2,2899	2,5143	2,9540	3,4387
11	1,2918	1,6479	2,2933	2,5201	2,9672	3,4669
12	1,2909	1,6476	2,2961	2,5250	2,9782	3,4899
13	1,2902	1,6474	2,2985	2,5290	2,9873	3,5091
14	1,2895	1,6472	2,3006	2,5325	2,9950	3,5253
15	1,2890	1,6470	2,3023	2,5355	3,0017	3,5392
16	1,2885	1,6469	2,3039	2,5381	3,0075	3,5512
17	1,2881	1,6467	2,3052	2,5404	3,0126	3,5617
18	1,2877	1,6466	2,3064	2,5424	3,0171	3,5710
19	1,2874	1,6465	2,3075	2,5442	3,0211	3,5792
20	1,2871	1,6464	2,3085	2,5458	3,0247	3,5866
21	1,2868	1,6464	2,3093	2,5473	3,0279	3,5932
22	1,2866	1,6463	2,3101	2,5486	3,0308	3,5992
23	1,2864	1,6462	2,3108	2,5498	3,0335	3,6047
24	1,2862	1,6462	2,3115	2,5509	3,0359	3,6096
25	1,2860	1,6461	2,3121	2,5519	3,0382	3,6142
26	1,2858	1,6461	2,3127	2,5529	3,0402	3,6118
27	1,2856	1,6460	2,3132	2,5538	3,0421	3,6293
28	1,2855	1,6460	2,3136	2,5546	3,0439	3,6258
29	1,2853	1,6459	2,3141	2,5553	3,0455	6,6292
30	1,2852	1,6459	2,3145	2,5560	3,0471	3,6323
40	1,2843	1,6456	2,3175	2,5610	3,0580	3,6545
50	1,2837	1,6455	2,3193	2,5640	3,0646	3,6677
60	1,2834	1,6454	2,3205	2,5660	3,0689	3,6764
70	1,2831	1,6453	2,3213	2,5674	3,0720	3,6826
100	1,2826	1,6452	2,3228	2,5699	3,0775	3,6936
500	1,2818	1,6449	2,3256	2,5747	3,0877	3,7140
1000	1,2817	1,6449	2,3260	2,5772	3,0890	3,7165

**Tabla D.** Valores críticos para la U del test de Rao (tomado de Batschelet, 1981)

$n$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,1$
4	221,0	186,5	171,7
5	212,0	183,6	168,8
6	206,0	180,7	166,3
7	202,7	177,8	164,9
8	198,4	175,7	163,4
9	195,1	173,5	162,4
10	192,2	172,1	161,3
11	189,7	170,3	160,2
12	187,6	169,2	159,2
13	185,5	167,8	158,4
14	184,0	166,7	157,7
15	182,2	165,6	157,0
16	180,7	164,9	156,6
17	179,6	164,2	155,9
18	178,2	163,1	155,2
19	177,1	162,4	154,8
20	176,0	161,6	154,4
25	171,9	158,9	152,7
30	168,8	156,7	151,4
35	166,4	155,0	150,3
40	164,4	153,6	149,5
45	162,7	152,4	148,7
50	161,2	151,4	148,1
100	152,8	146,8	143,7
200	146,8	142,6	140,4

**Tabla E.** Valores críticos para la K del test de Kuiper (tomado de Batschelet, 1981)

$n$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
5	1,102	1,330	1,458	1,565	1,682	1,763	1,838	1,920	1,970
6	1,109	1,341	1,471	1,582	1,711	1,793	1,867	1,957	2,020
7	1,116	1,350	1,483	1,598	1,727	1,814	1,894	1,987	2,051
8	1,122	1,357	1,493	1,608	1,741	1,830	1,911	2,009	2,077
9	1,127	1,364	1,500	1,618	1,752	1,843	1,926	2,027	2,097
10	1,131	1,370	1,507	1,625	1,761	1,854	1,938	2,041	2,113
11	1,134	1,374	1,513	1,631	1,769	1,862	1,948	2,053	2,126
12	1,138	1,378	1,517	1,637	1,776	1,876	1,957	2,062	2,137
13	1,141	1,382	1,522	1,642	1,782	1,870	1,964	2,071	2,146
14	1,143	1,385	1,525	1,646	1,787	1,882	1,970	2,078	2,154
15	1,146	1,388	1,529	1,650	1,791	1,887	1,976	2,085	2,161
16	1,148	1,391	1,532	1,653	1,795	1,892	1,981	2,090	2,168
17	1,150	1,393	1,534	1,657	1,799	1,896	1,986	2,096	2,173
18	1,152	1,396	1,537	1,659	1,802	1,899	1,990	2,100	2,178
19	1,154	1,398	1,539	1,662	1,805	1,903	1,993	2,104	2,183
20	1,155	1,400	1,541	1,664	1,808	1,906	1,997	2,108	2,187
21	1,157	1,401	1,543	1,667	1,810	1,908	2,000	2,112	2,191
22	1,158	1,403	1,545	1,669	1,813	1,911	2,003	2,115	2,194
23	1,159	1,405	1,547	1,670	1,815	1,913	2,005	2,118	2,198
24	1,161	1,406	1,549	1,672	1,817	1,916	2,008	2,121	2,201
25	1,162	1,407	1,550	1,674	1,819	1,918	2,010	2,123	2,203
30	1,167	1,413	1,556	1,681	1,826	1,926	2,019	2,134	2,215
35	1,171	1,417	1,561	1,686	1,832	1,933	2,026	2,141	2,223
40	1,174	1,421	1,565	1,690	1,837	1,938	2,032	2,148	2,23
45	1,177	1,424	1,568	1,694	1,841	1,942	2,036	2,152	2,235
50	1,179	1,427	1,571	1,697	1,844	1,946	2,040	2,157	2,239
100	1,191	1,442	1,588	1,714	1,862	1,965	2,060	2,178	2,262
200	1,200	1,453	1,600	1,726	1,876	1,979	2,075	2,194	2,279
500	1,208	1,462	1,610	1,737	1,887	1,990	2,087	2,207	2,292
$\infty$	1,216	1,472	1,620	1,747	1,898	2,001	2,098	2,218	2,303

Tabla F. Valores críticos para la chi-cuadrado

$u$	$\alpha = 0,20$	$0,10$	$0,05$	$0,025$	$0,02$	$0,01$	$0,005$	$0,001$
1	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,828
2	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,816
3	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,266
4	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,467
5	7,289	9,236	11,070	12,833	13,388	15,086	16,750	20,515
6	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,458
7	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,124
9	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	29,553	33,196	36,415	39,364	40,270	42,980	45,559	51,179
25	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	32,912	36,741	40,113	43,195	44,140	46,963	49,645	55,476
28	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,892
29	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,301
30	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,70

Tabla G. Valores críticos para la U2 del test de Watson (tomado de Zar, 1996)

$n_1$	$n_2$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
4	4	0,1172	0,1875	****	****	****	****	****	****	****
4	5	0,0815	0,2037	0,2037	****	****	****	****	****	****
4	6	0,0875	0,1333	0,2167	0,2167	****	****	****	****	****
4	7	0,0844	0,1299	1,1688	0,2273	****	****	****	****	****
4	8	0,0903	0,1319	0,1362	0,2631	****	****	****	****	****
4	9	0,0855	0,1282	0,1752	0,2436	0,2436	****	****	****	****
4	10	0,0804	0,1232	0,1571	0,2018	0,2500	****	****	****	****
4	11	0,0828	0,1253	0,1556	0,1949	0,2556	****	****	****	****
4	12	0,0781	0,1302	0,1563	0,2031	0,2604	0,2604	****	****	****
4	13	0,0792	0,1244	0,1538	0,1855	0,2647	0,2647	****	****	****
4	14	0,0780	0,1227	0,1534	0,1931	0,2298	0,2685	****	****	****
4	15	0,0789	0,1228	0,1561	0,1807	0,2298	0,2719	0,2719	****	****
4	16	0,0781	0,1250	0,1531	0,1836	0,2281	0,2750	0,2750	****	****
4	17	0,0775	0,1223	0,1531	0,1839	0,2330	0,2778	0,2778	****	****
4	18	0,0764	0,1212	0,1490	0,1818	0,2197	0,2481	0,2803	****	****
4	19	0,0755	0,1213	0,1533	0,1796	0,2220	0,2517	0,2826	****	****
4	20	0,0764	0,1201	0,1535	0,1842	0,2264	0,2451	0,2847	****	****
4	21	0,0752	0,1200	0,1514	0,1819	0,2143	0,2486	0,2867	0,2867	****
4	22	0,0756	0,1211	0,1508	0,1823	0,2185	0,2517	0,2885	0,2885	****
4	23	0,0751	0,1194	0,1508	0,1814	0,2177	0,2394	0,2636	0,2901	****
4	24	0,0755	0,1202	0,1499	0,1797	0,2184	0,2411	0,2660	0,2917	****
4	25	0,0752	0,1200	0,1497	0,1814	0,2152	0,2441	0,2600	0,2931	****
4	26	0,0752	0,1191	0,1486	0,1816	0,2175	0,2396	0,2624	0,2494	****
4	27	0,0753	0,1139	0,1505	0,1786	0,2151	0,2360	0,2646	0,2957	0,2957
4	28	0,0748	0,1203	0,1496	0,1775	0,2165	0,2388	0,2667	0,2969	0,2969
4	29	0,0749	0,1198	0,1491	0,1794	0,2165	0,2369	0,2557	0,2980	0,2980
4	30	0,0745	0,1196	0,1493	0,1797	0,2140	0,2395	0,2578	0,2990	0,2990
5	5	0,0890	0,1610	0,2250	0,2250	****	****	****	****	****
5	6	0,0848	0,1333	0,1818	0,2424	****	****	****	****	****
5	7	0,0846	0,1284	0,1712	0,1998	0,2569	****	****	****	****
5	8	0,0855	0,1308	0,1654	0,2154	0,2692	****	****	****	****
5	9	0,0798	0,1242	0,1591	0,1909	0,2798	0,2798	****	****	****
5	10	0,0836	0,1236	0,1609	0,1956	0,2409	0,2889	0,2889	****	****
5	11	0,0810	0,1241	0,1560	0,1901	0,2287	0,2969	0,2969	****	****
5	12	0,0784	0,1235	0,1549	0,1863	0,2255	0,2608	0,3039	****	****
5	13	0,0777	0,1256	0,1563	0,1837	0,2298	0,2692	0,3102	****	****
5	14	0,0782	0,1218	0,1534	0,1820	0,2211	0,2571	0,2767	0,3158	****
5	15	0,0782	0,1235	0,1515	0,1835	0,2248	0,2515	0,2835	0,3208	****
5	16	0,7660	0,1206	0,1552	0,1825	0,2230	0,2552	0,2897	0,3254	****
5	17	0,0761	0,1199	0,1520	0,1820	0,2205	0,2472	0,2782	0,3295	0,3295
5	18	0,0763	0,1208	0,1536	0,1797	0,2164	0,2464	0,2715	0,3338	0,3338
5	19	0,0754	0,1201	0,1517	0,1824	0,2193	0,2526	0,2745	0,3052	0,3368
5	20	0,0760	0,1216	0,1520	0,1824	0,2200	0,2416	0,2664	0,3096	0,3400

**Tabla G.** Valores críticos para la U2 del test de Watson (tomado de Zar, 1996) (cont.)

$n_1$	$n_2$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
5	21	0,0755	0,1195	0,1510	0,1810	0,2206	0,2448	0,2712	0,2990	0,3429
5	22	0,0756	0,1201	0,1524	0,1820	0,2191	0,2426	0,2689	0,3033	0,3457
5	23	0,0755	0,1196	0,1513	0,1811	0,2178	0,2451	0,2737	0,2963	0,3209
5	24	0,0747	0,1195	0,1511	0,1810	0,2190	0,2437	0,2736	0,2993	0,3241
5	25	0,0754	0,1197	0,1517	0,1810	0,2166	0,2461	0,2674	0,3021	0,3272
5	26	0,0749	0,1186	0,1514	0,1806	0,2189	0,2447	0,2675	0,2943	0,3176
5	27	0,0748	0,1193	0,1508	0,1804	0,2165	0,2443	0,2674	0,2975	0,3207
5	28	0,0746	0,1188	0,1512	0,1802	0,2170	0,2417	0,2694	0,2937	0,3136
5	29	0,0743	0,1189	0,1510	0,1802	0,2171	0,2443	0,2666	0,2790	0,3153
5	30	0,0743	0,1189	0,1512	0,1802	0,2160	0,2419	0,2678	0,2979	0,3181
6	6	0,0880	0,1319	0,1713	0,2060	0,2639	****	****	****	****
6	7	0,0806	0,1209	0,1538	0,1941	0,2821	0,2821	****	****	****
6	8	0,0833	0,1265	0,1607	0,1964	0,2455	0,2976	0,2976	****	****
6	9	0,0815	0,1259	0,1556	0,1926	0,2321	0,2617	0,3111	****	****
6	10	0,0771	0,1260	0,1563	0,1896	0,2313	0,2479	0,3229	0,3229	****
6	11	0,0784	0,1212	0,1569	0,1872	0,2246	0,2620	0,2888	0,3333	****
6	12	0,0802	0,1242	0,1551	0,1829	0,2261	0,2593	0,2747	0,3426	0,3426
6	13	0,0769	0,1215	0,1538	0,1849	0,2213	0,2497	0,2780	0,3509	0,3509
6	14	0,0768	0,1220	0,1536	0,1839	0,2250	0,2506	0,2821	0,3196	0,3583
6	15	0,0762	0,1217	0,1524	0,1852	0,2201	0,2487	0,2730	0,3058	0,3651
6	16	0,0758	0,1212	0,1534	0,1823	0,2235	0,2500	0,2789	0,3073	0,3357
6	17	0,0750	0,1211	0,1526	0,1833	0,2199	0,2472	0,2745	0,3129	0,3427
6	18	0,0760	0,1211	0,1535	0,1840	0,2199	0,2461	0,2739	0,2998	0,3298
6	19	0,0751	0,1200	0,1523	0,1832	0,2204	0,2498	0,2744	0,3060	0,3298
6	20	0,0747	0,1196	0,1526	0,1824	0,2196	0,2490	0,2734	0,3077	0,3333
6	21	0,0758	0,1205	0,1523	0,1834	0,2205	0,2475	0,2734	0,3057	0,3369
6	22	0,0749	0,1204	0,1518	0,1824	0,2202	0,2473	0,2752	0,3036	0,3260
6	23	0,0745	0,1194	0,1514	0,1824	0,2194	0,2469	0,2729	0,3073	0,3273
6	24	0,0743	0,1194	0,1519	0,1826	0,2206	0,2484	0,2715	0,3056	0,3289
6	25	0,0744	0,1191	0,1514	0,1819	0,2202	0,2473	0,2731	0,3015	0,3277
6	26	0,0739	0,1188	0,1510	0,1815	0,2198	0,2464	0,2710	0,3047	0,3265
6	27	0,0741	0,1193	0,1515	0,1822	0,2200	0,2469	0,2731	0,3053	0,3281
6	28	0,0737	0,1190	0,1507	0,1821	0,2201	0,2467	0,2731	0,3039	0,3270
6	29	0,0736	0,1189	0,1511	0,1816	0,2200	0,2473	0,2719	0,3038	0,3258
6	30	0,0736	0,1193	0,1509	0,1823	0,2194	0,2471	0,2725	0,3045	0,3262
7	7	0,0791	0,1345	0,1578	0,1986	0,2511	0,3036	0,3036	****	****
7	8	0,0794	0,1198	0,1556	0,1817	0,2246	0,2722	0,3222	****	****
7	9	0,0786	0,1223	0,1560	0,1818	0,2215	0,2552	0,2909	0,3385	****
7	10	0,0773	0,1227	0,1546	0,1866	0,2269	0,2622	0,2773	0,3529	0,3529
7	11	0,0771	0,1219	0,1551	0,1839	0,2214	0,2532	0,2806	0,3225	0,3657
7	12	0,0764	0,1216	0,1541	0,1855	0,2256	0,2519	0,2757	0,3083	0,3772
7	13	0,0765	0,1216	0,1545	0,1842	0,2227	0,2523	0,2776	0,3150	0,3479
7	14	0,0761	0,1228	0,1568	0,1840	0,2248	0,2530	0,2744	0,3210	0,3337

**Tabla G.** Valores críticos para la U2 del test de Watson (tomado de Zar, 1996) (cont.)

$n_1$	$n_2$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
7	15	0,0754	0,1213	0,1525	0,1845	0,2235	0,2503	0,2780	0,3118	0,3378
7	16	0,0753	0,1203	0,1530	0,1848	0,2236	0,2508	0,2772	0,3113	0,3432
7	17	0,0749	0,1204	0,1526	0,1827	0,2227	0,2500	0,2752	0,3109	0,3340
7	18	0,0749	0,1200	0,1524	0,1841	0,2235	0,2502	0,2768	0,3117	0,3346
7	20	0,0743	0,1198	0,1526	0,1832	0,2219	0,2499	0,2780	0,3081	0,3330
7	21	0,0751	0,1203	0,1534	0,1840	0,2240	0,2496	0,2782	0,3123	0,3336
7	22	0,0743	0,1196	0,1518	0,1832	0,2221	0,2512	0,2763	0,3090	0,3341
7	23	0,0739	0,1194	0,1522	0,1832	0,2226	0,2499	0,2780	0,3103	0,3327
8	8	0,0781	0,1250	0,1563	0,1836	0,2256	0,2500	0,2959	0,3438	****
8	9	0,0784	0,1225	0,1552	0,1863	0,2255	0,2582	0,2827	0,3627	0,3627
8	10	0,0775	0,1220	0,1546	0,1852	0,2220	0,2491	0,2796	0,3359	0,3796
8	11	0,0766	0,1220	0,1543	0,1842	0,2249	0,2524	0,2799	0,3194	0,3529
8	12	0,0766	0,1208	0,1557	0,1854	0,2229	0,2521	0,2807	0,3167	0,3396
8	13	0,0754	0,1212	0,1532	0,1853	0,2237	0,2531	0,2778	0,3135	0,3446
8	14	0,0751	0,1205	0,1530	0,1655	0,2224	0,2516	0,2796	0,3137	0,3381
8	15	0,0746	0,1210	0,1536	0,1855	0,2232	0,2507	0,2783	0,3130	0,3341
8	16	0,0761	0,1220	0,1542	0,1854	0,2222	0,2531	0,2795	0,3156	0,3414
8	17	0,0747	0,1200	0,1529	0,1841	0,2241	0,2524	0,2782	0,3124	0,3388
8	18	0,0748	0,1190	0,1528	0,1840	0,2244	0,2513	0,2813	0,3152	0,3397
8	19	0,0742	0,1196	0,1527	0,1839	0,2243	0,2526	0,2799	0,3143	0,3384
8	20	0,0741	0,1196	0,1527	0,1839	0,2239	0,2527	0,2795	0,3134	0,3393
9	9	0,0770	0,1250	0,1152	0,1867	0,2251	0,2663	0,2855	0,3404	0,3843
9	10	0,0760	0,1216	0,1544	0,1860	0,2257	0,2538	0,2865	0,3205	0,3614
9	11	0,0764	0,1208	0,1542	0,1845	0,2249	0,2552	0,2814	0,3168	0,3410
9	12	0,0767	0,1217	0,1543	0,1852	0,2257	0,2540	0,2804	0,3157	0,3395
9	13	0,0755	0,1205	0,1532	0,1850	0,2247	0,2526	0,2798	0,3187	0,3389
9	14	0,0752	0,1201	0,1532	0,1853	0,2243	0,2526	0,2809	0,3168	0,3409
9	15	0,0757	0,1201	0,1535	0,1850	0,2245	0,2541	0,2831	0,3152	0,3393
9	16	0,0744	0,1200	0,1533	0,1850	0,2244	0,2539	0,2822	0,3172	0,3439
10	10	0,0750	0,1225	0,1545	0,1850	0,2250	0,2545	0,2825	0,3170	0,3450
10	11	0,0756	0,1215	0,1544	0,1856	0,2237	0,2548	0,2791	0,3172	0,3405
10	12	0,0758	0,1212	0,1534	0,1848	0,2246	0,2545	0,2818	0,3155	0,3409
10	13	0,0749	0,1204	0,1532	0,1853	0,2254	0,2542	0,2816	0,3184	0,3452
10	14	0,0749	0,1201	0,1535	0,1847	0,2252	0,2550	0,2823	0,3181	0,3439
10	15	0,0747	0,1211	0,1536	0,1856	0,2256	0,2549	0,2837	0,3189	0,3440
11	11	0,0760	0,1211	0,1541	0,1857	0,2232	0,2540	0,2826	0,3194	0,3442
11	12	0,0751	0,1206	0,1535	0,1851	0,2253	0,2543	0,2839	0,3182	0,3439
11	13	0,0746	0,1206	0,1532	0,1853	0,2255	0,2546	0,2838	0,3193	0,3461
12	12	0,0752	0,1215	0,1528	0,1863	0,2266	0,2558	0,2844	0,3192	0,3438
14	14	0,0700	0,1170	0,1510	0,1830	0,2260	0,2580	0,2890	0,3300	0,3610
16	16	0,0700	0,1170	0,1510	0,1840	0,2270	0,2590	0,2910	0,3320	0,3640
18	18	0,0700	0,1170	0,1510	0,1840	0,2280	0,2600	0,2920	0,3340	0,3660
20	20	0,0590	0,1170	0,1510	0,1850	0,2280	0,2610	0,2930	0,3350	0,3670

**Tabla G.** Valores críticos para la U2 del test de Watson (tomado de Zar, 1996) (cont.)

$n_1$	$n_2$	$\alpha = 0,5$	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,002$	$\alpha = 0,001$
25	25	0,0690	0,1170	0,1520	0,1850	0,2290	0,2620	0,2950	0,3380	0,3700
30	30	0,0690	0,1170	0,1520	0,1860	0,2300	0,2630	0,2960	0,3390	0,3720
35	35	0,0690	0,1170	0,1520	0,1860	0,2310	0,2640	0,2970	0,3400	0,3730
40	40	0,0690	0,1170	0,1520	0,1860	0,2310	0,2640	0,2980	0,3410	0,3740
50	50	0,0690	0,1170	0,1520	0,1870	0,2310	0,2650	0,2990	0,3430	0,3760
60	60	0,0690	0,1170	0,1520	0,1870	0,2320	0,2660	0,2990	0,3430	0,3770
80	80	0,0690	0,1170	0,1520	0,1870	0,2320	0,2660	0,3000	0,3440	0,3780
100	100	0,0690	0,1170	0,1520	0,1870	0,2330	0,2670	0,3000	0,3450	0,3780
$\infty$	$\infty$	0,0710	0,1167	0,1518	0,1869	0,2333	0,2684	0,3035	0,3500	0,3851

Tabla H. Valores críticos para la F de Snedecor ( $\alpha = 0,10$ )

$\frac{u}{v}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	60,71	61,22	61,74	62,00	62,26	62,53	62,79	63,06
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,48
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,20	5,18	5,18	5,17	5,16	5,15	5,14
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,90	3,87	3,84	3,83	3,82	3,80	3,79	3,78
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,27	3,24	3,21	3,19	3,17	3,16	3,14	3,12
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,90	2,87	2,84	2,82	2,80	2,78	2,76	2,74
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,67	2,63	2,59	2,58	2,56	2,54	2,51	2,49
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,50	2,46	2,42	2,40	2,38	2,36	2,34	2,32
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,38	2,34	2,30	2,28	2,25	2,23	2,21	2,18
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,28	2,24	2,20	2,18	2,16	2,13	2,11	2,08
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,21	2,17	2,12	2,10	2,08	2,05	2,03	2,00
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,15	2,10	2,06	2,04	2,01	1,99	1,96	1,93
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14	2,10	2,05	2,01	1,98	1,96	1,93	1,90	1,88
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10	2,05	2,01	1,96	1,94	1,91	1,89	1,86	1,83
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	2,02	1,97	1,92	1,90	1,87	1,85	1,82	1,79
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03	1,99	1,94	1,89	1,87	1,84	1,81	1,78	1,75
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00	1,96	1,91	1,86	1,84	1,81	1,78	1,75	1,72
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1,93	1,89	1,84	1,81	1,78	1,75	1,72	1,69
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96	1,91	1,86	1,81	1,79	1,76	1,73	1,70	1,67
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,89	1,84	1,79	1,77	1,74	1,71	1,68	1,64
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92	1,87	1,83	1,78	1,75	1,72	1,69	1,66	1,62
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90	1,86	1,81	1,76	1,73	1,70	1,67	1,64	1,60
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89	1,84	1,80	1,74	1,72	1,69	1,66	1,62	1,59
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88	1,83	1,78	1,73	1,70	1,67	1,64	1,61	1,57
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87	1,82	1,77	1,72	1,69	1,66	1,63	1,59	1,56
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86	1,81	1,76	1,71	1,68	1,65	1,61	1,58	1,54
27	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85	1,80	1,75	1,70	1,67	1,64	1,60	1,57	1,53

**Tabla H.** Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,10$ ) (cont.)

$\begin{matrix} u \\ v \end{matrix}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84	1,79	1,74	1,69	1,66	1,63	1,59	1,56	1,52
29	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83	1,78	1,73	1,68	1,65	1,62	1,58	1,55	1,51
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,77	1,72	1,67	1,64	1,61	1,57	1,54	1,50
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,71	1,66	1,61	1,57	1,54	1,51	1,47	1,42
60	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74	1,71	1,66	1,60	1,54	1,51	1,48	1,44	1,40	4,35
120	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65	1,60	1,55	1,48	1,45	1,41	1,37	1,32	1,26
$\infty$	2,71	2,30	2,08	1,94	1,85	1,77	1,72	1,67	1,63	1,60	1,55	1,49	1,42	1,38	1,34	1,30	1,24	1,17

Tabla H. Valores críticos para la F de Snedecor ( $\alpha = 0,05$ ) (cont.)

$\nu$ \ $u$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40
6	6,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73

**Tabla H.** Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,05$ ) (cont.)

$\nu \backslash u$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35
$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22

Tabla H. Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,025$ ) (cont.)

$\backslash$ $v$	$u$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,7	984,9	984,9	993,1	997,2	1001	1006	1010	1014
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,43	39,45	39,46	39,47	39,48	39,49	39,50
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	14,25	14,17	14,12	14,08	14,04	13,99	13,95
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,66	8,56	8,51	8,46	8,41	8,36	8,31
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,43	6,33	6,28	6,23	6,18	6,12	6,07
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,27	5,17	5,12	5,07	5,01	4,96	4,90
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,57	4,47	4,42	4,36	4,31	4,25	4,20
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,10	4,00	3,95	3,89	3,84	3,78	3,73
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,77	3,67	3,61	3,56	3,51	3,45	3,39
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,52	3,42	3,37	3,31	3,26	3,20	3,14
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,43	3,33	3,33	3,23	3,17	3,12	3,06	3,00	2,94
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,18	3,07	3,02	2,96	2,91	2,85	2,79
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,15	3,05	3,05	2,95	2,89	2,84	2,78	2,72	2,66
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,05	2,95	2,95	2,84	2,79	2,73	2,67	2,61	2,55
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,86	2,76	2,70	2,64	2,59	2,52	2,46
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,89	2,79	2,79	2,68	2,63	2,57	2,51	2,45	2,38
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,82	2,72	2,72	2,62	2,56	2,50	2,44	2,38	2,32
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,77	2,67	2,67	2,56	2,50	2,44	2,38	2,32	2,26
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,72	2,62	2,62	2,51	2,45	2,39	2,33	2,27	2,20
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,57	2,46	2,41	2,35	2,29	2,22	2,16
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,64	2,53	2,53	2,42	2,37	2,31	2,25	2,18	2,11
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,60	2,50	2,50	2,39	2,33	2,27	2,21	2,14	2,08
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,57	2,47	2,47	2,36	2,30	2,24	2,18	2,11	2,04
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,54	2,44	2,44	2,33	2,27	2,21	2,15	2,08	2,01
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,51	2,41	2,41	2,30	2,24	2,18	2,12	2,05	1,98
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59	2,49	2,39	2,39	2,28	2,22	2,16	2,09	2,03	1,95
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57	2,47	2,36	2,36	2,25	2,19	2,13	2,07	2,00	1,93

**Tabla H.** Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,025$ ) (cont.)

$\begin{matrix} u \\ v \end{matrix}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55	2,45	2,34	2,23	2,17	2,11	2,05	1,98	1,91
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53	2,43	2,32	2,21	2,15	2,09	2,03	1,96	1,89
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43
$\infty$	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27

Tabla H. Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,01$ ) (cont.)

$\mu$ $\nu$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	4052	5000	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20

Tabla H. Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,01$ ) (cont.)

$\nu$ \ $u$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53
$\infty$	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32

Tabla H. Valores críticos para la  $F$  de Snedecor ( $\alpha = 0,005$ ) (cont.)

$\mu$ $\nu$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	24940	25044	25148	25253	25359
2	198,5	199,0	199,2	199,2	199,3	199,3	199,4	199,4	199,4	199,4	199,4	199,4	199,4	199,4	199,5	199,5	199,5	199,5
3	55,55	49,80	47,47	46,19	45,39	44,84	44,43	44,13	43,88	43,69	43,39	43,08	42,78	42,62	42,47	42,31	42,15	41,99
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,44	20,17	20,03	19,89	19,75	19,61	19,47
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,15	12,90	12,78	12,66	12,53	12,40	12,27
6	18,63	14,54	12,92	12,03	11,46	11,07	10,79	10,57	10,39	10,25	10,03	9,81	9,59	9,47	9,36	9,24	9,12	9,00
7	16,24	12,40	10,88	10,05	9,52	9,16	8,89	8,68	8,51	8,38	8,18	7,97	7,75	7,65	7,53	7,42	7,31	7,19
8	14,69	11,04	9,60	8,81	8,30	7,95	7,69	7,50	7,34	7,21	7,01	6,81	6,61	6,50	6,40	6,29	6,18	6,06
9	13,61	10,11	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	6,23	6,03	5,83	5,73	5,62	5,52	5,41	5,30
10	12,83	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,66	5,47	5,27	5,17	5,07	4,97	4,86	4,75
11	12,23	8,91	7,60	6,88	6,42	6,10	5,86	5,68	5,54	5,42	5,24	5,05	4,86	4,76	4,65	4,55	4,44	4,34
12	11,75	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,91	4,72	4,53	4,43	4,33	4,23	4,12	4,01
13	11,37	8,19	6,93	6,23	5,79	5,48	5,25	5,08	4,94	4,82	4,64	4,46	4,27	4,17	4,07	3,97	3,87	3,76
14	11,06	7,92	6,68	6,00	5,56	5,26	5,03	4,86	4,72	4,60	4,43	4,25	4,06	3,96	3,86	3,76	3,66	3,55
15	10,80	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	4,25	4,07	3,88	3,79	3,69	3,58	3,48	3,37
16	10,58	7,51	6,30	5,64	5,21	4,91	4,69	4,52	4,38	4,27	4,10	3,92	3,73	3,64	3,54	3,44	3,33	3,22
17	10,38	7,35	6,16	5,50	5,07	4,78	4,56	4,39	4,25	4,14	3,97	3,79	3,61	3,51	3,41	3,31	3,21	3,10
18	10,22	7,21	6,03	5,37	4,96	4,66	4,44	4,28	4,14	4,03	3,86	3,68	3,50	3,40	3,30	3,20	3,10	2,99
19	10,07	7,09	5,92	5,27	4,85	4,56	4,34	4,18	4,04	3,93	3,76	3,59	3,40	3,31	3,21	3,11	3,00	2,89
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,68	3,50	3,32	3,22	3,12	3,02	2,92	2,81
21	9,83	6,89	5,73	5,09	4,68	4,39	4,18	4,01	3,88	3,77	3,60	3,43	3,24	3,15	3,05	2,95	2,84	2,73
22	9,73	6,81	5,65	5,02	4,61	4,32	4,11	3,94	3,81	3,70	3,54	3,36	3,18	3,08	2,98	2,88	2,77	2,66
23	9,63	6,73	5,58	4,95	4,54	4,26	4,05	3,88	3,75	3,64	3,47	3,30	3,12	3,02	2,92	2,82	2,71	2,60
24	9,55	6,66	5,52	4,89	4,49	4,20	3,99	3,83	3,69	3,59	3,42	3,25	3,06	2,97	2,87	2,77	2,66	2,55
25	9,48	6,60	5,46	4,84	4,43	4,15	3,94	3,78	3,64	3,54	3,37	3,20	3,01	2,92	2,82	2,72	2,61	2,50
26	9,41	6,54	5,41	4,79	4,38	4,10	3,89	3,73	3,60	3,49	3,33	3,15	2,97	2,87	2,77	2,67	2,56	2,45
27	9,34	6,49	5,36	4,74	4,34	4,06	3,85	3,69	3,56	3,45	3,28	3,11	2,93	2,83	2,73	2,63	2,52	2,41

Tabla H. Valores críticos para la F de Snedecor ( $\alpha = 0,005$ ) (cont.)

$\frac{u}{v}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
28	9,28	6,44	5,32	4,70	4,30	4,02	3,81	3,65	3,52	3,41	3,25	3,07	2,89	2,79	2,69	2,59	2,48	2,37
29	9,23	6,40	5,28	4,66	4,26	3,98	3,77	3,61	3,48	3,38	3,21	3,04	2,86	2,76	2,66	2,56	2,45	2,33
30	9,18	6,35	5,24	4,62	4,23	3,95	3,74	3,58	3,45	3,34	3,18	3,01	2,82	2,73	2,63	2,52	2,42	2,30
40	8,83	6,07	4,98	4,37	3,99	3,71	3,51	3,35	3,22	3,12	2,95	2,78	2,60	2,50	2,40	2,30	2,18	2,06
60	8,49	5,79	4,73	4,14	3,76	3,49	3,29	3,13	3,01	2,90	2,74	2,57	2,39	2,29	2,19	2,08	1,96	1,83
120	8,18	5,54	4,50	3,92	3,55	3,28	3,09	2,93	2,81	2,71	2,54	2,37	2,19	2,09	1,98	1,87	1,75	1,61
$\infty$	7,88	5,30	4,28	3,72	3,35	3,09	2,90	2,74	2,62	2,52	2,36	2,19	2,00	1,90	1,79	1,67	1,53	1,36

Tabla I. Valores críticos para la R' del Test de Moore (tomado de Zar 1996)

$n$	$\alpha = 0,5$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,025$	$\alpha = 0,01$	$\alpha = 0,005$	$\alpha = 0,001$
2	0,791	1,049	1,058	1,060	1,061	1,061	1,061
3	0,693	1,039	1,095	1,124	1,143	1,149	1,154
4	0,620	1,008	1,090	1,146	1,192	1,212	1,238
5	0,588	0,988	1,084	1,152	1,216	1,250	1,298
6	0,568	0,972	1,074	1,152	1,230	1,275	1,345
7	0,556	0,959	1,066	1,150	1,238	1,291	1,373
8	0,546	0,949	1,059	1,148	1,242	1,300	1,397
9	0,538	0,940	1,053	1,146	1,245	1,307	1,416
10	0,532	0,934	1,048	1,144	1,248	1,313	1,432
12	0,523	0,926	1,042	1,140	1,252	1,322	1,456
14	0,518	0,920	1,037	1,136	1,252	1,325	1,470
16	0,514	0,914	1,031	1,132	1,250	1,327	1,480
18	0,510	0,910	1,027	1,129	1,248	1,328	1,487
20	0,507	0,906	1,024	1,127	1,247	1,329	1,492
22	0,505	0,903	1,022	1,126	1,246	1,330	1,496
24	0,503	0,901	1,021	1,125	1,246	1,331	1,499
26	0,502	0,899	1,019	1,124	1,246	1,332	1,501
28	0,500	0,897	1,018	1,124	1,246	1,333	1,502
30	0,499	0,896	1,016	1,123	1,245	1,334	1,502
40	0,494	0,891	1,012	1,119	1,243	1,332	1,504
60	0,489	0,887	1,007	1,115	1,241	1,329	1,506
80	0,487	0,883	1,005	1,113	1,240	1,329	1,508
100	0,485	0,881	1,004	1,112	1,240	1,329	1,509
$\infty$	0,481	0,876	0,999	1,109	1,239	1,329	1,517

**Tabla J.** Valores críticos para B del test de Mardia-Watson-Wheeler (tomado de Batschelet 1981)

$n_1$	$n_2$	$\alpha = 0,001$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,1$
8	4				6,83
9	3				6,41
9	4			8,29	4,88
10	3				6,85
10	4			9,47	6,24
10	5			10,47	6,85
11	3			7,20	5,23
11	4			10,42	7,43
11	5		12,34	8,74	6,6
12	3			7,46	5,73
12	4		11,20	8,46	7,46
12	5		13,93	10,46	7,46
12	6		14,93	11,20	7,46
13	3			7,68	6,15
13	4		11,83	9,35	7,03
13	5		15,26	10,15	7,39
13	6		17,31	10,42	8,04
14	3			7,85	6,49
14	4		12,34	9,30	7,6
14	5		16,39	10,30	7,85
14	6	19,20	15,59	12,21	7,94
14	7	20,20	16,39	11,65	8,85
15	3			7,99	6,78
15	4		12,78	8,74	7,91
15	5	17,35	14,52	10,36	7,91
15	6	20,92	17,48	11,61	9,12
15	7	22,88	16,14	11,57	9,06
16	3			8,11	5,83
16	4		13,14	9,44	7,38
16	5	18,16	15,55	10,44	9,03
16	6	22,43	16,98	11,54	9,11
16	7	25,57	18,16	12,66	9,78
17	3		8,21	7,23	6,14
17	4	13,44	11,76	9,74	7,64
17	5	18,86	16,44	11,03	8,76
17	6	23,73	17,76	12,21	9,41
17	7	27,40	17,98	12,63	10,11
17	8	29,37	19,11	13,36	10,15

**Tabla K.** Factor de corrección (K) para el test de Watson-Williams  
(tomado de Zar, 1996)

$r$	0	1	2	3	4	5	6	7	8	9
0		188,4989	94,7472	63,5015	47,8749	38,4992	32,2498	27,7851	24,4367	21,8335
0,01	19,7489	18,0444	16,6239	15,4219	14,3916	13,4986	12,7173	12,0278	11,4150	10,8667
0,02	10,3731	9,9266	9,5206	9,1500	8,8103	8,4976	8,2091	7,9419	7,6338	7,4628
0,03	7,2472	7,0455	6,8564	6,6787	6,5115	6,3539	6,2050	6,0641	5,9306	5,8040
0,04	5,6837	5,5693	5,4603	5,3564	5,2572	5,1625	5,7180	4,9850	4,9017	4,8219
0,05	4,7453	4,6717	4,6009	4,5328	4,4672	4,4039	4,3430	4,2841	4,2273	4,1724
0,06	4,1194	4,0680	4,0184	3,9703	3,9237	3,8785	3,8347	3,7922	3,7510	3,7109
0,07	3,6720	3,6342	3,5974	3,5616	3,5268	3,4930	3,4600	3,4278	3,3965	3,3660
0,08	3,3362	3,3072	3,2789	3,2512	3,2243	3,1979	3,1722	3,1470	3,1224	3,0984
0,09	3,0749	3,0519	3,0294	3,0074	2,9858	2,9668	2,9441	2,9239	2,9041	2,8846
0,10	2,8656	2,8469	2,8286	2,8107	2,7931	2,7758	2,7589	2,7423	2,7259	2,7099
0,11	2,6942	2,6787	2,6636	2,6487	2,6340	2,6196	2,6055	2,5915	2,5779	2,5644
0,12	2,5512	2,5382	2,5254	2,5128	2,5004	2,4882	2,4762	2,4644	2,4528	2,4413
0,13	2,4301	2,4189	2,4080	2,3912	2,3866	2,3792	2,3658	2,3557	2,3457	2,3358
0,14	2,3261	2,3165	2,3070	2,2977	2,2885	2,2794	2,2705	2,2616	2,2529	2,2443
0,15	2,2358	2,2275	2,2192	2,2110	2,2030	2,1950	2,1872	2,1794	2,1718	2,1642
0,16	2,1567	2,1494	2,1421	2,1349	2,1278	2,1208	2,1138	2,1070	2,1002	2,0945
0,17	2,0868	2,0803	2,0738	2,0674	2,0611	2,0549	2,0487	2,0426	2,0365	2,0305
0,18	2,0246	2,0188	2,0230	2,0072	2,0016	1,9960	1,9904	1,9849	1,9795	1,9741
0,19	1,9668	1,9635	1,9583	1,9532	1,9481	1,9430	1,9380	1,9331	1,9282	1,9233
0,20	1,9158	1,9137	1,9090	1,9043	1,8997	1,8951	1,8906	1,8661	1,8817	1,8772
0,21	1,8729	1,8685	1,8643	1,8600	1,8558	1,8516	1,8475	1,8434	1,8393	1,8353
0,22	1,8313	1,8274	1,8234	1,8195	1,8157	1,8119	1,8081	1,8043	1,8006	1,7969
0,23	1,7933	1,7896	1,7860	1,7825	1,7789	1,7754	1,7719	1,7685	1,7651	1,7617
0,24	1,7583	1,7550	1,7516	1,7484	1,7451	1,7419	1,7386	1,7355	1,7323	1,7292
0,25	1,7261	1,7230	1,7199	1,7169	1,7138	1,7108	1,7079	1,7049	1,7020	1,6991
0,26	1,6962	1,6933	1,6905	1,6877	1,6849	1,6821	1,6793	1,6766	1,6739	1,6712
0,27	1,6685	1,6658	1,6632	1,6606	1,6579	1,6554	1,6528	1,6502	1,6477	1,6452
0,28	1,6427	1,6402	1,6377	1,6353	1,6328	1,6304	1,6280	1,6256	1,6233	1,6209
0,29	1,6186	1,6162	1,6139	1,6116	1,6094	1,6071	1,6048	1,6026	1,6004	1,5982
0,30	1,5960	1,5938	1,5916	1,5895	1,5873	1,5852	1,5831	1,5810	1,5789	1,5768
0,31	1,5748	1,5727	1,5707	1,5687	1,5667	1,5647	1,5627	1,5607	1,5587	1,5568
0,32	1,5548	1,5529	1,5510	1,5491	1,5472	1,5453	1,5434	1,5419	1,5397	1,5369
0,33	1,5360	1,5342	1,5324	1,5306	1,5282	1,5270	1,5253	1,5235	1,5217	1,5200
0,34	1,5183	1,5165	1,5148	1,5131	1,5114	1,5097	1,5081	1,5064	1,5047	1,5031
0,35	1,5014	1,4998	1,4988	1,4966	1,4954	1,4934	1,4918	1,4902	1,4886	1,4871
0,36	1,4855	1,4839	1,4824	1,4809	1,4793	1,4778	1,4763	1,4748	1,4733	1,4718
0,37	1,4703	1,4689	1,4674	1,4659	1,4645	1,4630	1,4616	1,4602	1,4587	1,4573
0,38	1,4559	1,4545	1,4531	1,4517	1,4503	1,4490	1,4476	1,4462	1,4449	1,4435
0,39	1,4422	1,4408	1,4395	1,4382	1,4368	1,4355	1,4343	1,4329	1,4316	1,4303
0,40	1,4290	1,4277	1,4265	1,4252	1,4239	1,4227	1,4214	1,4202	1,4189	1,4177
0,41	1,4165	1,4152	1,4140	1,4128	1,4116	1,4104	1,4092	1,4080	1,4068	1,4056

**Tabla K.** Factor de corrección (K) para el test de Watson-Williams  
(tomado de Zar, 1996) (cont.)

$r$	0	1	2	3	4	5	6	7	8	9
0,42	1,4044	1,4033	1,4021	1,4009	1,3998	1,3986	1,3975	1,3963	1,3952	1,3940
0,43	1,3929	1,3918	1,3907	1,3895	1,3884	1,3873	1,3682	1,3851	1,3840	1,3829
0,44	1,3818	1,3808	1,3797	1,3786	1,3775	1,3765	1,3754	1,3744	1,3733	1,3723
0,45	1,3712	1,3702	1,3691	1,3681	1,3671	1,3660	1,3650	1,3640	1,3630	1,3620
0,46	1,3610	1,3600	1,3590	1,3580	1,357	1,3560	1,3550	1,3540	1,3530	1,3521
0,47	1,3510	1,3501	1,3492	1,3482	1,3472	1,3463	1,3453	1,3444	1,3434	1,3425
0,48	1,3416	1,3406	1,3397	1,3388	1,3378	1,3369	1,3360	1,3351	1,3342	1,3333
0,49	1,3324	1,3315	1,3306	1,3297	1,3288	1,3279	1,3270	1,3261	1,3252	1,3243
0,50	1,3235	1,3226	1,3217	1,3209	1,3200	1,3191	1,3183	1,3174	1,3166	1,3157
0,51	1,3148	1,3140	1,3132	1,3123	1,3115	1,3106	1,3098	1,3090	1,3081	1,3073
0,52	1,3065	1,3057	1,3049	1,3040	1,3032	1,3024	1,3016	1,3008	1,3000	1,2992
0,53	1,2984	1,2976	1,2968	1,2960	1,2952	1,2994	1,2936	1,2929	1,2921	1,2913
0,54	1,2905	1,2897	1,2890	1,2882	1,2874	1,2867	1,2859	1,2851	1,2844	1,2836
0,55	1,2829	1,2821	1,2814	1,2806	1,2799	1,2791	1,2784	1,2776	1,2769	1,2762
0,56	1,2754	1,2747	1,2740	1,2732	1,2725	1,2718	1,2710	1,2703	1,2696	1,2689
0,57	1,2682	1,2674	1,2627	1,2660	1,2653	1,2646	1,2639	1,2632	1,2625	1,2618
0,58	1,2611	1,2604	1,2597	1,2590	1,2582	1,2576	1,2569	1,2562	1,2555	1,2548
0,59	1,2542	1,2535	1,2598	1,2521	1,2514	1,2508	1,2501	1,2494	1,2487	1,2841
0,60	1,2474	1,2467	1,2461	1,2454	1,2447	1,2441	1,2434	1,2428	1,2421	1,2414
0,61	1,2408	1,2401	1,2395	1,2388	1,2382	1,2375	1,2369	1,2362	1,2356	1,2350
0,62	1,2343	1,2337	1,2330	1,2324	1,2318	1,2311	1,2305	1,2298	1,2292	1,2286
0,63	1,2280	1,2273	1,2267	1,2261	1,2254	1,2248	1,2242	1,2236	1,2230	1,2223
0,64	1,2217	1,2211	1,2205	1,2199	1,2192	1,2186	1,2180	1,2174	1,2168	1,2162
0,65	1,2156	1,2150	1,2144	1,2138	1,2132	1,2126	1,2120	1,2114	1,2108	1,2102
0,66	1,2096	1,2090	1,2084	1,2078	1,2072	1,2066	1,2060	1,2054	1,2048	1,2042
0,67	1,2036	1,2030	1,2024	1,2018	1,2013	1,2007	1,2001	1,1995	1,1989	1,1983
0,68	1,1797	1,1972	1,1966	1,1960	1,1954	1,1948	1,1943	1,1937	1,1931	1,1925
0,69	1,1920	1,1914	1,1908	1,1902	1,1897	1,1891	1,1885	1,1879	1,1874	1,1868
0,70	1,1862	1,1857	1,1851	1,1845	1,1840	1,1834	1,1828	1,1823	1,1817	1,1811
0,71	1,1806	1,1800	1,1794	1,1789	1,1783	1,1777	1,1772	1,1766	1,1761	1,1755
0,72	1,7449	1,1744	1,1738	1,1733	1,1727	1,1721	1,1716	1,1710	1,1705	1,1699
0,73	1,1694	1,1688	1,1682	1,1677	1,1671	1,1666	1,1660	1,1665	1,1649	1,1644
0,74	1,1638	1,1633	1,1627	1,1621	1,1616	1,1610	1,16,5	1,1599	1,1594	1,1588
0,75	1,1583	1,1577	1,1572	1,1566	1,1561	1,1555	1,1550	1,1544	1,1539	1,1533
0,76	1,1528	1,1522	1,1517	1,1511	1,1505	1,1500	1,1494	1,1489	1,1483	1,1478
0,77	1,1472	1,1467	1,1461	1,1456	1,1450	1,1445	1,1439	1,1434	1,1428	1,1423
0,78	1,1417	1,1412	1,1406	1,1401	1,1395	1,1389	1,1384	1,1378	1,1373	1,1367
0,79	1,1362	1,1356	1,1351	1,1345	1,1340	1,1334	1,1328	1,1323	1,1317	1,1312
0,80	1,1306	1,1300	1,1295	1,1289	1,1284	1,1278	1,1272	1,1267	1,1261	1,1256
0,81	1,1250	1,1244	1,1239	1,1233	1,1227	1,1222	1,1216	1,1210	1,1205	1,1199
0,82	4,1193	1,1188	1,1182	1,1176	1,1170	1,1165	1,1559	1,1153	1,1147	1,1142
0,83	1,1136	1,1130	1,1124	1,1119	1,1113	1,1107	1,1101	1,0950	1,1090	1,1084

**Tabla K.** Factor de corrección (K) para el test de Watson-Williams  
(tomado de Zar, 1996) (*cont.*)

<i>r</i>	0	1	2	3	4	5	6	7	8	9
0,84	1,1078	1,1072	1,1066	1,1060	1,1054	1,1049	1,1043	1,1037	1,1031	1,1025
0,85	1,1019	1,1013	1,1007	1,1001	1,0995	1,0989	1,0983	1,0977	1,0971	1,0965
0,86	1,0959	1,0953	1,0947	1,0941	1,0935	1,0928	1,0922	1,0916	1,0910	1,0904
0,87	1,0898	1,0892	1,0885	1,0879	1,0873	1,0867	1,0861	1,0854	1,0848	1,0842
0,88	1,0835	1,0829	1,0823	1,0816	1,0810	1,0804	1,0797	1,0791	1,0785	1,0778
0,89	1,0772	1,0765	1,0759	1,0752	1,0746	1,0740	1,0733	1,0727	1,0720	1,0713
0,90	1,0707	1,0700	1,0694	1,0687	1,0681	1,0674	1,0667	1,0661	1,0654	1,0647
0,91	1,0641	1,0634	1,0627	1,0621	1,0614	1,0607	1,0601	1,0594	1,0587	1,0580
0,92	1,0573	1,0567	1,0560	1,0553	1,0546	1,0539	1,0533	1,0526	1,0519	1,0512
0,93	1,0505	1,0498	1,0491	1,0484	1,0477	1,0470	1,0463	1,0456	1,0449	1,0443
0,94	1,0436	1,0429	1,0422	1,0414	1,0407	1,0400	1,0393	1,0386	1,0379	1,0372
0,95	1,0365	1,0358	1,0351	1,0344	1,0337	1,0330	1,0322	1,0315	1,0308	1,0301
0,96	1,0294	1,0287	1,0279	1,0272	1,0265	1,0258	1,0251	1,0243	1,0236	1,0229
0,97	1,0222	1,0214	1,0207	1,0200	1,0192	1,0185	1,0178	1,0170	1,0163	1,0156
0,98	1,0418	1,0141	1,0134	1,0126	1,0119	1,0112	1,0104	1,0097	1,0890	1,0082
0,99	1,0075	1,0067	1,0060	1,0052	1,0045	1,0037	1,0030	1,0022	1,0000	1,0000

Tabla L. Estimación de  $\tilde{\kappa}$  (tomado de Batschelet, 1981)

$n$	$r = 0,10$	$r = 0,15$	$r = 0,20$	$r = 0,25$	$r = 0,30$	$r = 0,35$	$r = 0,40$	$r = 0,45$	$r = 0,50$	$r = 0,55$	$r = 0,60$	$r = 0,65$	$r = 0,70$	$r = 0,75$	$r = 0,80$	$r = 0,85$	$r = 0,90$	$r = 0,95$
5	$\tilde{\kappa} = 0$	0	0	0	0	0	0	0,15	0,67	0,94	1,18	1,41	1,68	2,00	2,44	2,10	4,39	8,33
6	0	0	0	0	0	0	0	0,56	0,83	1,04	1,25	1,48	1,74	2,07	2,51	3,20	4,54	8,66
7	0	0	0	0	0	0	0,38	0,69	0,90	1,10	1,30	1,52	1,78	2,11	2,56	3,27	4,65	8,89
8	0	0	0	0	0	0	0,53	0,76	0,95	1,13	1,33	1,55	1,81	2,15	2,60	3,32	4,73	9,06
9	0	0	0	0	0	0,31	0,61	0,80	0,98	1,16	1,35	1,57	1,84	2,17	2,63	3,36	4,79	9,19
10	0	0	0	0	0	0,42	0,65	0,83	1,00	1,18	1,37	1,59	1,86	2,19	2,66	3,39	4,84	9,30
11	0	0	0	0	0	0,48	0,69	0,85	1,02	1,19	1,38	1,61	1,87	2,21	2,68	3,42	4,89	9,39
12	0	0	0	0	0,23	0,53	0,71	0,87	1,03	1,20	1,40	1,62	1,88	2,22	2,69	3,44	4,92	9,46
13	0	0	0	0	0,32	0,56	0,73	0,88	1,04	1,21	1,41	1,63	1,89	2,23	2,71	3,46	4,95	9,53
14	0	0	0	0	0,37	0,58	0,74	0,89	1,05	1,22	1,41	1,63	1,90	2,24	2,72	3,47	4,98	9,58
15	0	0	0	0	0,41	0,6	0,75	0,90	1,06	1,23	1,42	1,64	1,91	2,25	2,73	3,49	5,00	9,63
20	0	0	0	0,30	0,50	0,65	0,79	0,93	1,09	1,26	1,45	1,67	1,94	2,28	2,76	3,53	5,07	9,79
25	0	0	0	0,38	0,54	0,67	0,81	0,95	1,10	1,27	1,46	1,68	1,95	2,30	2,79	3,56	5,12	9,88
30	0	0	0,22	0,42	0,56	0,69	0,82	0,96	1,11	1,28	1,47	1,69	1,96	2,31	2,80	3,58	5,15	9,95
35	0	0	0,27	0,44	0,57	0,7	0,83	0,97	1,12	1,29	1,48	1,70	1,97	2,32	2,81	3,60	5,17	9,99
40	0	0	0,31	0,45	0,58	0,7	0,83	0,97	1,12	1,29	1,48	1,70	1,98	2,33	2,82	3,61	5,19	10,03
45	0	0,04	0,33	0,46	0,58	0,71	0,84	0,98	1,13	1,30	1,49	1,71	1,98	2,33	2,82	3,62	5,20	10,06
50	0	0,14	0,34	0,47	0,59	0,71	0,84	0,98	1,13	1,30	1,49	1,71	1,98	2,34	2,83	3,62	5,21	10,08
100	0	0,26	0,38	0,49	0,61	0,73	0,86	1,00	1,15	1,31	1,50	1,73	2,00	2,35	2,85	3,65	5,26	10,18
150	0,18	0,28	0,39	0,50	0,62	0,74	0,86	1,00	1,15	1,32	1,51	1,73	2,00	2,36	2,86	3,66	5,27	10,21
200	0,19	0,29	0,4	0,51	0,62	0,74	0,87	1,00	1,15	1,32	1,51	1,73	2,01	2,36	2,86	3,67	5,28	10,22
$\infty$	0,20	0,30	0,41	0,52	0,63	0,75	0,87	1,01	1,16	1,33	1,52	1,74	2,01	2,37	2,87	3,68	5,31	10,27