



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

BONDAD DE AJUSTE PARA PROBAR ALEATORIEDAD  
EN LA ESTADÍSTICA CIRCULAR

T E S I S  
QUE PARA OBTENER EL TÍTULO DE  
MATEMÁTICA CON ORIENTACIÓN EN  
ESTADÍSTICA MATEMÁTICA

PRESENTA:

LETICIA SANCHEZ LOPEZ

DIRECTORA DE TESIS:  
MAT. MARGARITA ELVIRA CHAVEZ CANO



287560



FACULTAD DE CIENCIAS  
UNAM

FACULTAD DE CIENCIAS  
SECCION ESCOLAR



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**MAT. MARGARITA ELVIRA CHÁVEZ CANO**  
**Jefa de la División de Estudios Profesionales**  
**P r e s e n t e**

Comunicamos a usted que hemos revisado el trabajo de Tesis:

**“BONDAD DE AJUSTE PARA PROBAR ALEATORIEDAD EN LA ESTADÍSTICA CIRCULAR”**

realizado por: Leticia Sánchez López

Con número de cuenta 8817660-8 , pasante de la carrera de Matemáticas.

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de tesis  
Propietario

Mat. Margarita Elvira Chávez Cano

Propietario

M. en C. Gerardo Rivas Lechuga

Propietario

M. en C. Beatriz Eugenia Rodríguez Fernández

Suplente

M. en C. Inocencio Rafael Madrid Ríos

Suplente

Act. Mónica Ángeles Velázquez

**Consejo Departamental de Matemáticas.**

*Héctor Méndez Lango*

Dr. Héctor Méndez Lango

A mi maestra Margarita, por su apoyo incondicional y su confianza en mí;  
A mis hijos, por la motivación que me dan;  
A todos quienes, aún sin saberlo, me han ayudado a llegar a esta meta.

# Índice General

1	GENERALIDADES.	9
2	PRUEBAS DE BONDAD DE AJUSTE PARA PROBAR UNIFORMIDAD	13
2.1	PRUEBA DE RAYLEIGH . . . . .	14
2.1.1	HISTORIA . . . . .	14
2.1.2	PROPÓSITO . . . . .	14
2.1.3	DATOS REQUERIDOS . . . . .	14
2.1.4	IDEA BÁSICA DE LA PRUEBA . . . . .	15
2.1.5	HIPÓTESIS A PROBAR . . . . .	16
2.1.6	DISTRIBUCION DE "R". . . . .	17
2.1.7	DISTRIBUCIÓN ASINTÓTICA DE $2R/n^2$ . . . . .	28
2.1.8	INVARIANZA BAJO ROTACIONES . . . . .	30
2.2	PRUEBA DE GREENWOOD Y DURAND . . . . .	33
2.2.1	HISTORIA . . . . .	33
2.2.2	PROPÓSITO . . . . .	33
2.2.3	DATOS REQUERIDOS . . . . .	34

2.2.4	IDEA BÁSICA DE LA PRUEBA . . . . .	34
2.2.5	HIPÓTESIS A PROBAR . . . . .	35
2.2.6	DISTRIBUCIÓN DE $\bar{x}$ Y $v$ . . . . .	35
2.2.7	DISTRIBUCIÓN DE "X" . . . . .	38
2.3	PRUEBA DEL RANGO . . . . .	40
2.3.1	HISTORIA . . . . .	40
2.3.2	DATOS . . . . .	41
2.3.3	PROPÓSITO . . . . .	41
2.3.4	IDEA BÁSICA DE LA PRUEBA . . . . .	41
2.3.5	OBSERVACIONES . . . . .	42
2.3.6	HIPÓTESIS . . . . .	42
2.3.7	ESTADÍSTICA DE PRUEBA . . . . .	42
2.3.8	REGLA DE DECISIÓN . . . . .	43
2.3.9	DISTRIBUCIÓN DE $w$ . . . . .	43
2.4	PRUEBA DE ESPACIOS DE RAO . . . . .	44
2.4.1	HISTORIA . . . . .	44
2.4.2	DATOS . . . . .	45
2.4.3	OBSERVACIONES . . . . .	45
2.4.4	PROPÓSITO . . . . .	45
2.4.5	IDEA BÁSICA DE LA PRUEBA . . . . .	45
2.4.6	HIPÓTESIS NULA . . . . .	46
2.4.7	ESTADÍSTICA PRUEBA . . . . .	46
2.4.8	REGLA DE DECISIÓN . . . . .	46

	5
2.4.9 DISTRIBUCIÓN DE LA ESTADÍSTICA . . . . .	17
2.5 PRUEBA DE HODGES Y AJNE . . . . .	48
2.5.1 HISTORIA . . . . .	48
2.5.2 PROPÓSITO . . . . .	48
2.5.3 DATOS REQUERIDOS . . . . .	49
2.5.4 IDEA BÁSICA DE LA PRUEBA . . . . .	49
2.5.5 HIPÓTESIS A PROBAR . . . . .	50
2.5.6 REGLA DE DECISIÓN . . . . .	50
2.5.7 INVARIANZA BAJO ROTACIONES . . . . .	50
<b>3 BONDAD DE AJUSTE, UN ENFOQUE GENERAL</b>	<b>53</b>
3.1 PRUEBA DE LA JI CUADRADA . . . . .	53
3.1.1 HISTORIA . . . . .	53
3.1.2 DATOS . . . . .	54
3.1.3 PROPÓSITO . . . . .	54
3.1.4 IDEA BÁSICA DE LA PRUEBA . . . . .	54
3.1.5 OBSERVACIONES . . . . .	55
3.1.6 ESTADÍSTICA DE PRUEBA . . . . .	55
3.1.7 DISTRIBUCIÓN ASINTÓTICA . . . . .	55
3.1.8 DEFINICIÓN DE LA PRUEBA . . . . .	59
3.2 PRUEBA DE KUIPER . . . . .	61
3.2.1 HISTORIA . . . . .	61
3.2.2 DATOS . . . . .	61

3.2.3	PROPÓSITO	62
3.2.4	IDEA BÁSICA DE LA PRUEBA	62
3.2.5	OBSERVACIONES	62
3.2.6	ESTADÍSTICA DE PRUEBA	63
3.2.7	DEFINICIÓN DE LA PRUEBA	63
3.2.8	INVARIANZA BAJO ROTACIONES	63
3.3	PRUEBA $U^2$ DE WATSON	67
3.3.1	HISTORIA	67
3.3.2	DATOS	67
3.3.3	PROPÓSITO	68
3.3.4	IDEA BÁSICA DE LA PRUEBA	68
3.3.5	OBSERVACIONES	68
3.3.6	ESTADÍSTICA DE PRUEBA	68
3.3.7	DEFINICIÓN DE LA PRUEBA	68
3.3.8	DISTRIBUCIÓN DE $U_n^2$	69
3.3.9	INVARIANZA DE LA PRUEBA	71
3.3.10	DISTRIBUCIÓN APROXIMADA DE $U_n^2$	72
<b>4</b>	<b>APLICACIONES</b>	<b>75</b>
4.1	PRUEBA DE RAYLEIGH	80
4.2	PRUEBA DEL RANGO	88
4.3	PRUEBA DE ESPACIOS DE RAO	92
4.4	PRUEBA DE HODGES Y AJNE	102
4.5	PRUEBA DE KUIPER	104
4.6	PRUEBA DE WATSON	112

# INTRODUCCIÓN

La teoría desarrollada sobre la Estadística Circular es muy reciente, ésta tiene como objetivo el análisis de datos distribuidos en una circunferencia, por lo que cualquier tipo de observaciones en un plano o referidas a un fenómeno cíclico están dentro de su estudio.

Pensando en una analogía con la Estadística clásica, se ha construido una base teórica en la que los conceptos y resultados conocidos en relación a estadística descriptiva, funciones de distribución, pruebas de hipótesis, análisis de regresión, análisis multivariado, etc., se han adaptado o replanteado para que en la Estadística Circular tengan sentido.

Esta tesis se centra en el estudio de la Bondad de Ajuste para pruebas de aleatoriedad de una muestra de observaciones circulares, el cual incluirá el análisis matemático de las mismas y su aplicación, para ampliar así el panorama práctico que hasta ahora ha sido limitado, de acuerdo a la bibliografía con que se cuenta.

No hay un material que conjunte de manera clara, sencilla y completa, estos dos aspectos sobre el tema: el teórico y el práctico.

Para cumplir con dicho objetivo, se plantea en el Capítulo 1, una breve explicación de conceptos básicos que se manejan en Estadística Circular, así como la definición de la función de distribución uniforme, que será la hipótesis nula en todas las pruebas que se presentarán.

En los Capítulos 2 y 3, se analizan varias pruebas; la razón de su separación está en el planteamiento de la hipótesis alternativa, ya que en el primer caso se definirá

una función de distribución específica y en el segundo caso, que es más general, no se requiere plantear una función de distribución para la hipótesis alternativa.

En el Capítulo 4, se realiza una aplicación inédita de las pruebas de aleatoriedad en estadística circular, ya que sólo se tienen aplicaciones en relación con la migración de animales, además se dan ideas sobre otras aplicaciones que son de interés en otras áreas del conocimiento como son la Medicina, la Geología y la Ecología.

Finalmente se encuentran las conclusiones de la investigación desarrollada y un apéndice donde están las tablas utilizadas para la aplicación de las pruebas.

# Capítulo 1

## GENERALIDADES.

En las pruebas de aleatoriedad se utilizan conceptos básicos generales que se analizan a continuación con el fin de comprender mejor la teoría que se presenta en esta tesis.

En estadística circular las direcciones se representan con puntos sobre el círculo unitario, a cada observación se le asocia un ángulo  $\alpha$  que mide el *acimut*, que es el ángulo que va desde el Norte hasta el lugar donde se encuentra la observación, en el sentido en el que avanzan las manecillas del reloj. El Norte corresponderá a la dirección cero o el origen de la circunferencia.

De la misma manera que a cada dirección de una muestra aleatoria, representada por un punto en la circunferencia unitaria, se le asocia un ángulo; también se le asocia un vector unitario:  $e_1, e_2, \dots, e_n$ .

Se define al *vector medio* como el vector resultante de la suma de los vectores unitarios, dividido entre el número de éstos:

$$m = \frac{(e_1 + e_2 + \dots + e_n)}{n}$$

Al ángulo que forma el vector medio con el eje positivo X se llama *ángulo medio* de la muestra y lo denotaremos como  $\bar{x}_0$ . Sean  $R$  la longitud del vector resultante y  $r$  la longitud del vector medio; esto es:

$$R = \left| \sum_{i=1}^n c_i \right|, \quad r = |m| \Rightarrow r = \frac{R}{n}$$

$$0 \leq R \leq n \quad 0 \leq r \leq 1$$

Se puede escribir a  $r$  usando coordenadas rectangulares:

Si  $\bar{x}$  y  $\bar{y}$  son las coordenadas rectangulares del centro de gravedad o de masa de los puntos, entonces:

$$\begin{aligned} \bar{x} &= \frac{1}{n} (x_1 + x_2 + \dots + x_n) & \bar{y} &= \frac{1}{n} (y_1 + y_2 + \dots + y_n) \\ &= \frac{1}{n} (\cos \theta_1 + \cos \theta_2 + \dots + \cos \theta_n) & &= \frac{1}{n} (\sin \theta_1 + \sin \theta_2 + \dots + \sin \theta_n) \\ R &= \sqrt{\left( \sum_{i=1}^n \cos \theta_i \right)^2 + \left( \sum_{i=1}^n \sin \theta_i \right)^2} \Rightarrow r = \frac{1}{n} \sqrt{\left( \sum_{i=1}^n \cos \theta_i \right)^2 + \left( \sum_{i=1}^n \sin \theta_i \right)^2} \end{aligned}$$

Al obtener la proyección del vector medio sobre otro vector con dirección  $\theta_0$  se pueden presentar dos situaciones:

i) Si  $\theta_0 = 0$  (Fig. 1), la proyección será precisamente:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i$$

ii) Si  $\theta_0 \neq 0$  (Fig.2), la proyección será:

$$p = r \cos(\bar{x}_0 - \theta_0)$$

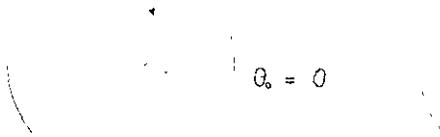


Fig. 1



Fig. 2

Los fenómenos de ocurrencia cíclica en los que pudiera existir más de una moda, son numerosos. Por ejemplo, se han registrado las direcciones que siguen ciertas especies de aves que al liberarlas en medio de un lago, vuelan siempre hacia la orilla más cercana; en ese caso, dependiendo de la situación geográfica se podría ajustar la muestra a una distribución bimodal.

El índice de contaminación máxima en un día debería presentarse en las horas en que hay más coches circulando, (cuando se inicia y termina la jornada de trabajo matutina); que también sería una distribución bimodal.

Si el ángulo entre las dos modas es arbitrario, no existe un método para transformar la muestra en otra que sea unimodal, pero si este ángulo es de  $180^\circ$ , es decir, los datos están diametralmente opuestos entre sí, se dice que se tiene una muestra aleatoria con datos axiales y en este caso se puede implementar una manera de convertir la muestra en otra que sea unimodal.

En Estadística Circular las muestras aleatorias con que se trabaja son muestras aleatorias de ángulos:  $\theta_1, \theta_2, \dots, \theta_n$ .

La distribución más importante en la Estadística Circular es la de "von Mises",

pues es análoga a la distribución Normal de la estadística convencional:

Una variable aleatoria  $\theta$  tiene distribución von Mises si su función de densidad de probabilidad está dada por:

$$f(\theta; \mu_0, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu_0)} \quad 0 < \theta \leq 2\pi, \quad \kappa > 0, \quad 0 \leq \mu_0 < 2\pi$$

donde  $I_0(\kappa)$  es la función de Bessel modificada:

$$I_0(\kappa) = \sum_{r=0}^{\infty} \frac{1}{r!^2} \left(\frac{1}{2}\kappa\right)^{2r}$$

$\mu_0$  es el parámetro que indica la dirección media de la distribución y  $\kappa$  describe su concentración, si este último es cero:

$$f(\theta; \mu_0, \kappa) = \frac{1}{2\pi}, \quad 0 \leq \theta \leq 2\pi$$

que es la f.d.p. uniforme en estadística circular. Conforme  $\kappa$  crece, la distribución se concentra más alrededor de  $\mu_0$

Otro concepto que es necesario revisar antes de analizar las pruebas de hipótesis es el de *longitud de arco*, que mide las diferencias en grados, de las observaciones adyacentes. Sean  $\theta_{(1)}, \theta_{(2)}, \dots, \theta_{(n)}$  las estadísticas de orden lineales de  $\theta_1, \theta_2, \dots, \theta_n$   
 $0 < \theta_i \leq 2\pi$

$$T_i = \theta_{(i+1)} - \theta_i, \quad i = 1, 2, \dots, n-1, \quad T_n = 2\pi - \theta_{(n)} + \theta_{(1)}$$

Desde los inicios de la estadística, los análisis se han comenzado proponiendo una distribución para las observaciones, posteriormente se verifica si ésta es cierta. El estudio de este tipo de procedimientos se conoce como *bondad de ajuste*.

## Capítulo 2

# PRUEBAS DE BONDAD DE AJUSTE PARA PROBAR UNIFORMIDAD

Se presentarán cinco pruebas de este tipo. La hipótesis nula para todos los casos establece que la población de la que se extrajo la muestra tiene distribución uniforme, en lo que respecta a la hipótesis alternativa, en algunos casos ésta podrá establecer una distribución específica, que por lo general se tratará de una distribución von Mises.

## 2.1 PRUEBA DE RAYLEIGH

### 2.1.1 HISTORIA

Como se analizará más adelante, la prueba de Rayleigh tiene como estadística de prueba a “ $r$ ”, la longitud del vector medio. La interpretación de este valor fue estudiada primero por lord Rayleigh en 1880, pero fue después de 25 años cuando J.C. Kluiver calcula la distribución exacta de  $r$  bajo la hipótesis nula, sin embargo las integrales que necesitan ser calculadas en la función de distribución de  $r$  son muy difíciles de manejar, pues quedan en términos de funciones de Bessel, por tal razón personas como A. Greenwood, D. Durand, K. Pearson y Rayleigh han realizado valiosos intentos para obtener aproximaciones.

### 2.1.2 PROPÓSITO

Probar si la población de la cual es extraída la muestra difiere significativamente de la aleatoriedad, es decir, si existe evidencia estadística de orientación o tendencia de la población hacia una orientación determinada.

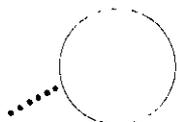
### 2.1.3 DATOS REQUERIDOS

Una muestra aleatoria de  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$  que representan direcciones en el plano o los tiempos de ocurrencia de un fenómeno cíclico.

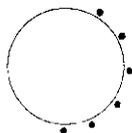
### 2.1.4 IDEA BÁSICA DE LA PRUEBA

La longitud del *vector medio*  $r$ , que varía entre 0 y 1, es una medida descriptiva en la estadística circular, éste se relaciona con la dispersión de las observaciones de la siguiente manera: si todas las observaciones coinciden,  $r = 1$ ; en este caso resulta ser obvio que no existe aleatoriedad, pero a medida que los datos se extienden en la circunferencia unitaria,  $r$  empieza a decrecer y en el caso extremo, es decir cuando los datos se dispersan totalmente en la circunferencia,  $r$  es muy cercano a 0, en este caso la muestra estaría indicando aleatoriedad o uniformidad en su comportamiento.

En las siguientes gráficas se puede constatar:



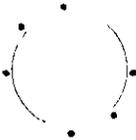
$$r = \left[ \left( \frac{1}{6} (6 \text{Cos} 250^\circ) \right)^2 + \left( \frac{1}{6} (6 \text{Sen} 250^\circ) \right)^2 \right]^{\frac{1}{2}} = 1$$



$\theta_i$	$x_i = \text{Cos } \theta_i$	$y_i = \text{Sen } \theta_i$
$\theta_1 = 30^\circ$	0.8660	0.5
$\theta_2 = 60^\circ$	0.5	0.8660
$\theta_3 = 90^\circ$	0	1
$\theta_4 = 120^\circ$	-0.5	0.8660
$\theta_5 = 150^\circ$	-0.8660	0.5
$\theta_6 = 180^\circ$	-1	0
Total	-1	3.732

$$\bar{x} = 0.1667, \bar{y} = 0.662$$

$$r = \left[ (0.1667)^2 + (0.552)^2 \right]^{\frac{1}{2}} = 0.6440$$



$\theta_i$	$x_i = \text{Cos } \theta_i$	$y_i = \text{Sen } \theta_i$
$\theta_1 = 0^\circ$	1	0
$\theta_2 = 90^\circ$	0	1
$\theta_3 = 135^\circ$	-0.7071	0.7071
$\theta_4 = 180^\circ$	-1	0
$\theta_5 = 270^\circ$	0	-1
$\theta_6 = 315^\circ$	0.7071	-0.7071
Total	0	0

$$\bar{x} = 0, \bar{y} = 0$$

$$r = 0$$

Esto conduce a pensar de manera natural que  $r$ , la **longitud del vector medio** debe funcionar como una buena estadística de prueba.

Si  $n$  es muy grande, la estadística de prueba que se usa es  $\frac{r^2}{n}$ , resultado que justificaremos después.

### 2.1.5 HIPÓTESIS A PROBAR

Las hipótesis nula y alternativa, son:

$H_0$  : La población se distribuye uniformemente (hay aleatoriedad)

$H_a$  : La población se distribuye bajo una dirección definida.

En este caso se define una prueba por cociente de verosimilitudes donde  $H_0$  indicará que la población se distribuye como una von Mises( $\kappa, \Lambda$ )

### 2.1.6 DISTRIBUCION DE "R".

Considere el siguiente problema:

Una persona está en el punto  $(0,0)$  y camina cierta distancia en línea recta, entonces gira un ángulo  $\theta$  y camina por segunda vez una nueva distancia; el proceso se repite  $n$  veces. Queremos ahora encontrar la probabilidad de que después de  $n$  pasos, se encuentre a una distancia dada del punto inicial de partida, es decir, del origen.

Obsérvese la similitud que existe entre el problema de encontrar la distribución de  $r$  y el mencionado anteriormente. La diferencia principal consiste en el tipo de observaciones que se tienen, en el primero, se pueden pensar como coordenadas polares y en el segundo como coordenadas rectangulares, esto conducirá a trabajar con algunos cambios de variable para pasar de un problema a otro.

Se definen entonces:

$l_1, l_2, \dots, l_n$  las longitudes de los  $n$  pasos con direcciones  $\theta_1, \theta_2, \dots, \theta_n$ .

Se supone que se trata de una muestra aleatoria y que :

$$f(\theta_i) = \frac{1}{2\pi}, 0 \leq \theta_i \leq 2\pi \text{ (Bajo } H_0)$$

Los  $n$  vectores con componentes :  $(l_i \sin \theta_i, l_i \cos \theta_i)$

Sean:

$$X = \sum_{i=1}^n l_i \cos \theta_i$$

$$Y = \sum_{i=1}^n l_i \sin \theta_i$$

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

Se necesita encontrar la distribución de  $R$ . Reduciendo el problema definiéndolo en términos más simples:

$$x = l \cos \theta$$

$$y = l \sin \theta$$

Por el Teorema de Inversión sabemos que es posible escribir la f.d.p. conjunta de  $x$  y  $y$  en términos de su función característica.

Sea  $\psi(t_1, t_2)$  la función característica conjunta de  $x$  y  $y$ :

$$\psi(t_1, t_2) = E(e^{it_1x + it_2y})$$

entonces:

$$f_{(x,y)}(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{(-it_1x - it_2y)} \psi(t_1, t_2) dt_1 dt_2$$

Sean:

$$t_1 = \rho \cos \Phi$$

$$t_2 = \rho \sin \Phi$$

Sustituyendo con los cambios de variable  $x, y, t_1, t_2$  tenemos que:

$$e^{(it_1x + it_2y)} = e^{[i(\rho \cos \Phi)(l \cos \theta) + i(\rho \sin \Phi)(l \sin \theta)]} = e^{[i\rho l(\cos \Phi \cos \theta + \sin \Phi \sin \theta)]} = e^{[i\rho l \cos(\theta - \Phi)]}$$

entonces

$$E(e^{i\rho l \cos(\theta - \Phi)}) = \Psi(\rho, \Phi)$$

Conociendo  $f_{x,y}$ , se puede obtener  $f_{l,\theta}$ :

$$f_{l,\theta}(l, \theta) = f_{x,y}(x, y) |J(x, y)|^{-1}$$

$$\text{Cuando } \quad x = l \cos \theta \quad \quad y = l \sin \theta :$$

$$|J(x, y)|^{-1} = l$$

$$\text{y cuando } t_1 = \rho \cos \Phi \quad \quad t_2 = \rho \sin \Phi$$

$$|J(t_1, t_2)|^{-1} = \rho$$

por lo cual:

$$f_{l,\theta}(l, \theta) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{\infty} e^{(-i\rho l \cos(\theta - \Phi))} l \rho \Psi(\rho, \Phi) d\rho d\Phi$$

Integrando sobre el recorrido de  $\theta$  y acomodando de manera conveniente el orden de integración, obtenemos:

$$f_l(l) = \frac{1}{(2\pi)^2} \int_0^{\infty} \int_0^{2\pi} \left[ \int_0^{2\pi} e^{(-i\rho l \cos(\theta - \Phi))} d\theta \right] \rho \Psi(\rho, \Phi) d\rho d\Phi$$

Esto es válido ya que:

$$H = \int_0^{2\pi} e^{(-i\rho l \cos(\theta - \Phi))} d\rho d\Phi$$

no depende de  $\Phi$ , además aplicando la teoría de las funciones de Bessel[2], sabemos

que

$$\frac{1}{2\pi}H = J_0(\rho, l)$$

y sustituyendo en la f.d.p. de  $l$ , obtenemos .

$$f_l(l) = \frac{1}{2\pi}l \int_0^\infty \int_0^{2\pi} \rho J_0(\rho, l) \Psi(\rho, \Phi) d\rho d\Phi$$

Ahora es tiempo de volver a nuestro problema inicial, que era encontrar la f.d.p.

de  $R = \sqrt{X^2 + Y^2}$ . En este caso la función característica de  $x$  y  $y$  será.

$$\prod_{i=1}^n J_0(l_i \rho)$$

y

$$f(R) = R \int_0^\infty u J_0(l_i u) du$$

Como en nuestro caso  $l_1 = l_2 = \dots = l_n = 1$ ,

$$f(r) = r \int_0^\infty u J_0(ur) J_0^n(u) du$$

La teoría de funciones de Bessel nos ayuda de nuevo pues :

$$\frac{d(xJ_1(x))}{dx} = xJ_0(x)$$

por lo tanto

$$F(r) = r \int_0^\infty J_1(ur) J_0^n(u) du$$

que es la expresión mínima que podemos encontrar de la función de distribución de

$r$ .

Existe otra manera de encontrar la distribución de  $r$ , se analizará el caso en el que se tengan dos observaciones:

Se definen  $l_i, \theta_i, \quad i = 1, 2$  como antes y nos referiremos además como  $S_2$  al vector resultante de  $l_1$  y  $l_2$

Sea  $P_2(r, l_1, l_2)$  la probabilidad de que después de dos desplazamientos, la distancia al punto de partida sea menor que  $r$ :

$$P_2(r; l_1, l_2) = \frac{1}{2\pi} \int d\theta_1$$

el problema es conocer qué valores puede tomar  $\theta_1$ , se sabe que sólo puede tomar los valores tales que  $S_2 \leq r$ . Para que se cumpla esto, Kluyver introdujo un "factor de discontinuidad" [5] definido como:

$$r \int_0^{\infty} J_1(rx) J_0(S_2 x) dx = \begin{cases} 0 & \text{si } S_2 > r \\ 1 & \text{si } S_2 < r \end{cases}$$

entonces:

$$P_2(r; l_1, l_2) = \frac{r}{2\pi} \int_0^{2\pi} d\theta_1 \int_0^{\infty} J_1(rx) J_0(l_2 x)$$

además por ser  $S_2$  el vector resultante,  $S_2 = l_1^2 + l_2^2 - 2l_1 l_2 \cos \theta_1$ , por lo que:

$$\int_0^{2\pi} J_0 \left( \left( l_1^2 + l_2^2 - 2l_1 l_2 \cos \theta \right)^{\frac{1}{2}} \right) d\theta = 2\pi J_0(l_1) J_0(l_2)$$

De esta manera se puede escribir

$$P_2(r; l_1, l_2) = r \int_0^{\infty} J_1(rx) J_0(l_1 x) J_0(l_2 x) dx$$

Si se tienen  $n$  observaciones, realizando el procedimiento anterior  $n-1$  veces se obtiene que

$$P_n(r; l_1, l_2, \dots, l_n) = r \int_0^{\infty} J_1(rx) J_0(l_1x) J_0(l_2x) \dots J_0(l_nx) dx$$

y como  $l_i = 1 \quad \forall i$

$$P_n(r, l) = r \int_0^{\infty} J_1(rx) J_0^n(x) dx$$

que es el mismo resultado que obtenido con el primer procedimiento.

Con la hipótesis nula que se está estudiando, la hipótesis alternativa puede establecer que la población tenga una distribución von Mises, sin embargo de esta última se derivan en total doce posibles modalidades, ya que de acuerdo a los intereses que se tengan, se pueden variar los parámetros de la siguiente manera:

Para  $\mu$  :

$$0 \leq \mu \leq 2\pi$$

$$\mu_1 \leq \mu \leq \mu_2$$

$$\mu = \mu_1$$

Para  $\kappa$  :

$$\kappa = \kappa^* > 0 \quad (\kappa^* \text{ desconocido}) \quad \kappa = \kappa_1$$

$$\kappa_1 \leq \kappa \quad \kappa_1 \leq \kappa \leq \kappa_2$$

Se analizará el caso más difícil, que es cuando desconocemos ambos parámetros, es decir:

Bajo  $H_0: \kappa = 0$  y Bajo  $H_a: \kappa = \kappa^* > 0$

$$0 \leq \mu < 2\pi$$

$$\Theta = \{(\kappa, \mu) / 0 \leq \kappa < \infty; 0 \leq \mu < 2\pi\}$$

$$\omega = \{(\kappa, \mu) / \kappa = 0\}$$

Las funciones de verosimilitud son:

$$L(\omega) = \prod_{i=1}^n \frac{1}{2\pi} = \frac{1}{(2\pi)^n} \quad \text{si } 0 < \theta_i \leq 2\pi, \quad i = 1, 2, \dots, n$$

$$L(\Theta) = \prod_{i=1}^n \frac{e^{\kappa^* \cos(\theta_i - \mu)}}{2\pi I_0(\kappa^*)} \quad \text{si } 0 < \theta_i \leq 2\pi, \kappa > 0, 0 \leq \mu < 2\pi$$

$$= [2\pi I_0(\kappa^*)]^{-n} e^{\left[\sum_{i=1}^n \kappa^* \cos(\theta_i - \mu)\right]}$$

$$\ln L(\Theta) = -n \ln [2\pi I_0(\kappa^*)] + \sum_{i=1}^n \kappa^* \cos(\theta_i - \mu)$$

$$= -n \ln 2\pi - n \ln I_0(\kappa^*) + \kappa^* \sum_{i=1}^n \cos(\theta_i - \mu)$$

Derivando con respecto a  $\mu$  para obtener el estimador máximo verosímil de  $\mu$ :

$$\frac{\partial \ln L(\Theta)}{\partial \mu} = \kappa^* \sum_{i=1}^n \text{Sen}(\theta_i - \mu)$$

$$\text{Si } \kappa^* \sum_{i=1}^n \text{Sen}(\theta_i - \mu) = 0 \implies \sum_{i=1}^n \text{Sen } \theta_i \cos \mu - \sum_{i=1}^n \text{Sen } \mu \cos \theta_i = 0$$

$$\implies \frac{\text{Sen } \mu}{\cos \mu} = \frac{\sum_{i=1}^n \text{Sen } \theta_i}{\sum_{i=1}^n \cos \theta_i} = \frac{Y}{X}$$

pero el ángulo que cumple que  $\frac{\text{Sen } \mu}{\cos \mu} = \frac{Y}{X}$  es  $\bar{x}_0$ : ángulo formado con el vector resultante y el eje de las  $x$ 's

Se deriva con respecto a  $\kappa^*$  para obtener el estimador máximo verosímil de  $\kappa^*$ .

$$\frac{\partial \ln L(\Theta)}{\partial \kappa^*} = \frac{-n \partial \ln I_0(\kappa^*)}{\partial \kappa^*} + \sum_{i=1}^n \text{Cos}(\theta_i - \mu)$$

Por definición [2]:

$$I_p(\kappa^*) = \sum_{r=0}^{\infty} \{\Gamma(p+r+1) \Gamma(r+1)\}^{-1} \left(\frac{\kappa^*}{2}\right)^{2r+p}$$

por lo cual:

$$I_0(\kappa^*) = \sum_{r=0}^{\infty} \frac{1}{\Gamma(r+1)\Gamma(r+1)} \left(\frac{\kappa^*}{2}\right)^{2r} = 1 + \left(\frac{\kappa^*}{2}\right)^2 + \frac{1}{2! \cdot 2!} \left(\frac{\kappa^*}{2}\right)^4 + \frac{1}{3! \cdot 3!} \left(\frac{\kappa^*}{2}\right)^6 + \dots$$

$$\frac{\partial I_0(\kappa^*)}{\partial \kappa^*} = \left(\frac{\kappa^*}{2}\right) + \frac{1}{2} \left(\frac{\kappa^*}{2}\right)^3 + \frac{1}{2 \cdot 3!} \left(\frac{\kappa^*}{2}\right)^5 + \dots$$

Por otro lado:

$$I_1(\kappa^*) = \sum_{r=0}^{\infty} \{\Gamma(r+2) \Gamma(r+1)\}^{-1} \left(\frac{\kappa^*}{2}\right)^{2r+1}$$

$$= \left(\frac{\kappa^*}{2}\right) + \frac{1}{2} \left(\frac{\kappa^*}{2}\right)^3 + \frac{1}{2 \cdot 3!} \left(\frac{\kappa^*}{2}\right)^5 + \dots$$

Lo cual muestra que :

$$I_1(\kappa^*) = \frac{\partial I_0(\kappa^*)}{\partial \kappa^*} \implies \frac{\partial \ln I_0(\kappa^*)}{\partial \kappa^*} = \frac{I_1(\kappa^*)}{I_0(\kappa^*)}$$

Entonces:

$$\frac{\partial \ln L(\Theta)}{\partial \kappa^*} = -n \frac{I_1(\kappa^*)}{I_0(\kappa^*)} + \sum_{i=1}^n \text{Cos}(\theta_i - \mu)$$

Igualando a cero:

$$\begin{aligned} \frac{I_1(\kappa^*)}{I_0(\kappa^*)} &= \frac{1}{n} \sum_{i=1}^n \text{Cos}(\theta_i - \mu) = \frac{1}{n} \sum_{i=1}^n \text{Cos} \theta_i \text{Cos} \mu + \sum_{i=1}^n \text{Sen} \theta_i \text{Sen} \mu \\ &= \frac{1}{n} \text{Cos} \mu \sum_{i=1}^n \text{Cos} \theta_i + \text{Sen} \mu \sum_{i=1}^n \text{Sen} \theta_i \end{aligned}$$

y como  $\mu$  es estimado por el vector  $\bar{x}_0$ :

$$\begin{aligned} \frac{I_1(\kappa^*)}{I_0(\kappa^*)} &= \frac{1}{n} \left[ \sum_{i=1}^n \frac{\text{Cos} \theta_i}{R} \sum_{i=1}^n \text{Cos} \theta_i + \sum_{i=1}^n \frac{\text{Sen} \theta_i}{R} \sum_{i=1}^n \text{Sen} \theta_i \right] \\ &= \frac{1}{n} \frac{R^2}{R} = \frac{R}{n} = A(\kappa^*) \end{aligned}$$

Entonces el estimador máximo verosímil de  $\kappa$  es la solución de  $A(\kappa^*) = \bar{R}$  o de manera equivalente  $\hat{\kappa} = A^{-1}(\bar{R})$ . La solución de la ecuación anterior se obtiene sólo mediante métodos numéricos. Una aproximación para estimar  $\hat{\kappa}$  la dieron Abramowitz y Stegun[4]:

$$\text{a) Si } 0 \leq R < 0.45 \quad \hat{\kappa} \doteq \frac{1}{6} \bar{R} (12 + 6\bar{R}^2 + 5\bar{R}^4)$$

$$\text{b) Si } 0.45 \leq \bar{R} < 0.8 \quad \hat{\kappa} \doteq 2\bar{R}$$

$$\text{c) Si } 0.8 \leq \bar{R} \leq 1 \quad \hat{\kappa} \doteq \frac{1}{2} (1 - \bar{R})$$

Sabiendo ya cuáles son los estimadores  $\hat{\kappa}$  y  $\hat{\mu}$ , se realiza el cociente de verosimilitudes:

$$\frac{L(\hat{\omega})}{L(\hat{\theta})} = \frac{(2\pi)^{-n}}{(2\pi I_0(\hat{\kappa}))^{-n} e^{[\hat{\kappa} \sum_{i=1}^n \text{Cos}(\theta_i - \hat{\mu})]}} = \frac{(2\pi)^{-n}}{(2\pi I_0(\hat{\kappa}))^{-n} e^{(\hat{\kappa} R)}} = \frac{(I_0(\hat{\kappa}))^n}{e^{(\hat{\kappa} R)}} = \lambda$$

$$\text{Como } A(\hat{\kappa}) = \frac{R}{n}$$

$$\implies \lambda = \frac{(I_0(\hat{\kappa}))^n}{e^{(n\hat{\kappa}A(\hat{\kappa}))}}$$

Se analizará ahora qué tipo de función es  $\lambda$ , para conocer así cuál será la región crítica de la prueba:

$$\ln \lambda = n \ln I_0(\hat{\kappa}) - n\hat{\kappa}A'(\hat{\kappa}) = n(\ln I_0(\hat{\kappa}) - \hat{\kappa}A'(\hat{\kappa}))$$

Se sabe además que:

$$\frac{\partial I_0(\kappa)}{\partial \kappa} = I_1(\kappa)$$

por lo cual

$$\frac{\partial \ln \lambda}{\partial \hat{\kappa}} = \frac{nI_1(\hat{\kappa})}{I_0(\hat{\kappa})} - n\hat{\kappa}A'(\hat{\kappa}) - nA'(\hat{\kappa}) = nA'(\hat{\kappa}) - n\hat{\kappa}A'(\hat{\kappa}) - nA'(\hat{\kappa}) = -n\hat{\kappa}A'(\hat{\kappa})$$

Por la definición de  $R$ , se sabe que  $\bar{R} \geq 0$ , por lo tanto  $\hat{\kappa} \geq 0$  pues  $A(\hat{\kappa}) = \bar{R}$ ; ahora se verá cómo es  $A'(\hat{\kappa})$ , para lo cual se construirá una ecuación que involucre esa expresión y así se pueda interpretar su comportamiento.

Sea  $\theta$  una variable aleatoria con f.d.p. von Mises  $(0, \kappa)$  :

$$f(\theta) = \frac{e^{(\kappa \cos \theta)}}{2\pi I_0(\kappa)} \quad 0 < \theta \leq 2\pi, \quad \kappa > 0$$

Sea  $\cos \theta$  una transformación de  $\theta$ , entonces:

$$\begin{aligned} E(\cos \theta) &= \int_0^{2\pi} \cos \theta \frac{e^{(\kappa \cos \theta)}}{2\pi I_0(\kappa)} d\theta = \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} \cos \theta e^{(\kappa \cos \theta)} d\theta \\ &= \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} \cos \theta \sum_{r=0}^{\infty} \frac{(\kappa \cos \theta)^r}{r!} d\theta = \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} \sum_{r=0}^{\infty} \frac{\kappa^r \cos^{r+1} \theta}{r!} d\theta \\ &= \frac{1}{2\pi I_0(\kappa)} \left[ \int_0^{2\pi} \cos \theta d\theta + \int_0^{2\pi} \kappa \cos^2 \theta d\theta + \int_0^{2\pi} \frac{\kappa^2 \cos^3 \theta}{2!} d\theta + \int_0^{2\pi} \frac{\kappa^3 \cos^4 \theta}{3!} d\theta + \dots \right] \\ &= \frac{1}{2\pi I_0(\kappa)} \left[ 0 + \pi \kappa + \frac{\kappa^3}{3!} \frac{3}{4} \pi + 0 + \frac{\kappa^5}{5!} \frac{5}{6} \frac{3}{4} \pi + \dots \right] = \frac{1}{I_0(\kappa)} \left[ \frac{\kappa}{2} + \frac{\kappa^3}{3!} \frac{3}{4} + \frac{\kappa^5}{5!} \frac{5}{6} \frac{3}{4} + \dots \right] \\ &= \frac{1}{I_0(\kappa)} \left[ \frac{\kappa}{2} + \left(\frac{\kappa}{2}\right)^3 \frac{1}{2!1!} + \left(\frac{\kappa}{2}\right)^5 \frac{1}{3!2!} + \dots \right] = \frac{I_1(\kappa)}{I_0(\kappa)} \end{aligned}$$

Por lo tanto:

$$E(\cos \theta) = \frac{I_1(\kappa)}{I_0(\kappa)}$$

y :

$$A'(\kappa) = \frac{I_0(\kappa) I_1'(\kappa) - I_1(\kappa) I_0'(\kappa)}{I_0^2(\kappa)} = \frac{I_1'(\kappa)}{I_0(\kappa)} - \frac{I_1^2(\kappa)}{I_0^2(\kappa)}$$

Se sabe que:

$$I_1(\kappa) = \frac{\kappa}{2} + \left(\frac{\kappa}{2}\right)^3 \frac{1}{2!1!} + \left(\frac{\kappa}{2}\right)^5 \frac{1}{3!2!} + \left(\frac{\kappa}{2}\right)^7 \frac{1}{4!3!} + \dots$$

por lo que:

$$I_1'(\kappa) = \frac{1}{2} + \left(\frac{\kappa}{2}\right)^2 \frac{3}{3!1!2} + \left(\frac{\kappa}{2}\right)^4 \frac{5}{3!2!2} + \left(\frac{\kappa}{2}\right)^6 \frac{7}{4!3!2} + \dots$$

escribiéndolo de una manera diferente:

$$\begin{aligned} &= \frac{1}{2\pi} \left[ \pi + \frac{\kappa^2}{2!} \frac{3}{4} \pi + \frac{\kappa^4}{4!} \frac{5}{6} \frac{3}{4} \pi + \frac{\kappa^6}{6!} \frac{7}{8} \frac{5}{6} \frac{3}{4} \pi + \dots \right] \\ &= \frac{1}{2\pi} \left[ \int_0^{2\pi} \cos^2 \theta \, d\theta + \frac{\kappa^2}{2!} \int_0^{2\pi} \cos^4 \theta \, d\theta + \frac{\kappa^4}{4!} \int_0^{2\pi} \cos^6 \theta \, d\theta + \frac{\kappa^6}{6!} \int_0^{2\pi} \cos^8 \theta \, d\theta + \dots \right] \\ &= \frac{1}{2\pi} \left[ \int_0^{2\pi} \cos^2 \theta \sum_{r=0}^{\infty} \frac{(\kappa \cos \theta)^r}{r!} d\theta \right] = \int_0^{2\pi} \frac{\cos^2 \theta e^{(\kappa \cos \theta)}}{2\pi} d\theta \end{aligned}$$

Por lo tanto:

$$\frac{I_1'(\kappa)}{I_0(\kappa)} = \int_0^{2\pi} \frac{\cos^2 \theta e^{(\kappa \cos \theta)}}{2\pi I_0(\kappa)} d\theta = E(\cos^2 \theta)$$

cuando  $\theta$  tiene una distribución von Mises(0,  $\kappa$ ), como se mencionó antes, y:

$$A'(\kappa) = \frac{I_1'(\kappa)}{I_0(\kappa)} - \frac{I_1^2(\kappa)}{I_0^2(\kappa)} = E(\cos^2 \theta) - E^2(\cos \theta) = Var(\cos \theta)$$

entonces  $A'(\kappa) \geq 0$ ,  $\lambda$  es una función monótona creciente de  $\overline{R}$ . Del cociente de verosimilitudes que tenemos llegamos a:

$$\ln I_0(\widehat{\kappa}) - \kappa A(\widehat{\kappa}) < \lambda' \implies -\widehat{\kappa}\overline{R} < \lambda'' \implies \overline{R} > \lambda'''$$

donde  $\lambda'''$  está dada por:

$$P[\overline{R} > \lambda''' / H_0] = \alpha$$

En efecto, la región crítica se esperaba de esta forma, pues si se rechaza  $H_0$ , es porque tenemos una  $R$  "grande", ya que bajo la hipótesis de uniformidad,  $\overline{R}$ , la longitud del vector medio, debe ser un valor pequeño, cercano a 0

### 2.1.7 DISTRIBUCIÓN ASINTÓTICA DE $2R/n^2$

A pesar de haber obtenido la distribución exacta de  $R$ , es evidente la dificultad que se presenta al tratar de evaluarla, por lo cual se hace necesario buscar una aproximación que sea más sencilla:

*"Cuando  $\kappa = 0$ ,  $X = \sum_{i=1}^n \text{Cos } \theta_i$ , y  $Y = \sum_{i=1}^n \text{Sen } \theta_i$  tienen distribuciones normales con media 0 y varianza  $n/2$ ".*

Para demostrar el resultado anterior se encontrará primero la media y la varianza de  $\text{Cos } \theta_i$ , partiendo de que  $\theta$  se distribuye como una von Mises( $0, \kappa$ ).

Hemos demostrado que para  $i = 1, 2, \dots, n$ :

$$\begin{aligned} E(\cos \theta_i) &= \frac{I_1(\kappa)}{I_0(\kappa)} \\ \text{Var}(\cos \theta_i) &= \frac{I_1'(\kappa)}{I_0(\kappa)} - \frac{I_1^2(\kappa)}{I_0^2(\kappa)} \end{aligned}$$

De la teoría de las funciones de Bessel se sabe que [2]:

$$I_1'(\kappa) = \frac{I_0(\kappa) + I_2(\kappa)}{2}$$

Por lo que:

$$Var(\cos \theta_i) = \frac{1}{2} \left[ 1 + \frac{I_2(\kappa)}{I_0(\kappa)} - \frac{2I_1^2(\kappa)}{I_0^2(\kappa)} \right]$$

Y como bajo la hipótesis de uniformidad  $\kappa = 0$ , para  $i = 1, 2, \dots, n$ :

$$\begin{aligned} E(\cos \theta_i) &= \frac{I_1(0)}{I_0(0)} = \frac{0}{1} = 0 \\ Var(\cos \theta_i) &= \frac{1}{2} \left[ 1 + \frac{0}{1} - \frac{2(0)^2}{1} \right] = \frac{1}{2} \end{aligned}$$

Aplicando el teorema del Límite Central:

$$\frac{\sum_{i=1}^n \cos \theta_i}{\left(\frac{n}{2}\right)^{\frac{1}{2}}} \sim N(0, 1)$$

Por lo que:

$$\sum_{i=1}^n \cos \theta_i \sim N\left(0, \frac{n}{2}\right)$$

Bajo las mismas hipótesis  $E(\sin \theta) = 0$ , pues el  $\sin \theta$  es una función simétrica con respecto a 0, por lo que  $Var(\sin \theta) = E(\sin^2 \theta)$  y:

$$\begin{aligned} E(\sin^2 \theta) &= \int_0^{2\pi} \sin^2 \theta \frac{e^{(\kappa \cos \theta)}}{2\pi I_0(\kappa)} d\theta = \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} \sin^2 \theta \sum_{r=0}^{\infty} \frac{(\kappa \cos \theta)^r}{r!} d\theta \\ &= \frac{1}{2\pi I_0(\kappa)} \left[ \int_0^{2\pi} \sin^2 \theta d\theta + \kappa \int_0^{2\pi} \sin^2 \theta \cos \theta d\theta + \frac{\kappa^2}{2!} \int_0^{2\pi} \sin^2 \theta \cos^2 \theta d\theta + \frac{\kappa^3}{3!} \int_0^{2\pi} \sin^2 \theta \cos^3 \theta d\theta + \dots \right] \\ &= \frac{1}{2\pi I_0(\kappa)} \left[ \pi + 0 + \frac{\kappa^2}{2!} \frac{\pi}{4} + 0 + \frac{\kappa^4}{4!} \frac{\pi}{4} \frac{1}{2} + 0 + \frac{\kappa^6}{6!} \frac{\pi}{4} \frac{1}{2} \frac{5}{8} + \dots \right] \end{aligned}$$

Y bajo la hipótesis nula,  $\kappa = 0$ .

$$E(\operatorname{sen}^2 \theta) = \frac{1}{2I_0(0)} = \frac{1}{2}$$

Entonces como ocurrió con  $X$ :

$$Y = \sum_{i=1}^n \operatorname{sen} \theta_i \sim N\left(0, \frac{n}{2}\right)$$

Y:

$$\frac{(\sum_{i=1}^n \operatorname{sen} \theta_i)^2 + (\sum_{i=1}^n \operatorname{cos} \theta_i)^2}{\frac{n}{2}} \sim \chi_{(2)}^2$$

esto es,

$$\frac{2R^2}{n} \sim \chi_{(2)}^2$$

siendo ésta la aproximación que en realidad se usa en la prueba y a partir de la cual se construyen las tablas.

## 2.1.8 INVARIANZA BAJO ROTACIONES

Como se habrá notado, el hecho de suponer que  $\theta$  tiene función de densidad von Mises con  $\mu = 0$  facilita en gran manera los cálculos y es que en realidad no es necesario mostrarlo cuando  $\mu \neq 0$ , ya que la prueba de Rayleigh tiene la ventaja de ser una prueba invariante bajo rotaciones, la idea de la demostración es la siguiente:

Con la hipótesis nula de nuestro interés:

$$H_0 : f(\theta) = \frac{1}{2\pi} \quad 0 < \theta \leq 2\pi,$$

Se define a la hipótesis alternativa  $H_a$  como una clase de funciones de densidad invariantes bajo cualquier cambio de dirección, de manera más clara, si  $(\xi_1, \xi_2, \dots, \xi_n)$  representa

una muestra aleatoria perteneciente a la familia de  $H_a$ , entonces también  $(\xi_1 e^{i\alpha}, \xi_2 e^{i\alpha}, \dots, \xi_n e^{i\alpha})$  pertenece a la clase de funciones de densidad de  $H_a$ , donde  $\alpha$  es cualquier constante real.

De acuerdo con Lehmann[3], con las características dadas, la prueba de  $H_0$  contra  $H_a$ , debe cumplir que para su función potencia  $\Psi$ :

$$\Psi(\xi_1, \xi_2, \dots, \xi_n) = \Psi(\xi_1 e^{i\alpha}, \xi_2 e^{i\alpha}, \dots, \xi_n e^{i\alpha})$$

Si se hace un cambio de dirección, tomando a  $\theta_n$  como nuestra nueva dirección cero, se define:

$$S_k = \frac{\xi_k}{\xi_n} \quad k = 1, 2, \dots, n-1$$

o bien.

$$S_k = e^{i\alpha_k} \quad \alpha_k = \theta_k - \theta_n$$

entonces  $\Psi$  sólo dependerá de  $n-1$  variables, pues  $S_n = 1$ , por lo que la función de densidad conjunta de  $\alpha_1, \alpha_2, \dots, \alpha_n$  será:

$$\begin{aligned} g(\alpha_1, \alpha_2, \dots, \alpha_{n-1}) &= \int_0^{2\pi} g(\theta_n) \prod_{k=1}^{n-1} g(\theta_k) d\theta_n \\ &= \int_0^{2\pi} g(\theta_n) \prod_{k=1}^{n-1} g(\theta_n + \alpha_k) d\theta_n \end{aligned}$$

Pero bajo la hipótesis de uniformidad  $g(\alpha_1, \alpha_2, \dots, \alpha_{n-1})$  es una constante, por lo que podemos encontrar una constante  $C$  tal que la función  $g(C_1, C_2, \dots, C_{n-1}) > C$ , concluyendo que ésta es la prueba uniformemente más potente para probar la hipótesis  $H_0$  contra  $H_a$ , de todas las pruebas invariantes con el mismo nivel de significancia.

Podemos también expresar la función  $g$ , en términos de las variables originales

$$\xi_k = e^{i\theta_k} .$$

$$g(\xi_1, \xi_2, \dots, \xi_n) = \int_0^{2\pi} \prod_{i=1}^n g(x + \theta_i) dx > C$$

y particularmente para una población con f.d.p. von Mises( $\mu_0, \kappa$ ) :

$$\begin{aligned} \prod_{i=1}^n g(x + \theta_i) &= \prod_{i=1}^n \frac{1}{2\pi I_0(\kappa)} e^{[\kappa \cos(x + \theta_i - \mu_0)]} \quad 0 \leq \theta_i, \mu_0 \leq 2\pi \quad \kappa > 0 \\ &= \frac{1}{[2\pi I_0(\kappa)]^n} e^{[\kappa \sum_{i=1}^n (\cos x - \mu_0 + \theta_i)]} \end{aligned}$$

Pero:

$$\begin{aligned} \sum_{i=1}^n \cos[(x - \mu_0) + \theta_i] &= \cos(x - \mu_0) \sum_{i=1}^n \cos \theta_i - \text{sen}(x - \mu_0) \sum_{i=1}^n \text{sen} \theta_i \\ &= \cos(x - \mu_0) R \cos \bar{x}_0 - \text{sen}(x - \mu_0) R \text{sen} \bar{x}_0 \\ &= R \cos[(x + \bar{x}_0) - \mu_0] \end{aligned}$$

De esta manera la función de densidad conjunta de  $\xi_1, \xi_2, \dots, \xi_n$  queda simplificada

así:

$$g(\xi_1, \xi_2, \dots, \xi_n) = \int_0^{2\pi} \frac{e^{[\kappa R \cos(x + \bar{x}_0 - \mu_0)]}}{[2\pi I_0(\kappa)]^n}$$

Pero esta expresión es muy parecida a la función de densidad de una variable von Mises( $x + \bar{x}_0, \kappa R$ ), evaluada sobre todo el recorrido, para verlo con más claridad se

multiplicará la expresión por  $\frac{I_0(\kappa R)}{I_0(\kappa R)}$  :

$$\frac{I_0(\kappa R)}{(2\pi)^{n-1} [I_0(\kappa)]^n} \int_0^{2\pi} \frac{e^{[\kappa R \cos(x + \bar{x}_0 - \mu_0)]}}{2\pi I_0(\kappa R)} = \frac{I_0(\kappa R)}{(2\pi)^{n-1} [I_0(\kappa)]^n} \implies \frac{I_0(\kappa R)}{[I_0(\kappa)]^n} > C$$

Queda así comprobada la invarianza de la prueba, siendo ésta una propiedad muy útil en la práctica

## 2.2 PRUEBA DE GREENWOOD Y DURAND

### 2.2.1 HISTORIA

En 1955 Greenwood y Durand introdujeron y desarrollaron esta prueba, que es en realidad un caso especial de la prueba de Rayleigh, se obtiene cuando en ésta, la hipótesis alternativa sí especifica, al menos teóricamente, la dirección esperada de las observaciones, mientras que la hipótesis nula permanece igual. Este cambio que se le hace a la prueba de Rayleigh tiene efectos prácticos muy útiles, ya que al realizar experimentos, en muchas ocasiones se tiene una idea de la dirección que tomarán las observaciones, pero se necesita evidencia estadística para confirmarla o de lo contrario para no rechazar la aleatoriedad. Si no se tomara en consideración la dirección teórica, se estaría desperdiciando una parte muy importante de la información.

Las aplicaciones de la prueba se centran principalmente en experimentos relacionados con la orientación de algunos animales bajo condiciones especiales.

### 2.2.2 PROPÓSITO

Probar si los valores angulares tienden a concentrarse cerca de  $\theta_0$  y así saber si la distribución de la que provienen difiere significativamente de la aleatoriedad.

Hay que aclarar en este punto que  $\theta_0$  debe conocerse antes de la realización del

experimento. Por otro lado, si se sabe que, por ejemplo, cierta especie animal sigue cierta orientación y se desea probar si la dirección media difiere significativamente de la dirección teórica, no se podrá aplicar la prueba de Greenwood y Durand, ésta solo debe ser usada para probar aleatoriedad.

### 2.2.3 DATOS REQUERIDOS

Una muestra aleatoria de  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$  y una dirección angular predeterminada,  $\theta_0$ .

### 2.2.4 IDEA BÁSICA DE LA PRUEBA

La estadística en la que se basa el criterio para definir la región crítica de la prueba será:

$$\begin{aligned}\bar{X} &= \frac{1}{n} \sum_{i=1}^n \cos \theta_i && \text{si } \theta_0 = 0 \text{ ó} \\ v &= r \cos(\bar{x}_0 - \theta_0) && \text{si } \theta_0 \neq 0\end{aligned}$$

donde  $r$  y  $\bar{x}_0$  son la longitud y dirección del vector medio. A estas estadísticas se les puede dar la misma interpretación:

Un simple análisis visual de las figuras 1 y 2 del capítulo 1, basta para saber que  $\bar{X}$  y  $v$  representan en ambos casos la proyección de  $m$ , el vector medio, sobre el vector con dirección  $\theta_0$ , por tal motivo pueden tomar valores de -1 a 1; más aún, dependen de la longitud del vector medio, que como se sabe varía en relación a la dispersión de

las observaciones alrededor de la dirección media, entonces, entre más parecidos sean los valores  $\theta_i$ , a  $\theta_0$ ,  $\bar{X}$  y  $v$  serán más cercanos a 1.

Lo anterior justifica, aunque de manera informal, la idea básica de la prueba y su relación con  $\bar{X}$  y  $v$ .

## 2.2.5 HIPÓTESIS A PROBAR

En el caso I:

$$\begin{array}{ll} H_0 : \kappa = 0 & \text{Bajo } H_a : \quad \kappa = \kappa^* > 0 \quad (\text{desconocido}) \\ \mu = 0 & \mu = 0 \end{array}$$

## 2.2.6 DISTRIBUCIÓN DE $\bar{x}$ Y $v$ .

Primero se mostrará que el cociente de verosimilitudes que establece la prueba, conduce a la necesidad de saber cómo se distribuyen estas estadísticas.

En el caso I:

$$\begin{array}{ll} \text{Bajo } H_0 : \kappa = 0 & \text{Bajo } H_a : \quad \kappa = \kappa^* > 0 \quad (\text{desconocido}) \\ \mu = 0 & \mu = 0 \end{array}$$

$$\Theta = \{(\mu, \kappa) / \mu = 0, 0 < \kappa < \infty\}$$

$$\omega = \{(\mu, \kappa) / \mu = \kappa = 0\}$$

Como  $\kappa$  es desconocido, se encontrará su Estimador Máximo Verosímil, para sustituirlo en el cociente de verosimilitudes:

$$\begin{aligned}
 L(\Theta) &= \prod_{i=1}^n \frac{e^{(\kappa \cos \theta_i)}}{2\pi I_0(\kappa)} \\
 \ln L(\Theta) &= \kappa \sum_{i=1}^n \cos \theta_i - n \ln 2\pi - n \ln I_0(\kappa) \\
 \frac{\partial \ln L(\Theta)}{\partial \kappa} &= \sum_{i=1}^n \cos \theta_i - n \frac{I_1(\kappa)}{I_0(\kappa)}
 \end{aligned}$$

Igualando a 0:

$$\sum_{i=1}^n \frac{\cos \theta_i}{n} = \frac{I_1(\kappa)}{I_0(\kappa)} = A(\kappa) \implies \bar{X} = A(\kappa)$$

Por lo que  $\hat{\kappa}_{MV}$  es la solución de la ecuación anterior, la cual se puede aproximar sólo por métodos numéricos.

El cociente de verosimilitudes queda determinado como:

$$\begin{aligned}
 \frac{L(\omega)}{L(\Theta)} &= \frac{(2\pi)^{-n}}{[2\pi I_0(\hat{\kappa})]^{-n} e^{(\hat{\kappa}\bar{X})}} \\
 &= \frac{I_0(\hat{\kappa})^n}{e^{(\hat{\kappa}\bar{X})}} < \lambda \implies n \ln I_0(\hat{\kappa}) - n\hat{\kappa}\bar{X} < \lambda' \implies \bar{X} > \lambda''
 \end{aligned}$$

Donde  $\lambda''$  se obtiene de  $P[\bar{X} > \lambda''/H_0] = \alpha$ .

Si se modifica la hipótesis alternativa:

$$\begin{aligned}
 H_a &: \kappa = \kappa^* > 0 \\
 &: \mu = \theta_0
 \end{aligned}$$

$\hat{\kappa}_{MV}$  será ahora la solución de la ecuación:

$$\sum_{i=1}^n \frac{\cos(\theta_i - \theta_0)}{n} = \frac{I_1(\kappa)}{I_0(\kappa)} = A(\kappa)$$

Al desarrollar el lado izquierdo de la ecuación:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \theta_0) &= \frac{1}{n} \sum_{i=1}^n \cos \theta_i \cos \theta_0 + \text{sen } \theta_i \text{ sen } \theta_0 \\ &= \frac{1}{n} \left( \cos \theta_0 \sum_{i=1}^n R \cos \bar{x}_0 + \text{sen } \theta_0 \sum_{i=1}^n R \text{sen } \bar{x}_0 \right) \\ &= \cos \theta_0 R \cos \bar{x}_0 + \text{sen } \theta_0 R \text{sen } \bar{x}_0 = R \cos(\bar{x}_0 - \theta_0) \end{aligned}$$

Entonces  $\hat{\kappa}_{MV}$  es la solución de :

$$R \cos(\bar{x}_0 - \theta_0) = A(\kappa)$$

El cociente de verosimilitudes es por tanto:

$$\begin{aligned} \frac{L(\hat{\omega})}{L(\hat{\Theta})} &= \frac{(2\pi)^{-n}}{[2\pi I_0(\hat{\kappa})]^{-n} e^{[\hat{\kappa} n^2 r \cos(\bar{x}_0 - \theta_0)]}} \\ \text{pues } r &= \frac{R}{n} \\ &= \frac{[I_0(\kappa)]^n}{e^{[\hat{\kappa} n^2 r \cos(\bar{x}_0 - \theta_0)]}} < \lambda \\ \implies n \ln I_0(\hat{\kappa}) - \hat{\kappa} n^2 r \cos(\bar{x}_0 - \theta_0) &< \lambda' \\ \implies v = r \cos(\bar{x}_0 - \theta_0) &> \lambda'' \end{aligned}$$

donde  $\lambda''$ , se obtiene de  $P[v > \lambda''] = \alpha$

En realidad la estadística de prueba que se usa es  $(2n)^{\frac{1}{2}} v$ , por su aproximación asintótica que se analizará después.

## 2.2.7 DISTRIBUCIÓN DE “X”

Al igual que la prueba de Rayleigh, la prueba “ $v$ ”, es invariante bajo rotaciones, pues es sólo un caso particular de aquélla. Gracias a este resultado que fue analizado en la sección anterior, solo se encontrará la distribución de :

$$X = \sum_{i=1}^n \cos \theta_i$$

estadística obtenida cuando en la hipótesis alternativa  $\mu = 0$ .

Para deducir la f.d.p. de  $X$ , se utilizará el Teorema de Inversión, el cual involucra a la función característica de  $X$  :

$$\phi_x(t) = \prod_{i=1}^n \phi_{\cos \theta_i}(t) = \phi_x^n(t) \quad \text{donde } x = \cos \theta$$

Partiendo de la función característica:

$$\phi_x(t) = E[e^{it \cos \theta}] = \int_0^{2\pi} e^{it \cos \theta} \frac{e^{[\kappa \cos(\theta - \mu_0)]}}{2\pi I_0(\kappa)} d\theta$$

Sean:

$$\mu = \cos \mu_0$$

$$v = \text{sen } \mu_0$$

Como:

$$it \cos \theta + \kappa \cos(\theta - \mu_0) = (it + \kappa \mu) \cos \theta + i(\kappa v/i) \text{sen } \theta$$

Si se define:

$$a = it + \kappa\mu$$

$$b = \kappa v/i$$

$$\phi_x(t) = \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} e^{(a \cos \theta + ib \sin \theta)} d\theta$$

Usando el desarrollo en serie de la función exponencial:

$$\begin{aligned} \int_0^{2\pi} e^{(a \cos \theta + ib \sin \theta)} d\theta &= \int_0^{2\pi} \sum_{n=0}^{\infty} \frac{(a \cos \theta + ib \sin \theta)^n}{n!} d\theta \\ &= \int_0^{2\pi} \left[ 1 + \frac{(a \cos \theta + ib \sin \theta)^2}{2!} + \frac{(a \cos \theta + ib \sin \theta)^4}{4!} + \frac{(a \cos \theta + ib \sin \theta)^6}{6!} + \dots \right] d\theta \\ &= 2\pi + \frac{\pi(a^2 - b^2)}{2!} + \frac{3\pi(a^2 - b^2)^2}{4 \cdot 4!} + \frac{5\pi(a^2 - b^2)^3}{8 \cdot 6!} + \dots \\ &= 2\pi \left[ 1 - \frac{(b^2 - a^2)}{2 \cdot 2!} + \frac{3(b^2 - a^2)^2}{8 \cdot 4!} - \frac{5(b^2 - a^2)^3}{16 \cdot 6!} + \dots \right] \\ &= 2\pi \left[ 1 - \frac{(b^2 - a^2)}{2^2} + \frac{(b^2 - a^2)^2}{2^2 \cdot 4^2} - \frac{(b^2 - a^2)^3}{2^2 \cdot 4^2 \cdot 6^2} + \dots \right] \\ &= 2\pi \left\{ J_0 \left[ (b^2 - a^2)^{\frac{1}{2}} \right] \right\} \end{aligned}$$

ya que[2] :

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \frac{x^6}{2^2 \cdot 4^2 \cdot 6^2} + \dots$$

Volviendo a la expresión de la función característica de  $\cos \theta$  :

$$\begin{aligned} \phi_x(t) &= \frac{2\pi \left[ (b^2 - a^2)^{\frac{1}{2}} \right]}{2\pi I_0(\kappa)} \\ &= \frac{J_0 \left\{ \left[ (t - i\kappa\mu)^2 - \kappa^2 v^2 \right]^{\frac{1}{2}} \right\}}{I_0(\kappa)} \end{aligned}$$

Aplicando el Teorema de Inversión:

$$\begin{aligned} g(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(-ixt)} \phi_x^n(t) dt \\ &= \frac{1}{2\pi J_0^n(\kappa)} \int_{-\infty}^{\infty} e^{(-ixt)} J_0^n \left\{ [(t - i\kappa\mu)^2 - \kappa^2\nu^2]^{\frac{1}{2}} \right\} dt \end{aligned}$$

Bajo la hipótesis  $\mu_0 = 0$ , se tiene  $\nu = 0$  y  $\mu = 1$  :

$$g(x) = \frac{1}{2\pi J_0^n(\kappa)} \int_{-\infty}^{\infty} e^{(-ixt)} J_0^n(t - i\kappa) dt$$

Bajo la hipótesis de uniformidad, es decir si  $\kappa = 0$  :

$$\begin{aligned} g(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(-ixt)} J_0^n(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\cos xt + i \operatorname{sen} xt) J_0^n(t) dt \\ &= \frac{1}{\pi} \int_0^{\infty} \cos Xt J_0^n(t) dt \end{aligned}$$

Esta es la función de densidad de  $X$ , sin embargo debido a su complejidad, se usa su aproximación asintótica, resultado que ya se analizó.

## 2.3 PRUEBA DEL RANGO

### 2.3.1 HISTORIA

Los primeros en formularla y estudiarla fueron Laubscher y Rudolph en 1968, posteriormente Rao se interesó. Fisher y H.A. David no trabajaron directamente sobre el tema, pero realizaron trabajos de cálculo que facilitaron la elaboración de las tablas

necesarias para la prueba. Como se verá más adelante, ésta es una de las pruebas más rápidas, por lo que los investigadores frecuentemente la prefieren sobre otras.

### 2.3.2 DATOS

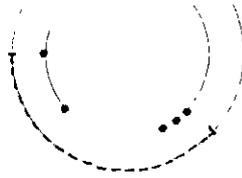
Se necesitan  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$  que representen direcciones en el plano o los tiempos de ocurrencia de algún fenómeno cíclico.

### 2.3.3 PROPÓSITO

Probar si la población de la cual es extraída la muestra difiere significativamente de la aleatoriedad, o bien probar si hay evidencia estadística de tendencia a cierta orientación.

### 2.3.4 IDEA BÁSICA DE LA PRUEBA

Se considerará una muestra de una población con distribución unimodal. Sea  $r$  la longitud del arco más pequeño que contenga todas las observaciones de la muestra. Si  $r$  es pequeño, físicamente se tendría una gráfica con las observaciones concentradas hacia una dirección y entonces se estaría en condiciones de rechazar la aleatoriedad.



r

### 2.3.5 OBSERVACIONES

Los datos deberán ordenarse y no deberán estar agrupados. En caso de tener una muestra bimodal, con las modas opuestas entre sí, es decir una muestra axial, podemos obtener el doble de los ángulos, reduciéndolos si fueran mayores que  $360^\circ$ .

### 2.3.6 HIPÓTESIS

$H_0$ : La población de la cual proviene la muestra se distribuye uniforme (hay aleatoriedad)

### 2.3.7 ESTADÍSTICA DE PRUEBA

$w$  : Rango Circular

El rango circular es la longitud del arco más pequeño que contiene todas las observaciones. En términos de las variables  $T_i$ 's definidas anteriormente:

$$w = 360^\circ - \max(T_1, T_2, \dots, T_n)$$

### 2.3.8 REGLA DE DECISIÓN

Si  $w < w(\alpha)$ , la hipótesis nula de aleatoriedad se rechaza, con lo que se concluye que existe algún tipo de orientación.

### 2.3.9 DISTRIBUCIÓN DE $w$

Para obtener la función de distribución de  $w$ , bajo la hipótesis de uniformidad, se verá bajo qué condiciones se cumple que  $w \leq r$ . Por la definición misma de  $w$ , lo anterior sucede cuando al menos una de las longitudes de los arcos  $T_i$ 's sobrepasa el valor  $r' = 2\pi - r$ . Sea  $E_i$  el evento  $T_i > r'$ , entonces por el Teorema de la probabilidad total:

$$\begin{aligned} F(r) &= P\left(\bigcup_{i=1}^n E_i\right) \\ &= n \Pr(E_1) - \binom{n}{2} \Pr(E_1 E_2) + \dots + (-1)^{s-1} \binom{n}{s} \Pr(E_1 \dots E_s) \end{aligned}$$

donde  $s$  es la parte de integración de  $2\pi/r'$ .

Se encontrará la probabilidad de  $E_1 E_2 \dots E_k$ . Como los  $E_i$ 's están en función de los  $T_i$ 's, se tratará primero de estudiar ésta.

Si una estadística  $\eta(\theta_1, \theta_2, \dots, \theta_n)$  es invariante bajo rotaciones se puede decir que la misma se basa en sólo  $n - 1$  variables, escogiendo por ejemplo a  $\theta_n$  como la dirección cero, se tiene que  $\eta(\theta_1, \theta_2, \dots, \theta_n) = \eta(\theta_1^*, \theta_2^*, \dots, 0)$ , y la función de verosimilitud de  $\theta_1^*, \theta_2^*, \dots, \theta_{n-1}^*$  sería:

$$\int_0^{2\pi} f(\theta) \prod_{i=1}^{n-1} f(\theta_i + \theta) d\theta \quad , \quad 0 < \theta_i^* \leq 2\pi \quad , \quad i = 1, 2, \dots, n - 1.$$

Bajo la hipótesis nula  $f(\theta)$  tiene una distribución uniforme sobre  $(0, 2\pi)$  y si además se considera que las observaciones pueden ordenarse de  $(n-1)!$  maneras distintas, la distribución conjunta de las estadísticas de orden  $\theta_{(1)}^*, \theta_{(2)}^*, \dots, \theta_{(n-1)}^*$  es

$$\frac{(n-1)!}{(2\pi)^{n-1}} \quad 0 < \theta_i^* \leq 2\pi$$

Entonces:

$$\begin{aligned} g(T_1, T_2, \dots, T_k) &= (n-1)! \frac{(n-1)!}{(2\pi)^{n-1}} \int_0^{2\pi-T_1} \dots \int_0^{2\pi-\dots-T_{n-2}} dT_{n-1} \dots dT_{k+1} \\ &= [(n-1)! / \{(n-k-1)!(2\pi)^{n-1}\}] (2\pi - T_1 - T_2 - \dots - T_k)^{n-k-1} \end{aligned}$$

Aplicando este resultado a:

$$P(E_1, E_2, \dots, E_k) = \int_w^{2\pi} \dots \int_w^{2\pi-T_1-\dots-T_{k-1}} g(T_1, T_2, \dots, T_k) dT_k \dots dT_1$$

se obtiene:

$$\Pr(E_1 \dots E_k) = \left\{ 1 - (kr'/2\pi) \right\}^{n-1}$$

$$\therefore F(r) = \sum_{k=1}^{\infty} (-1)^{k-1} \binom{n}{k} [1 - k \{(2\pi - r)/2\pi\}]^{n-1}$$

## 2.4 PRUEBA DE ESPACIOS DE RAO

### 2.4.1 HISTORIA

Esta prueba es relativamente reciente. Rao la formuló por primera vez y Puri y Yoon la modificaron posteriormente, haciéndola más eficiente. Las aplicaciones de esta prueba se han dado en el campo de la Biología, por Freid y Wagner, Ireland, entre otros.

## 2.4.2 DATOS

Se requiere de una muestra aleatoria de  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$ , que representan direcciones en el plano o tiempos de ocurrencia de algún fenómeno cíclico.

## 2.4.3 OBSERVACIONES

Además de que los datos deben ordenarse, no deben estar agrupados. En caso de tener una muestra bimodal, con las dos modas opuestas entre sí, es decir, una muestra axial, se puede obtener el doble de los ángulos y trabajar con ellos.

## 2.4.4 PROPÓSITO

Probar si la población de la cual proviene la muestra difiere significativamente de la aleatoriedad o si existe evidencia estadística de alguna tendencia direccional. dirección.

## 2.4.5 IDEA BÁSICA DE LA PRUEBA

Bajo la hipótesis de aleatoriedad, se esperaría que al graficar las observaciones en el círculo, éstas tenderían a estar igualmente espaciadas entre sí. La longitud del arco entre cualesquiera dos observaciones adyacentes sería  $\frac{360^\circ}{n}$ . Podrían permitirse diferencias pequeñas con respecto a esta cantidad, sin embargo diferencias grandes no indicarían una distribución uniforme de las observaciones de la muestra.



## 2.4.6 HIPÓTESIS NULA

$H_0$ : La población de la que se extrajo la muestra se distribuye uniforme (hay aleatoriedad).

## 2.4.7 ESTADÍSTICA PRUEBA

Haciendo uso de la definición de los  $T_i$ 's de la sección anterior, se puede saber que :

$$\sum_{i=1}^n T_i = 360^\circ$$

$\left| T_i - \frac{360^\circ}{n} \right| \quad i = 1, 2, \dots, n$ ; son las desviaciones entre la longitud del arco esperada y las longitudes de arco muestrales. Se puede pensar entonces que:

$$R = \sum_{i=1}^n \left| T_i - \frac{360^\circ}{n} \right|$$

funcione como una estadística de prueba.

## 2.4.8 REGLA DE DECISIÓN

Se rechaza  $H_0$ , la hipótesis de aleatoriedad, al nivel de significancia  $\alpha$ , si  $R > R(\alpha)$

(valor de tablas).

## 2.4.9 DISTRIBUCIÓN DE LA ESTADÍSTICA

Para encontrar la distribución asintótica de  $R = \sum_{i=1}^n |T_i - \frac{360^\circ}{n}|$ , se usará un resultado de L. Sherman[6], quien demostró que:

$$E(R^k) = \binom{n+k}{k}^{-1} \sum_{s=0}^{\mu-1} \binom{n+1}{s+1} \binom{k-1}{s} \left(\frac{n-s}{n+1}\right)^{n+k}$$

donde  $\mu = \min(k, n)$ .

Entonces si  $k=1$  :

$$E(R) = \binom{n+1}{1}^{-1} \binom{n+1}{1} \binom{0}{0} \left(\frac{n}{n+1}\right)^{n+1} = \left(\frac{n}{n+1}\right)^{n+1}$$

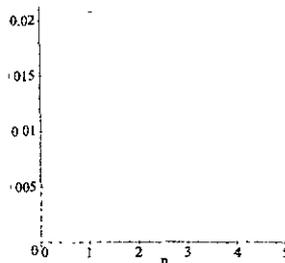
Si  $k=2$ :

$$\begin{aligned} E(R^2) &= \binom{n+2}{2}^{-1} \sum_{s=0}^1 \binom{n+1}{s+1} \binom{1}{s} \left(\frac{n-s}{n+1}\right)^{n+2} \\ &= \binom{n+2}{2}^{-1} \left[ \binom{n+1}{1} \binom{1}{0} \left(\frac{n}{n+1}\right)^{n+2} + \binom{n+1}{2} \binom{1}{1} \left(\frac{n-1}{n+1}\right)^{n+2} \right] \\ &= \frac{2}{(n+1)(n+2)} \left[ \frac{n^{n+2}}{(n+1)^{n+1}} + \frac{n(n-1)^{n+2}}{2(n+1)^{n+1}} \right] = \frac{2n^{n+2} + n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}} \end{aligned}$$

Podemos observar que  $E(R) \rightarrow \frac{1}{e}$  si  $n \rightarrow \infty$ , además:

$$Var(R) = \frac{2n^{n+2} + n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}} - \left(\frac{n}{n+1}\right)^{2n+2},$$

para distintos valores de  $n$  se comporta como:



por lo que  $Var(R) \simeq \frac{c}{n}$ , con  $c$  adecuada.

Se define la variable estandarizada:

$$R^* = \frac{R - E(R)}{(VarR)^{\frac{1}{2}}} = \left(\frac{n}{c}\right)^{\frac{1}{2}} \left(R - \frac{1}{c}\right)$$

Si los momentos de  $R^*$  se aproximan a los momentos de una variable aleatoria con función de distribución  $F(x)$ , entonces la función de distribución de  $R^*$  se aproximará a  $F(x)$ .

Sherman demuestra que :

$$E' = \lim_{n \rightarrow \infty} E \left[ \left(\frac{n}{c}\right)^m \left(L - \frac{1}{e}\right)^{2m} \right] = \frac{a_m(2m)!}{(2e-5)^m}$$

con  $a_m = \frac{1}{m!2^m} - (2e-5)^m$ , por lo que  $E'$  representa los momentos pares de una distribución normal, entonces  $R_{n \rightarrow \infty}^* \simeq \text{Normal}$ .

## 2.5 PRUEBA DE HODGES Y AJNE

### 2.5.1 HISTORIA

En 1968 Ajne estudió la prueba para detectar uniformidad en una distribución circular, sin embargo Hodges anteriormente (1955), había desarrollado una prueba acerca de la ubicación de una distribución bivariada. Aunque parezca no haber relación entre ellas, en realidad son idénticas, es por eso que la prueba se atribuye a ambos.

### 2.5.2 PROPÓSITO

Probar si existe evidencia estadística de orientación o de manera equivalente, si la muestra obtenida dista de seguir un comportamiento aleatorio.

### 2.5.3 DATOS REQUERIDOS

Una muestra aleatoria de  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$  que representan direcciones en el plano o los tiempos de ocurrencia de un fenómeno cíclico.

Los datos no deben estar agrupados, sin embargo se puede debilitar esta condición pidiendo que, en caso de tenerlos, su número debe ser grande comparado con el tamaño de la muestra y en este caso se tendrá una modificación en la estadística de prueba.

En caso de tener datos axiales, se debe obtener en primer lugar el doble de los ángulos dados.

### 2.5.4 IDEA BÁSICA DE LA PRUEBA

Se considera un conjunto finito de puntos, localizados sobre la circunferencia unitaria, se quiere probar la hipótesis de que la distribución subyacente de los puntos es uniforme; para tal propósito se dibuja un diámetro  $l$ , éste divide a la circunferencia en dos semicírculos, ahora se rota  $l$  sobre el centro del círculo hasta tener el número máximo de puntos de la muestra de un lado de  $l$ , y el número mínimo del otro lado. Si el número mínimo de puntos que quedan de un lado de la circunferencia es “pequeño”, es lógica la sospecha de que no haya aleatoriedad en las observaciones de la muestra dada. Esto conduce a pensar de manera natural que el número mínimo de puntos de la muestra que quedan de un lado del diámetro  $l$ , llámese  $m$ , puede funcionar como una estadística de prueba.

## 2.5.5 HIPÓTESIS A PROBAR

Las hipótesis nula y alternativa, son:

$H_0$  : La población se distribuye uniformemente (hay aleatoriedad).

$H_a$  : La población se distribuye bajo una dirección definida (puede proponerse que la población se distribuye como una von Mises, por ejemplo)

## 2.5.6 REGLA DE DECISIÓN

Se rechaza  $H_0$  al nivel de significancia  $\alpha$  si  $P < \alpha$

donde  $P$  que está en función de  $n$  y  $m$ , es un valor encontrado en las tablas.

## 2.5.7 INVARIANZA BAJO ROTACIONES

Para comprender esta propiedad se propondrán las siguientes hipótesis:

$H_0$  : La muestra proviene de una población que se distribuye uniforme.

$H_a$  : La muestra proviene de una clase específica de distribuciones continuas, la cual es invariante bajo el grupo de todas las rotaciones.

Con el propósito de aclarar el significado de  $H_a$ , se puede decir que si  $(\xi_1, \xi_2, \dots, \xi_n)$  tiene una distribución perteneciente a  $H_a$ , entonces para toda constante real  $\beta$ ,  $(\xi_1 e^{i\beta}, \xi_2 e^{i\beta}, \dots, \xi_n e^{i\beta})$  también tendrá una distribución perteneciente a  $H_a$ .

Bajo este supuesto, la prueba de  $H_0$  vs.  $H_a$  es invariante, en otras palabras, la función potencia  $\phi$ , cumple que :

$$\phi(\xi_1, \xi_2, \dots, \xi_n) = \phi(\xi_1 e^{i\beta}, \xi_2 e^{i\beta}, \dots, \xi_n e^{i\beta})$$

esto es, que la probabilidad con la cual  $H_0$  es rechazada, dada la muestra  $(\xi_1, \xi_2, \dots, \xi_n)$  es igual a la probabilidad con la cual  $H_0$  es rechazada, dada la muestra transformada:

$$(\xi_1 e^{i\alpha}, \xi_2 e^{i\alpha}, \dots, \xi_n e^{i\alpha})$$

Sabemos que  $\xi_k = c^{i\theta_k}$   $0 < \theta \leq 2\pi$ , si se define

$$\begin{aligned} S_k &= \frac{\xi_k}{n} & k = 1, 2, \dots, (n-1) \\ &= e^{i\theta_k - i\theta_n} = e^{i\alpha_k} & \alpha_k = \theta_k - \theta_n \end{aligned}$$

$$0 \leq \alpha_k \leq 2\pi \quad \text{para } k = 1, 2, \dots, n$$

Los  $S'_k$ s describen qué tan parecidos son los  $\xi'_k$ s de  $\xi_n$ , ya que esta transformación sólo rotó la muestra  $\theta_n$  grados.

$\phi$  dependerá de  $n-1$  variables:  $S_1, S_2, \dots, S_{n-1}$  pues  $S_n = 1$ . Si se supone que  $\theta_1, \theta_2, \dots, \theta_n$  es una muestra aleatoria y utilizando el Teorema de cambio de variable, la función de densidad conjunta de  $\alpha_1, \alpha_2, \dots, \alpha_n$  es:

$$g(\alpha_1, \alpha_2, \dots, \alpha_n/f) = \int_0^{2\pi} f(\theta_n) \prod_{i=1}^n f(\alpha_k + \theta_n) d\theta_n$$

Si  $H_0$  es cierta,  $f$  y  $g$  son constantes, además por el lema de Neyman-Pearson, si una función invariante tiene función potencia  $\phi$  de la forma:

$$\phi(S_1, S_2, \dots, S_n) = \begin{cases} 1 & \text{si } g(\alpha_1, \alpha_2, \dots, \alpha_{n-1}/f) > C \\ 0 & \text{si } g(\alpha_1, \alpha_2, \dots, \alpha_{n-1}/f) < C \end{cases}$$

para alguna  $\alpha$ , entonces ésta es una prueba uniformemente más potente para probar  $H_0$  contra la alternativa correspondiente a la densidad  $f$ , entre todas las pruebas invariantes de igual o menor nivel de significancia.

## Capítulo 3

# BONDAD DE AJUSTE, UN ENFOQUE GENERAL

### 3.1 PRUEBA DE LA JI CUADRADA

#### 3.1.1 HISTORIA

Fue Karl Pearson, quien en el desarrollo de sus “Contribuciones Matemáticas a la Teoría de la Evolución”, se vió en la necesidad de probar “ajuste”, por lo que en 1900 desarrolló su prueba de la Ji Cuadrada, una de las pruebas más usadas debido a su simplicidad y a la gran diversidad de situaciones en la que puede ser usada, ya que sólo se requiere que las  $n$  observaciones puedan ser clasificadas en un número de  $k$  categorías ajenas y que la probabilidad de que las observaciones caigan en esas categorías pueda ser calculada cuando la hipótesis nula se supone cierta. Pero fue hasta 1972 cuando Rao, J.S. la aplicó a la estadística circular.

La prueba de la Ji Cuadrada en el contexto circular fue usada para el estudio de la orientación de ciertas especies de animales por Arendse, Tesch y Lelek, Waterman y Forward.

### 3.1.2 DATOS

Una muestra aleatoria de  $n$  valores angulares  $\theta_1, \theta_2, \dots, \theta_n$  de una distribución circular. Por ser nuestro propósito probar la uniformidad o aleatoriedad de las observaciones se propondrá una distribución uniforme, sin embargo para otros fines se podría proponer otra distribución, como una von Mises con cierta  $\kappa > 0$ .

### 3.1.3 PROPÓSITO

Probar si la distribución propuesta se ajusta a la muestra dada.

### 3.1.4 IDEA BÁSICA DE LA PRUEBA

Para poder aplicar la prueba de la Ji Cuadrada se debe dividir el círculo que contiene las observaciones, en arcos. En cada uno se contará la frecuencia de las observaciones de la muestra y además se calculará la frecuencia esperada de la distribución circular dada. El ajuste de la distribución se considera satisfactorio si las frecuencias esperadas no difieren mucho de las frecuencias observadas. Se puede aplicar como en el caso lineal para el circular sin hacer ajuste alguno, sin embargo Rao la mejoró: como el resultado de la prueba depende en cierta manera de la posición y longitud de los arcos que sirven como “intervalos”, Rao, J.S. sugirió una aproximación de la Ji Cuadrada

con el fin de eliminar el punto inicial, pero aún sin hacer esta consideración la prueba resulta ser buena.

### 3.1.5 OBSERVACIONES

Los datos deben ser agrupados de manera conveniente, la longitud de los arcos no debe ser necesariamente igual, pero la frecuencia esperada en cada intervalo debe ser de al menos cuatro.

### 3.1.6 ESTADÍSTICA DE PRUEBA

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

Donde:

$k$  = número de grupos en que se divide el círculo

$n_i$  = frecuencia observada del  $i$ -ésimo grupo en  $n$  experimentos

$e_i$  = frecuencia esperada del  $i$ -ésimo grupo en  $n$  experimentos

### 3.1.7 DISTRIBUCIÓN ASINTÓTICA

En 1949 Neyman demostró que la estadística de prueba obtenida por el procedimiento de razón de verosimilitudes puede escribirse de esta manera, la estadística tiene una distribución asintótica Ji Cuadrada con  $k - 1$  grados de libertad:

En nuestro contexto tenemos  $k$  arcos de la circunferencia, no necesariamente de longitud igual, en los cuales deben estar las  $n$  observaciones angulares de nuestro

experimento. Sea  $p_i, i = 1, 2, \dots, k$  la probabilidad de que el resultado esté en el  $i$ -ésimo arco, el problema consiste en determinar si un conjunto de datos experimentales se ajusta con lo "esperado", esto es, basándose en los valores de las  $p_i$ 's propuestas.

De manera formal :  $H_0 : p_i = \rho_i, i = 1, 2, \dots, k$ . Para plantear la razón de verosimilitudes se debe considerar que las observaciones que caen en el  $i$ -ésimo arco se distribuyen como una v.a. Binomial  $(n, \rho_i)$ , además la función de densidad conjunta de la muestra, para una v.a. discreta es la probabilidad de obtener los valores muestrales observados en el orden en el que fueron obtenidos , por lo que :

$$L(\Theta) = p_1^{n_1} p_2^{n_2} p_3^{n_3} \dots p_k^{n_k}$$

Se sabe que  $\sum_{i=1}^k p_i = 1$ , este resultado conduce a pensar que  $k - 1$  de las  $p_i$ 's son parámetros independientes por lo que para encontrar los estimadores de máxima verosimilitud de las  $p_i$ 's, se sustituirá  $p_k$  por  $1 - \sum_{i=1}^{k-1} p_i$ .

De esta manera :

$$\begin{aligned} L(\Theta) &= p_1^{n_1} p_2^{n_2} p_3^{n_3} \dots p_{k-1}^{n_{k-1}} \left( 1 - \sum_{i=1}^{k-1} p_i \right) \\ \ln L(\Theta) &= n_1 \ln p_1 + n_2 \ln p_2 + \dots + n_{k-1} \ln p_{k-1} + n \ln \left( 1 - \sum_{i=1}^{k-1} p_i \right) \\ &= \sum_{i=1}^{k-1} n_i \ln p_i + n_k \ln \left( 1 - \sum_{i=1}^{k-1} p_i \right) \\ \frac{\partial \ln L(\Theta)}{\partial p_i} &= \frac{n_i}{p_i} - \frac{n_k}{\left( 1 - \sum_{i=1}^{k-1} p_i \right)} = 0 \\ \implies \frac{n_i}{p_i} &= \frac{n_k}{p_k} \quad i = 1, 2, \dots, k \implies n_i p_k = n_k p_i \quad i = 1, 2, \dots, k \end{aligned}$$

Sumando los  $k$  términos:

$$p_k \sum_{i=1}^k n_i = n_k \sum_{i=1}^k p_i,$$

donde:

$$\begin{aligned} \sum_{i=1}^k n_i &= n \\ \sum_{i=1}^k p_i &= 1 \\ \implies p_k n &= n_k \implies p_k = \frac{n_k}{n} \implies \hat{p}_k = \frac{N_k}{n} \end{aligned}$$

Generalizando con los parámetros restantes:

$$\hat{p}_i = \frac{N_i}{n} \quad i = 1, 2, \dots, k$$

Bajo la hipótesis nula el cociente de verisimilitudes :

$$\begin{aligned} \lambda &= \frac{L(\omega)}{L(\Theta)} = \frac{\rho_1^{n_1} \rho_2^{n_2} \dots \rho_k^{n_k}}{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2} \dots \left(\frac{n_k}{n}\right)^{n_k}} \\ &= \left(\frac{n\rho_1}{n_1}\right)^{n_1} \left(\frac{n\rho_2}{n_2}\right)^{n_2} \dots \left(\frac{n\rho_k}{n_k}\right)^{n_k} \end{aligned}$$

y como ya se había considerado, cada  $N_i$  tiene distribución Bin  $(n, \rho_i)$ . El tratar de obtener la función de distribución exacta de  $\lambda$  no parece ser sencillo, pero se puede usar su distribución asintótica, pues se sabe que  $-2 \ln \lambda$  se distribuye asintóticamente como una Ji cuadrada en la que los grados de libertad se obtienen calculando la diferencia del número de parámetros independientes desconocidos bajo  $\Theta$  y  $\omega$ , que en este caso son  $k - 1$ . La región crítica queda determinada de la siguiente manera:

$C = \{x/ -2 \ln \lambda \geq c\}$  pues  $\lambda \leq \lambda_0 \Leftrightarrow -2 \ln \lambda \geq -2 \ln \lambda_0 = c$  y donde  $P\{-2 \ln \lambda \geq c/H_0\} =$

$\alpha$

Se demostrará ahora que  $-2 \ln \lambda$  puede escribirse como :

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

Para hacerlo, se utilizará la expansión en serie de Taylor de  $\ln p_i$  alrededor de

$$\widehat{p}_i = \frac{n_i}{n} :$$

$$\begin{aligned} \ln p_i &= \ln \widehat{p}_i + (p_i - \widehat{p}_i) \frac{1}{\widehat{p}_i} + (p_i - \widehat{p}_i)^2 \frac{1}{2!} \left( \frac{-1}{\widehat{p}_i^2} \right) + \varepsilon \\ &= \ln \frac{n_i}{n} + \left( p_i - \frac{n_i}{n} \right) \frac{n}{n_i} - \left( p_i - \frac{n_i}{n} \right)^2 \left( \frac{n}{n_i} \right)^2 + \varepsilon \\ \Rightarrow \ln p_i - \ln \frac{n_i}{n} &= \left( p_i - \frac{n_i}{n} \right) \frac{n}{n_i} - \left( p_i - \frac{n_i}{n} \right)^2 \frac{n}{2n_i^2} + \varepsilon \end{aligned}$$

Donde  $\varepsilon$  es la suma de los términos con signos alternantes:

$$\sum_{j=3}^{\infty} (-1)^{j+1} \left( p_i - \frac{n_i}{n} \right)^j \frac{n^j}{j! n_i^j}$$

Bajo  $H_0$ ,  $p_i = \rho_i$  y sustituyendo en la ecuación:

$$\begin{aligned} -2 \ln \lambda &= -2 \sum_{i=1}^k n_i \ln \left( \frac{n_i \rho_i}{n_i} \right) = -2 \sum_{i=1}^k n_i \ln \left( \frac{\rho_i}{n} \right) \\ &= -2 \sum_{i=1}^k n_i \left( \ln \rho_i - \ln \frac{n_i}{n} \right) \\ &= -2 \sum_{i=1}^k n_i \left[ \left( \rho_i - \frac{n_i}{n} \right) \frac{n}{n_i} - \left( \rho_i - \frac{n_i}{n} \right)^2 \frac{n^2}{2n_i^2} + \varepsilon \right] \\ &= -2 \sum_{i=1}^k n_i (n \rho_i - n_i) + \sum_{i=1}^k \frac{(n \rho_i - n_i)^2}{n_i} + \sum_{i=1}^k \varepsilon' \end{aligned}$$

Pero:

$$\begin{aligned} \sum_{i=1}^k (n\rho_i - n_i) &= n \sum_{i=1}^k \rho_i - \sum_{i=1}^k n_i = n - n = 0 \\ \Rightarrow -2 \ln \lambda &= \sum_{i=1}^k \frac{(n\rho_i - n_i)^2}{n_i} + \varepsilon'' \end{aligned}$$

Donde  $\varepsilon'' \rightarrow 0$  si  $n \rightarrow \infty$  con probabilidad 1.

Como  $-2 \ln \lambda$  se aproxima a la estadística :

$$Q = \sum_{i=1}^k \frac{(n\rho_i - n_i)^2}{n_i}$$

la f.d.p. de  $Q$  converge a la de  $-2 \ln \lambda$ , que es una Ji Cuadrada con  $k - 1$  grados de libertad.

### 3.1.8 DEFINICIÓN DE LA PRUEBA

$$H_0 : F = F_0 \quad H_a : F \neq F_0$$

Para nuestro interés  $F_0 = \frac{\theta}{2\pi} \quad 0 < \theta \leq 2\pi$

La hipótesis nula se rechaza si el valor crítico  $P$ , obtenido de tablas y con base en la estadística de prueba  $\chi^2$  es menor al nivel de significancia  $\alpha$  preasignado.

Los métodos gráficos juegan un papel importante al tratar de resolver el problema de decidir si una v.a. tiene una forma específica de distribución. Parte de estas pruebas se basan en la función de distribución empírica o muestral. *FDE*.

Después de haber tomado una dirección fija inicial, o dirección cero, se define la función de distribución empírica  $S_n(\theta)$  de la misma manera que se hace en el caso lineal. Se define  $\theta_0 = 0, \theta_{n+1} = 2\pi$

$$S_n(\theta) = \frac{k}{n}, \quad \text{si } \theta_i \leq \theta < \theta_{i+1} \quad i = 0, 1, \dots, n$$

Cualquier estadística que mida la diferencia entre  $S_n(\theta)$  y  $F(\theta)$ , función de distribución de la que se sospecha se ajusta a la muestra aleatoria que se tiene, será llamada una estadística *FDE*. Las más importantes son siete [7], éstas se basan en las diferencias verticales entre  $S_n(\theta)$  y  $F(\theta)$ . para su estudio se han dividido en dos clases, la de supremos y las cuadráticas.

Dentro de la primera clase están las estadísticas  $D^+$  y  $D^-$ , que son respectivamente las distancias mayores medidas verticalmente, cuando  $S_n(\theta)$  es mayor que  $F(\theta)$  y cuando  $S_n(\theta)$  es menor que  $F(\theta)$ , la muy bien conocida estadística  $D = \max(D^+, D^-)$  de Kolmogorov y la estadística  $V_n = D^+ + D^-$  de Kuiper, que por ciertas propiedades que tiene y se analizarán con detalle, es útil en el campo de la estadística circular.

Dentro de la clase de estadísticas cuadráticas se encuentra la gran familia de estadísticas *FDE* dada por Crámer-Von Mises:

$$Q = n \int_{-\infty}^{\infty} \{S_n(\theta) - F(\theta)\}^2 \psi(\theta) dF(\theta)$$

Cuando  $\psi(\theta) = \{[F(\theta)][1 - F(\theta)]\}^{-1}$ , la estadística resultante se conoce con el nombre de Anderson-Darling; por último, la estadística  $U^2$  de Watson definida como

$$U^2 = n \int_{-\infty}^{\infty} \left\{ S_n(\theta) - F(\theta) - \int_{-\infty}^{\infty} [S_n(\theta) - F(\theta)] dF(\theta) \right\}^2 dF(\theta)$$

y que como se verá posteriormente, también es de gran utilidad en estadística circular.

### 3.2.3 PROPÓSITO

Probar si la distribución dada se ajusta a la muestra.

### 3.2.4 IDEA BÁSICA DE LA PRUEBA

Para entender la idea intuitiva de la prueba podemos dibujar la función de distribución teórica, que debe ser una curva continua, y la función de distribución empírica, que será una función escalonada. Si se analiza qué tanto las dos funciones difieren una de la otra, se podrán distinguir dos tipos de diferencias: cuando la gráfica de la función empírica rebasa a la de la función teórica y viceversa.

Llámesese  $D^+$  la distancia mayor, medida verticalmente, que haya de la gráfica de la función teórica a la gráfica de la función empírica. De manera análoga llámesese  $D^-$  a la distancia máxima medida verticalmente de la gráfica de la función empírica hacia la gráfica de la función teórica.

Si se define :  $V_n = D^+ + D^-$ , resulta trivial pensar que un valor pequeño de  $V_n$  indicará un buen ajuste de la función de distribución muestral a la función de distribución teórica.

### 3.2.5 OBSERVACIONES

Las observaciones no se deberán agrupar, pero cuando sea necesario se podrán construir intervalos de clase que no excedan de 5°.

### 3.2.6 ESTADÍSTICA DE PRUEBA

$$K = n^{\frac{1}{2}}V_n$$

Transformación de  $V_n$  cuya distribución asintótica se puede encontrar.

### 3.2.7 DEFINICIÓN DE LA PRUEBA

Al igual que en el caso de la prueba de la Ji Cuadrada las hipótesis se pueden plantear, por el interés de probar la aleatoriedad de las observaciones como:

$$H_0 : F_\theta = \frac{\theta}{2\pi} \quad 0 < \theta \leq 2\pi \quad \text{vs} \quad H_a : F \neq F_0$$

La hipótesis nula se rechaza a un nivel de significancia  $\alpha$  si la estadística de prueba es mayor que el valor crítico correspondiente en tablas.

### 3.2.8 INVARIANZA BAJO ROTACIONES

Como se señaló al principio de esta sección, existen más estadísticas que permiten saber la diferencia entre la función de distribución empírica y la teórica; en este momento es conveniente aclarar porqué es mejor usar  $V_n$  y no otra como  $D_n$ , la famosa estadística de Kolmogorov que como se sabe es muy usada en este tipo de pruebas

Considérese un ejemplo sencillo, una muestra de cuatro valores tomando al Norte como origen (Fig.1) y posteriormente esa misma muestra pero tomando al Este como origen (Fig. 2). Se desea comprobar la invarianza de la prueba.

$$\theta_1 = 108^\circ \quad \theta'_1 = 18^\circ$$

$$\theta_2 = 144^\circ \quad \theta'_2 = 54^\circ$$

$$\theta_3 = 180^\circ \quad \theta'_3 = 90^\circ$$

$$\theta_4 = 324^\circ \quad \theta'_4 = 234^\circ$$

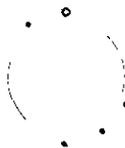


Fig. 1



Fig. 2

Haciendo la transformación  $U_i(\theta_i) = \theta_i/360$  :

$$u_1 = 0.3 \quad u'_1 = 0.05$$

$$u_2 = 0.4 \quad u'_2 = 0.15$$

$$u_3 = 0.5 \quad u'_3 = 0.25$$

$$u_4 = 0.9 \quad u'_4 = 0.65$$

Dibujando la *FDE* en ambos casos se obtienen las siguientes gráficas:

Se puede constatar así que:

$$D^+ = 0.25 \quad D'^+ = 0.50$$

$$D^- = 0.30 \quad D'^- = 0.05$$

$$D = 0.30 \quad D' = 0.50$$

$$V_n = 0.55 \quad V'_n = 0.55$$

Escójase una nueva dirección  $\theta = c$ , entonces  $2\pi U'_j$ ,  $j = 1, 2, \dots, n$ , serán los nuevos ángulos también en orden creciente. Para una  $c$  fija y una muestra dada existe un entero  $k$  tal que  $U_k < c < U_{k+1}$ . Entonces

$$\begin{aligned} U'_j &= U_{k+j} - c \quad \text{para } j = 1, \dots, n-k \\ &= U_{k+j-n} + 1 - c \quad \text{para } j = n-k+1, \dots, n \end{aligned}$$

De esta manera para  $i = k+j$ :

$$\frac{j}{n} - U'_j = \left[ \frac{(i-k)}{n} \right] - (U_i - c)$$

y para  $i = k+j-n$

$$\frac{j}{n} - U'_j = \left[ \frac{(i-k+n)}{n} \right] - (U_i + 1 - c)$$

De lo anterior se establece que:

$$\frac{j}{n} - U'_j = \frac{i}{n} - U_i + c - \frac{k}{n} \quad , \quad i = (k+j) \bmod n \quad j = 1, \dots, n$$

En realidad  $k$  es una variable aleatoria, pero dada una muestra y una  $c$ , es una constante. Para concluir se debe recordar un resultado muy sencillo.

Si se tiene  $\theta_0, \theta_1, \dots, \theta_n$ ,  $n$  números cualesquiera y una constante  $k$ , entonces :

$$\max_{0 \leq i \leq n} (\theta_i + k) = k + \max_{0 \leq i \leq n} \theta_i$$

De las igualdades que se habían obtenido:

$$\begin{aligned} V_n &= \max_{0 \leq j \leq n} \left[ \frac{j}{n} - U'_j \right] + \max_{0 \leq j \leq n} \left[ U'_j - \frac{(j-1)}{n} \right] \\ &= \max_{0 \leq i \leq n} \left[ \frac{i}{n} - U_i + c - \frac{k}{n} \right] + \max_{0 \leq i \leq n} \left[ U_i - c - \frac{(i-1-k)}{n} \right] \end{aligned}$$

### 3.3.4 IDEA BÁSICA DE LA PRUEBA

La estadística de prueba  $U_n^2$  es una medida de desviación cuadrática, por lo que si se observa un valor "grande" se podrá concluir que la distribución no se ajusta a la muestra.

### 3.3.5 OBSERVACIONES

Las observaciones no se deberán agrupar y tendrán que ordenarse de manera ascendente:  $\theta_1 \leq \theta_2 \leq \dots \leq \theta_n$ .

### 3.3.6 ESTADÍSTICA DE PRUEBA

$$U_n^2 = \sum_{i=1}^n \left( Y_i - \frac{2i-1}{2n} \right)^2 + \frac{1}{12n} - n \left( \frac{1}{2} - \bar{Y} \right)^2$$

que es la forma simplificada que se obtiene al realizar la desviación cuadrática mencionada anteriormente, y que se deducirá más adelante.

### 3.3.7 DEFINICIÓN DE LA PRUEBA

$$H_0 : F = F(\theta) \quad vs \quad H_a : F \neq F(\theta)$$

Para el caso que se analiza:

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

La hipótesis nula se rechaza si el valor de la estadística  $U_n^2$  es mayor que el valor crítico obtenido en tablas.

### 3.3.8 DISTRIBUCIÓN DE $U_n^2$

Para obtener las funciones de distribución de las estadísticas, se procederá en primer lugar a escribir  $U_n^2$  de otra manera, equivalente a la primera forma que fue dada:

Si se tienen  $n$  observaciones en orden ascendente  $\theta_1, \theta_2, \dots, \theta_n$ .

$$U_n^2 = n \int_0^{2\pi} \left\{ S_n(\theta) - F(\theta) - \int_0^{2\pi} [S_n(y) - F(y)] dF(y) \right\}^2 dF(\theta)$$

donde  $y = F(\theta) = P[\theta_i < \theta]$

se definirán unas variables aleatorias que servirán para expresar de otra manera la función  $S_n(\theta)$ :

$$\delta_i(t) = \begin{cases} 1 & \text{si } \theta_i < t \\ 0 & \text{e.o.c.} \end{cases}$$

Por construcción  $\delta_1(t), \delta_2(t), \dots, \delta_n(t)$  es un conjunto de v.a. con distribución Bernoulli ( $\lambda$ ), donde  $\lambda = P[\delta_i(t) = 1] = P[\theta_i < t] = F_\theta(t)$

$$\Rightarrow S_n(\theta) = \frac{1}{n} \sum_{i=1}^n \delta_i(\theta)$$

que será útil para simplificar la expresión inicial de la estadística  $U_n^2$ :

$$\int S_n(y) dF(y) = \int \sum_{i=1}^n \frac{\delta_i(\theta)}{n} dF(\theta) = \frac{1}{n} \sum_{i=1}^n F(\theta)$$

Por otro lado:

$$\int_0^{2\pi} F(y) dF(y) = \frac{[F(y)]^2}{2} \Big|_0^{2\pi} = \frac{1}{2}$$

$$\Rightarrow U_n^2 = \int_0^{2\pi} \left\{ S_n(\theta) - F(\theta) - \frac{1}{n} \sum_{i=1}^n F(\theta_i) + \frac{1}{2} \right\}^2 dF(\theta)$$

Existe una estadística que servirá para encontrar la forma simplificada de  $U_n^2$ , a esta estadística se le conoce como  $W_n^2$  y se define como:

$$W_n^2 = n \int_0^{2\pi} \{S_n(\theta) - F(\theta)\}^2 dF(\theta)$$

Que como se observa es más tratable que la misma  $U_n^2$ . Se puede ver que la estadística anterior tiene la forma del segundo momento de  $S_n(\theta) - F(\theta)$ , además  $U_n^2$  tiene la forma de la varianza, pues claramente las integrales representan la esperanza de  $S_n(\theta) - F(\theta)$  menos la esperanza de  $S_n(\theta) - F(\theta)$ , todo elevado al cuadrado. Hechas estas observaciones se puede establecer una relación entre las dos estadísticas:

$$\begin{aligned} U_n^2 &= n \operatorname{Var} [S_n(\theta) - F(\theta)] \\ &= n \{E[(S_n(\theta) - F(\theta))^2] - E^2[S_n(\theta) - F(\theta)]\} \\ &= W_n^2 - n \left[ \frac{1}{2} - \bar{Y} \right]^2 \end{aligned}$$

Esta última expresión se obtiene sustituyendo el valor de  $E[S_n(\theta) - F(\theta)]$ , obtenido anteriormente. Si en la fórmula de  $W_n^2$  se sustituye  $S_n(\theta)$  por  $\frac{x_i}{n}$ , cuando  $\theta_i \leq \theta \leq \theta_{i+1}$   $i = 0, 1, \dots, n$  y  $F(\theta)$  por  $y_i$ , donde las  $y_i$ 's son las  $n$  estadísticas uniformes ordenadas, se obtiene que:

$$\begin{aligned} W_n^2 &= \sum_{i=1}^n \left[ Y_i - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n} \\ \Rightarrow U_n^2 &= \left\{ \sum_{i=1}^n \left[ Y_i - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n} \right\} - n \left[ \frac{1}{2} - \bar{Y} \right]^2 \end{aligned}$$

Otra forma en la que se puede ver a  $U_n^2$  es :

$$\left\{ \sum_{i=1}^n 2 \left( Y_i - \frac{i}{n} + \frac{1}{2n} \right) \left( \frac{1}{2} - \bar{Y} \right) \right\} + n \left( \frac{1}{2} - \bar{Y} \right)^2 = -n \left( \frac{1}{2} - \bar{Y} \right)^2$$

Por lo que:

$$\begin{aligned} & \sum_{i=1}^n \left( Y_i - \frac{2i-1}{2n} \right)^2 + \sum_{i=1}^n 2 \left( Y_i - \frac{i}{n} + \frac{1}{2n} \right) \left( \frac{1}{2} - \bar{Y} \right) + \sum_{i=1}^n \left( \frac{1}{2} - \bar{Y} \right)^2 \\ = & \sum_{i=1}^n \left( Y_i - \frac{2i-1}{2n} \right)^2 - n \left( \frac{1}{2} - \bar{Y} \right)^2 \end{aligned}$$

Pero la parte derecha de la ecuación es igual a:

$$\begin{aligned} & \sum_{i=1}^n \left[ \left( Y_i - \frac{2i-1}{2n} \right) + \left( \frac{1}{2} - \bar{Y} \right) \right]^2 \\ \Rightarrow U_n^2 = & \sum_{i=1}^n \left[ Y_i - \frac{2i-1}{2n} + \frac{1}{2} - \bar{Y} \right]^2 + \frac{1}{12n} \end{aligned}$$

### 3.3.9 INVARIANZA DE LA PRUEBA

Para demostrar que el valor de:

$$U_n^2 = n \int_0^{2\pi} \left\{ S_n(\theta) - F(\theta) - \int_0^{2\pi} [S_n(y) - F(y)] dF(y) \right\}^2 dF(\theta)$$

es independiente del valor de la observación elegida como inicio, se supone que  $dF(\theta)$  es la función de densidad de probabilidad hipotética en  $\theta$ . Supóngase que existe otra función de distribución sobre el círculo y que  $G(\theta; \theta_1)$  y  $G(\theta; \theta_2)$  son sus funciones de distribución acumulada, pero empezando en  $\theta_1$  y  $\theta_2$ . Si se toma  $0 \leq \theta_2 - \theta_1 \leq 2\pi$ :

$$\begin{aligned}
G(\theta; \theta_1) &= G(\theta; \theta_2) + G(\theta_2; \theta_1) & \theta_2 \leq \theta \leq \theta_1 + 1 \\
&= G(\theta; \theta_2) - 1 + G(\theta_2; \theta_1) & \theta_1 + 1 \leq \theta \leq \theta_2 + 1 \\
G(\theta; \theta_1) - \int_0^{2\pi} G(\theta; \theta_1) dF(\theta) &= G(\theta; \theta_2) - \int_0^{2\pi} G(\theta; \theta_2) dF(\theta) + \int_{\theta_1}^{\theta_2} dF(\theta)
\end{aligned}$$

Si hay otra distribución sobre el círculo con funciones de distribución acumulada  $H(\theta; \theta_1), H(\theta; \theta_2)$  la fórmula anterior seguiría teniendo la misma forma, sólo que en vez de  $G$  tendríamos  $H$  de esta manera:

$$\{G(\theta; \theta_1) - H(\theta; \theta_1)\} - \int_0^{2\pi} [G(\theta; \theta_1) - H(\theta; \theta_1)] dF(\theta)$$

es independiente del valor  $\theta_1$  elegido. Sólo resta relacionar la función  $G$  con la función de distribución empírica  $S_n(\theta)$  y la función  $G$  con la distribución  $F(\theta)$ .

### 3.3.10 DISTRIBUCIÓN APROXIMADA DE $U_n^2$

Pearson obtuvo una aproximación de una distribución Ji-cuadrada no central, en términos de una distribución Ji-cuadrada central.

Esto da una idea de que se puede aproximar la distribución de  $U_n^2$  si se supone que  $U_n^2 \sim a + b\chi^2$ , donde  $\chi^2$  tiene distribución Ji cuadrada con  $f$  grados de libertad y además  $a, b$  y  $f$  se determinan de tal forma que  $a + b\chi^2$ , tenga los tres primeros momentos iguales a los de  $U_n^2$ , es decir, se deberá cumplir que :

$$\begin{aligned}
\mu_1 &= \frac{1}{12} \\
\mu_2 &= \frac{n-1}{360n} + \left(\frac{1}{12}\right)^2
\end{aligned}$$

$$\mu_3 = \frac{2n^3 - 5n + 3}{7560n^2}$$

que son los momentos obtenidos de la forma simple de  $U_n^2$ . Es importante mencionar que para calcular estos momentos se usan métodos muy complejos, ya que como puede observarse para cada valor de  $n$ , la función  $U_n^2$  varía en gran manera; Maguire, Pearson y Wynn hicieron estos cálculos considerando el caso en que  $n \rightarrow \infty$ .

Resolviendo el sistema de ecuaciones con tres incógnitas se obtiene que :

$$\begin{aligned} f &= \frac{49n(n-1)}{20\left(n-\frac{3}{2}\right)^2} \\ b &= \frac{1}{42n}\left(n-\frac{3}{2}\right) \\ a &= \frac{21n-56}{840\left(n-\frac{3}{2}\right)} \end{aligned}$$

En efecto, si se toma por ejemplo  $n = 2$  :

$$a = -\frac{1}{30} \quad b = \frac{1}{168} \quad f = \frac{196}{10}$$

Por lo que:

$$E(a + b\chi^2) = -\frac{1}{30} + \frac{1}{168} \frac{196}{10} = \frac{35}{420} = \frac{1}{12}$$

Lo anterior se esperaba pues  $\mu_1 = \frac{1}{12}$ .

$$\begin{aligned} E\left[(a + b\chi^2)^2\right] &= \text{Var}(a + b\chi^2) + \left(\frac{1}{12}\right)^2 \\ &= \left(\frac{1}{168}\right)^2 \text{Var} \chi^2 + \frac{1}{144} \end{aligned}$$

y como  $\text{Var } \chi^2$  es  $2f = \frac{196}{5}$ , la expresión anterior se reduce a  $\frac{6}{720}$ , que coincide con

$\mu_2 = \frac{1}{720} + \frac{1}{144} = \frac{6}{720}$ . Lo mismo ocurre para el tercer momento.

Aunque este es un método artificial y sencillo, da resultados muy parecidos a los que se obtienen de otras aproximaciones asintóticas de la distribución de  $U_n^2$ , pero que son mucho más complejas.

## Capítulo 4

# APLICACIONES

La teoría de la Estadística Circular desarrollada hasta ahora se ha aplicado básicamente al estudio de fenómenos que tienen que ver con la migración de especies animales, para probar si la orientación que siguen es aleatoria o no; pero es claro que ésta no debe ser la única situación en la que se puede aplicar la teoría.

Se puede hablar por ejemplo de la medicina, que ha jugado un papel importantísimo en el desarrollo de la humanidad. Hasta hace poco a los médicos se les enseñaba que el organismo humano vive en homeostasis, característica de los organismos vivos, especialmente el hombre, de reaccionar frente a las variaciones del medio externo, con el fin de alcanzar siempre un equilibrio fisiológico. Sin embargo desde hace algunos años se ha podido demostrar cuánto cambia el organismo a causa de los ritmos circadianos. Estos ritmos biológicos naturales son vitales, como por ejemplo, los latidos del corazón. Se dice que no hay una sola función del organismo que no tenga su propio ritmo, la ausencia de ritmo significa muerte. La presión arterial, la temperatura, la frecuencia cardiaca, la producción de diversas sustancias en nuestro

cuerpo ¿ cómo cambian en el transcurso del día ? Antes de estudiar los ritmos, cabe preguntarse si en realidad existen y esto lleva a pensar en la aleatoriedad de algunas variables que ya han sido mencionadas, en el siguiente sentido:

Se pueden registrar los cambios del número de latidos del corazón, de la temperatura corporal, de la presión arterial, etc. a lo largo de un día, durante un periodo predeterminado. Al preguntar si los cambios máximos entre las unidades de medición de estas variables se dan en algunas horas fijas o no, se puede pensar en realizar una prueba de aleatoriedad. Además si se consideran individuos fumadores, alcohólicos, etc., se esperaría encontrar algún cambio en la realización de las pruebas de aleatoriedad.

Al preguntar cuánto tiempo tarda un bebé en estabilizar su horario de sueño o alimentación, se podrían registrar las horas de estos sucesos y realizar una prueba de uniformidad (aleatoriedad). Se esperaría que al inicio la uniformidad no fuera rechazada, pero en algún momento este resultado cambiaría.

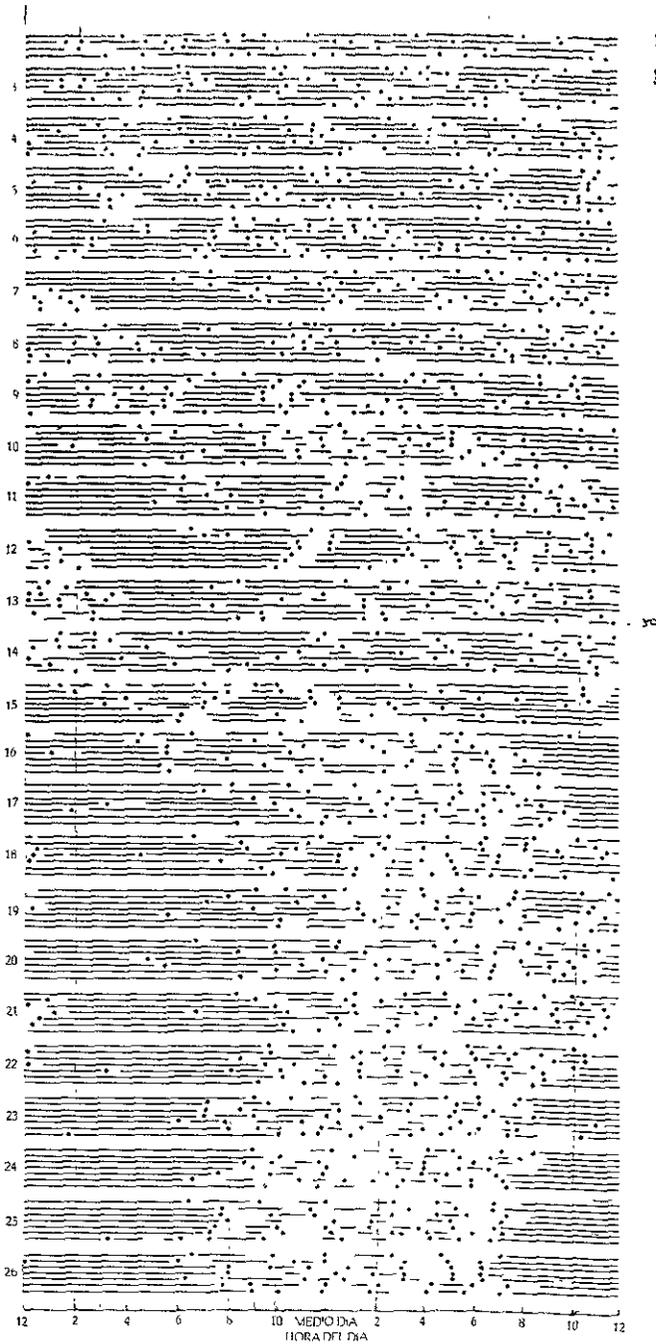
Otra situación pudiera ser, comparar las horas en las que los índices de contaminación son más altos o más bajos, al paso de varios años, lo cual pudiera ser de utilidad en un estudio o evaluación de un plan para combatir la contaminación.

Desde hace un siglo, el viento se utiliza para generar electricidad; estimaciones del Instituto de Investigaciones Eléctricas ( IIE ) ubican a México como el poseedor del mayor potencial de recursos eólicos a nivel mundial. A partir de 1983 se han intensificado en México los estudios de energía eólica con el propósito de alcanzar una mejor administración de los recursos energéticos nacionales, ante esto se puede pensar en una prueba de aleatoriedad para estudiar y, por ejemplo, analizar la dirección de

los vientos en determinados lugares y periodos de tiempo

Con esto se ve que sí se puede pensar en ampliar el campo de aplicación de la estadística circular. Se ha escogido uno de estos ejemplos al que se aplicarán las pruebas explicadas en los capítulos anteriores. Los datos se obtuvieron de un libro editado por la Secretaría de Educación Pública, que contiene ejemplos interesantes del tratamiento de la información [8]. Las tablas usadas fueron tomadas de Batschelet[1].

Los ciclos de alimentación de un bebé lactante: se pretende ver los datos de la gráfica mostrada en la página siguiente con un enfoque distinto. Se tomará el último día de cada uno de los siete meses de los que se tiene información y se aplicarán las pruebas.

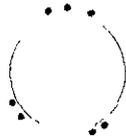


### Un ejemplo de tratamiento gráfico de la información

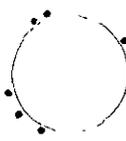
Ciclos de alimentación de un niño que lacta el pecho registrados desde el 11° hasta el 182° día de vida. Las líneas continuas representan los períodos de sueño, los puntos indican las tetadas y, los espacios en blanco, los períodos de vigilia. Es evidente que la madre ha acostumbrado progresivamente al niño a renunciar a la alimentación nocturna, por lo que se merma el efecto contraceptivo de la lactancia.



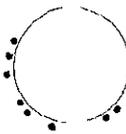
1er. mes



3er. mes



5o. mes



7o. mes



2o. mes



4o. mes



6o. mes

Un análisis visual permite establecer algunas diferencias importantes entre las observaciones del primer y séptimo mes:

En la gráfica correspondiente al primer mes, las horas de alimentación del bebé están más dispersas sobre la circunferencia, en cambio la última gráfica, que corresponde al séptimo mes, muestra una clara concentración de las horas de alimentación. Pero la pregunta es, si estadísticamente se puede hablar de un establecimiento de horario o no del lactante. Esta pregunta se traduce en un planteamiento de hipótesis propio: ¿se puede pensar que la muestra se ajusta a una distribución uniforme o aleatoria en el primer mes? ¿se ajusta a otra distribución o patrón en el séptimo mes?

**ESTA TESIS NO SALE  
DE LA BIBLIOTECA**

El resultado que se espera es que en los primeros meses la hipótesis nula de uniformidad no se rechace, ya que el bebé “come a cualquier hora”, sin embargo al paso del tiempo su horario se va estabilizando y termina por tener un horario “especial” para su alimentación durante el día y para dormir por las noches, en este caso se espera que la hipótesis nula de uniformidad sea rechazada.

La prueba de Greenwood y Durand no se aplicará al ejemplo, ya que por la naturaleza del mismo no se tiene una dirección teórica predeterminada para la hipótesis alternativa; tampoco se aplicará la prueba de la Ji Cuadrada, ya que no se cumple la condición de que la frecuencia esperada en cada intervalo sea al menos 4, pues en el ejemplo hay meses en los que solo se tienen 6 y 7 observaciones.

Se utilizarán dos valores de  $\alpha$ , error tipo I (probabilidad de rechazar la hipótesis nula, cuando ésta es cierta),  $\alpha = 0.05$  y  $\alpha = 0.01$

## 4.1 PRUEBA DE RAYLEIGH

Debe considerarse una sencilla transformación para pasar de las horas a los grados que en esta prueba se requieren:

$$g_i = 15h_i + 0.25m_i \quad i = 1, \dots, n$$

$g_i$  : grados correspondientes a la  $i$  -ésima observación

$h_i$  : horas enteras de la  $i$  -ésima observación

$m_i$  : minutos de la  $i$  -ésima observación

Así por ejemplo, en la primera observación del primer mes, las 3:40 equivalen a

$$(3)(15) + (40)(0.25) = 15 + 10 = 55^\circ$$

		1er.	mes
HORA	GRADOS	COSENO	SENO
3:40	55.0	0.5735	0.8191
7:30	112.5	-0.3826	0.9238
8:10	122.5	-0.5372	0.8433
9:00	135.0	-0.7071	0.7071
10:30	157.5	-0.9238	0.3826
15:00	225.0	-0.7071	-0.7071
19:10	287.5	0.3007	-0.9537
20:40	310.0	0.6427	-0.7660

$$\bar{x} = \sum_{i=1}^8 \cos \theta_i = 0.2176 \quad \bar{y} = \sum_{i=1}^8 \sen \theta_i = 0.1561$$

$$r = 0.2678$$

Para  $n = 8$  y  $\alpha = 0.05$ ,  $P = 0.597$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $n = 8$  y  $\alpha = 0.10$ ,  $P = 0.597$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir

que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

2o mes

HORA	GRADOS	COSENO	SENO
2:30	37.5	0.7933	0.6087
7:00	105.0	-0.2588	0.9659
9:00	135.0	-0.7071	0.7071
9:30	142.5	-0.7933	0.6087
14:00	210.0	-0.8660	-0.5000
15:00	225.0	-0.7071	-0.7071
16:00	240.0	-0.5000	-0.8660
17:00	255.0	-0.2588	-0.9659
23:00	345.0	0.9659	-0.2588

$$\bar{x} = \sum_{i=1}^9 \cos \theta_i = -0.2591 \quad \bar{y} = \sum_{i=1}^9 \sin \theta_i = -0.0452$$

$$r = 0.2630$$

Para  $n = 9$  y  $\alpha = 0.05$ ,  $P = 0.558$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $n = 9$  y  $\alpha = 0.10$ ,  $P = 0.558$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Nuevamente, como en el primer mes, lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que el bebé

se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

3er. mes			
HORA	GRADOS	COSENO	SENO
1:50	27.5	0.8870	0.4617
9:00	135.0	-0.7071	0.7071
11:30	172.5	-0.9914	0.1305
15:30	232.5	-0.6087	-0.7933
16:40	250.0	-0.3420	-0.9396
22:30	337.5	0.9238	-0.3826
24:00	360.0	1.0000	0.0000

$$\bar{x} = \sum_{i=1}^7 \cos \theta_i = 0.0230 \quad \bar{y} = \sum_{i=1}^7 \text{sen} \theta_i = -0.1166$$

$$r = 0.1188$$

Para  $n = 7$  y  $\alpha = 0.05$ ,  $P > 0.9$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $n = 7$  y  $\alpha = 0.10$ ,  $P > 0.9$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que el bebé se alimenta en tiempos aleatorios. lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

4o. mes

HORA	GRADOS	COSENO	SENO
0:00	0.0	1.0000	0.0000
4:30	67.5	0.3826	0.9238
6:20	95.0	-0.0871	0.9961
10:00	150.0	-0.8660	0.5000
15:00	225.0	-0.7071	-0.7071
17:00	255.0	-0.2588	-0.9659
21:30	322.5	0.7933	-0.6087

$$\bar{x} = \sum_{i=1}^7 \cos \theta_i = 0.0367 \quad \bar{y} = \sum_{i=1}^7 \sen \theta_i = 0.0197$$

$$r = 0.0416$$

Para  $n = 7$  y  $\alpha = 0.05$ ,  $P > 0.9$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $n = 7$  y  $\alpha = 0.10$ ,  $P > 0.9$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

La hipótesis de uniformidad no se rechaza, es decir, hay evidencia estadística para afirmar que el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

5o. mes

HORA	GRADOS	COSENO	SENO
5:30	82.5	0.1305	0.9914
13:00	195.0	-0.9659	-0.2588
15:30	232.5	-0.6087	-0.7933
17:50	267.5	-0.0436	-0.9990
19:50	297.5	0.4617	-0.8870
23:00	345.0	0.9659	-0.2588

$$\bar{x} = \sum_{i=1}^6 \cos \theta_i = -0.0100 \quad \bar{y} = \sum_{i=1}^6 \sin \theta_i = -0.3675$$

$$r = 0.3676$$

Para  $n = 6$  y  $\alpha = 0.05$ ,  $P = 0.477$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $n = 6$  y  $\alpha = 0.10$ ,  $P = 0.477$

$P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

Go. mes

HORA	GRADOS	COSENO	SENO
6:30	97.5	-0.1305	0.9914
11:00	165.0	-0.9659	0.2588
12:40	190.0	-0.9848	-0.1736
13:45	206.25	-0.8968	-0.4422
15:00	225.0	-0.7071	-0.7071
16:20	245.0	-0.4226	-0.9063
17:15	258.75	-0.1950	-0.9807
19:00	285.0	0.2588	-0.9659

$$\bar{x} = \sum_{i=1}^8 \cos \theta_i = -0.5054 \quad \bar{y} = \sum_{i=1}^8 \sin \theta_i = -0.3657$$

$$r = 0.6238$$

Para  $n = 8$  y  $\alpha = 0.05$ ,  $P = 0.041$

$P < \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *se rechaza*

Para  $n = 8$  y  $\alpha = 0.10$ ,  $P = 0.041$

$P < \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *se rechaza*

Ahora han cambiado los resultados que arroja la prueba, pues se tiene que se rechaza la hipótesis de uniformidad, por lo que hay evidencia estadística para afirmar que el bebé ya no se alimenta en tiempos aleatorios, lo que nos lleva a decir que ha establecido un horario de alimentación que abarca un mínimo número de horas.

7o. mes

HORA	GRADOS	COSENO	SENO
7:20	140.0	-0.7660	0.6427
10:20	155.0	-0.9063	0.4226
12:15	183.75	-0.9978	-0.0654
14:30	217.5	-0.7933	-0.6087
15:00	225.0	-0.7071	-0.7071
16:00	240.0	-0.5000	-0.8660
18:00	270.0	0	-1.0000
19:00	285.0	0.2588	-0.9659

$$\bar{x} = \sum_{i=1}^8 \cos \theta_i = -0.5514 \quad \bar{y} = \sum_{i=1}^8 \sin \theta_i = -0.3934$$

$$r = 0.6773$$

Para  $n = 8$  y  $\alpha = 0.05$ ,  $P = 0.022$

$P < \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *se rechaza*

Para  $n = 8$  y  $\alpha = 0.10$ ,  $P = 0.022$

$P < \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *se rechaza*

Si en el sexto mes se rechazó la uniformidad, se espera que en los meses siguientes también se rechace. Nuevamente los resultados de la prueba indican que se rechaza la hipótesis de uniformidad, por lo que hay evidencia estadística para afirmar que el bebé ya no se alimenta en tiempos aleatorios, lo que nos lleva a decir que ha establecido un horario de alimentación que abarca un mínimo número de horas.

## 4.2 PRUEBA DEL RANGO

Recuérdese que para la aplicación de esta prueba se necesita encontrar el arco más pequeño que contenga todas las observaciones. Se explicará con detalle el proceso del cálculo en el primer mes y se simplificará en los meses siguientes.

### 1er. mes

De acuerdo a la definición de  $T_i$ , de la pág. 4 y con las horas convertidas a medidas angulares como en la prueba anterior:

$i$	$T_i$
1	57.5
2	10.0
3	12.5
4	22.5
5	67.5
6	62.5
7	22.5
8	105

$$\text{Rango: } w = 360^\circ - \max(T_1, \dots, T_n) = 360^\circ - 105^\circ = 255^\circ$$

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 174.4^\circ$$

$$w = 255^\circ > w(\alpha) = 174.4^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 192.5^\circ$$

$$w = 255^\circ > w(\alpha) = 192.5^\circ$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechaza la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

**2o. mes**

$$\text{Rango: } w = 360^\circ - 67.5^\circ = 292.5^\circ$$

$$\text{Para } n = 9 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 188.1^\circ$$

$$w = 292.5^\circ > w(\alpha) = 188.1^\circ$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 9 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 205.1^\circ$$

$$w = 292.5^\circ > w(\alpha) = 205.1^\circ$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Como en el mes anterior, no se rechaza la hipótesis de uniformidad por lo que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

**3er. mes**

$$\text{Rango: } w = 360^\circ - 107.5^\circ = 252.5^\circ$$

$$\text{Para } n = 7 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 158^\circ$$

$$w = 252.5^\circ > w(\alpha) = 158^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 7 \quad \text{y} \quad \alpha = 0.10, \quad w(\alpha) = 177.3^\circ$$

$$w = 252.5^\circ > w(\alpha) = 177.3^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad nuevamente, es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

#### 4o. mes

$$\text{Rango: } w = 360^\circ - 75^\circ = 285^\circ$$

$$\text{Para } n = 7 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 158^\circ$$

$$w = 258^\circ > w(\alpha) = 158^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 7 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 177.3^\circ$$

$$w = 258^\circ > w(\alpha) = 177.3^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Al no rechazarse la hipótesis de uniformidad se deduce que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

**5o. mes**

$$\text{Rango: } w = 360^\circ - 112.5^\circ = 247.5^\circ$$

$$\text{Para } n = 6 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 138.2^\circ$$

$$w = 247.5^\circ > w(\alpha) = 138.2^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 6 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 158.7^\circ$$

$$w = 247.5^\circ > w(\alpha) = 158.7^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Como no se rechaza la hipótesis de uniformidad se infiere que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

**6o. mes**

$$\text{Rango: } w = 360^\circ - 172.5^\circ = 187.5^\circ$$

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 174.4^\circ$$

$$w = 187.5^\circ > w(\alpha) = 174.4^\circ$$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 192.5^\circ$$

$$w = 187.5 < w(\alpha) = 192.5$$

$\therefore$  La hipótesis nula de uniformidad *se rechaza*

En este mes con distintos valores de  $\alpha$ , obtenemos resultados distintos de la prueba.

Recuérdese que  $\alpha$  representa la probabilidad del error tipo I, que es de rechazar la

hipótesis nula, cuando ésta fuere cierta, es decir la probabilidad de que el lactante aún no haya establecido su horario de alimentación, aunque la prueba indique lo contrario es de 0.1.

**7o. mes**

$$\text{Rango: } w = 360^\circ - 215^\circ = 145^\circ$$

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.05; \quad w(\alpha) = 174.4^\circ$$

$$w = 145^\circ < w(\alpha) = 174.4^\circ$$

$\therefore$  La hipótesis nula de uniformidad *se rechaza*

$$\text{Para } n = 8 \quad \text{y} \quad \alpha = 0.10; \quad w(\alpha) = 192.5^\circ$$

$$w = 145^\circ < w(\alpha) = 192.5^\circ$$

$\therefore$  La hipótesis nula de uniformidad *se rechaza*

Si en el sexto mes se rechazó la uniformidad con un valor de  $\alpha$ , se espera que en los meses siguientes también se rechace. Ahora los resultados de la prueba indican que se rechaza la hipótesis de uniformidad con los dos valores de  $\alpha$ , por lo que hay evidencia estadística para afirmar que el bebé ya no se alimenta en tiempos aleatorios, lo que nos lleva a decir que ha establecido un horario de alimentación que abarca un mínimo número de horas.

### 4.3 PRUEBA DE ESPACIOS DE RAO

Se aclara que las diferencias positivas y negativas de las tablas se refieren a las de la expresión  $T_i - (360^\circ/n)$ ,  $i = 1, 2, \dots, n$ . La forma de obtener  $T_i$  se explica en la página

Por ejemplo, en la primera observación del primer mes, la diferencia es positiva:

$$T_1 - (360^\circ/8) = 57.5 - 45 = 12.5$$

En la segunda observación del primer mes, la diferencia es negativa:

$$T_2 - (360^\circ/8) = 10 - 45 = -35$$

1er. mes

$$n = 8 \quad 360/8 = 45$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	57.5	12.5	
2	10.0		35.0
3	12.5		32.5
4	22.5		32.5
5	67.5	22.5	
6	62.5	17.5	
7	22.5		22.5
8	105	60.0	
Total		112.5	112.5

La estadística de prueba  $U = 112.5^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 175.7^\circ$

$U < U(\alpha)$  · La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 163.4^\circ$

$U < U(\alpha)$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

## 2o. mes

$$n = 9 \quad 360/9 = 40$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	67.5	27.5	
2	30.0		10.0
3	7.5		32.5
4	67.5	27.5	
5	15.0		25.0
6	15.0		25.0
7	15.0		25.0
8	90	50.0	
9	52.5	12.5	
Total		117.5	117.5

La estadística de prueba  $U = 117.5^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 173.5^\circ$

$U < U(\alpha)$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 162.4^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*.

Al no rechazarse la hipótesis de uniformidad se deduce que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 3er. mes

$$n = 7 \quad 360/7 = 51.4$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	107.5	56.1	
2	37.5		13.9
3	60.0	8.6	
4	17.5		33.9
5	87.5	36.1	
6	22.5		28.9
7	27.5		23.9
Total		100.8	100.6

La estadística de prueba  $U = 100.8^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 177.8^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 164.9^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Como nuevamente no se rechaza la hipótesis de uniformidad, hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

4o. mes

$$n = 7 \quad 360/8 = 51.4$$

$z$	$T_i$	Diferencias positivas	Diferencias negativas
1	67.5	16.1	
2	27.5		23.9
3	55.0	3.6	
4	75.0	23.6	
5	30.0		21.4
6	67.5	16.1	
7	37.5		13.9
Total		59.4	59.2

La estadística de prueba  $U = 59.4^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 177.8^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 164.9^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución

uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

5o. mes

$$n = 6 \quad 360/6 = 60$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	112.5	52.5	
2	37.5		22.5
3	35.0		25.0
4	30.0		30.0
5	47.5		12.5
6	97.5	37.5	
Total		90.0	90.0

La estadística de prueba  $U = 90^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 180.7^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 166.3^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Si *no se rechaza* la hipótesis de uniformidad, hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

6o. mes

$$n = 8 \quad 360/8 = 45$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	67.5	22.5	
2	25.0		20.0
3	16.25		28.75
4	18.75		26.25
5	20.0		25.00
6	13.75		31.25
7	26.25		18.75
8	172.5	127.5	
Total		150	150

La estadística de prueba  $U = 150^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 175.7^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 163.4^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, es decir, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

7o. mes

$$n = 8 \quad 360/8 = 45$$

$i$	$T_i$	Diferencias positivas	Diferencias negativas
1	15.0		30.0
2	28.75		16.25
3	33.75		11.25
4	7.5		37.5
5	15.0		30.0
6	30.0		15.0
7	15.0		30.0
8	215.0	170.0	
Total		170.0	170.0

La estadística de prueba  $U = 170.0^\circ$

Para  $\alpha = 0.05$ ,  $U(\alpha) = 175.7^\circ$

$U < U(\alpha) \therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $\alpha = 0.10$ ,  $U(\alpha) = 163.4^\circ$

$U > U(\alpha) \therefore$  La hipótesis nula de uniformidad *se rechaza*

En este mes con distintos valores de  $\alpha$ , obtenemos resultados distintos de la prueba.

Recuérdese que  $\alpha$  representa la probabilidad del error tipo I. que es de rechazar la hipótesis nula, cuando ésta fuere cierta, es decir la probabilidad de que el lactante aún no haya establecido su horario de alimentación, aunque la prueba indique lo contrario, es de 0.1.

## 4.4 PRUEBA DE HODGES Y AJNE

La estadística de prueba  $K$ , que representa el número mínimo de puntos que pueda quedar de un lado de un diámetro, puede obtenerse visualmente de las gráficas están al inicio de este capítulo.

### 1er. mes

Para  $K = 1$  y  $n = 8$ ,  $P = 0.375$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 1$  y  $n = 8$ ,  $P = 0.375$

Si  $\alpha = 0.10$ ,  $P > \alpha$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 2o. mes

Para  $K = 2$  y  $n = 9$ ,  $P = 0.703$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 2$  y  $n = 9$ ,  $P = 0.703$

Si  $\alpha = 0.10$ ,  $P > \alpha$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 3er. mes

Para  $K = 3$  y  $n = 7$ ,  $P > 0.9$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 1$  y  $n = 8$ ,  $P > 0.9$

Si  $\alpha = 0.10$ ,  $P > \alpha$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 4o. mes

Para  $K = 3$  y  $n = 7$ ,  $P > 0.9$

Si  $\alpha = 0.05$ ,  $P > \alpha$ .  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 3$  y  $n = 7$ ,  $P > 0.9$

Si  $\alpha = 0.10$ ,  $P > \alpha$  . La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechaza la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 5o. mes

Para  $K = 1$  y  $n = 6$ ,  $P = 0.75$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 1$  y  $n = 6$ ,  $P = 0.75$

Si  $\alpha = 0.10$ ,  $P > \alpha$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechaza la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

### 6o. mes

Para  $K = 1$  y  $n = 8$ ,  $P = 0.375$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 1$  y  $n = 8$ ,  $P = 0.375$

Si  $\alpha = 0.10$ ,  $P > \alpha$   $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechaza la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

#### 7o. mes

Para  $K = 0$  y  $n = 8$ ,  $P = 0.063$

Si  $\alpha = 0.05$ ,  $P > \alpha$ ,  $\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Para  $K = 1$  y  $n = 8$ ,  $P = 0.06$

Si  $\alpha = 0.10$ ,  $P < \alpha$   $\therefore$  La hipótesis nula de uniformidad *se rechaza*

En este mes con distintos valores de  $\alpha$ , obtenemos resultados distintos de la prueba. Recuérdese que  $\alpha$  representa la probabilidad del error tipo I, que es de rechazar la hipótesis nula, cuando ésta fuere cierta, es decir la probabilidad de que el lactante aún no haya establecido su horario de alimentación, aunque la prueba indique lo contrario es de 0.1.

## 4.5 PRUEBA DE KUIPER

Recuérdese que la estadística  $K$ , de esta prueba se basa en las diferencias entre la distribución teórica, que es la distribución uniforme y la función de distribución empírica, basada en las observaciones, que en este caso son las horas de alimentación del bebé. Para realizar las tablas se aclara que:

$$z_i = 15h_i + 0.25m_i \quad i = 1, \dots, n$$

$z_i$  : grados correspondientes a la  $i$  - ésima observación

$h_i$  : horas enteras de la  $i$  - ésima observación

$m_i$  : minutos de la  $i$  - ésima observación

Así por ejemplo, en la primera observación del primer mes, las 3:40 equivalen a  
 $(3)(15) + (40)(0.25) = 45 + 10 = 55^\circ$

$$F(z_i) = \frac{z_i}{360}$$

Entonces  $F(55^\circ) = \frac{55}{360} = 0.1528$  (Redondeando hasta diezmilésimos)

La definición de  $S_n(z_i)$  es:

$$S_n(z_i) = \begin{cases} 0 & \text{si } z < z_1 \\ \frac{i}{n} & \text{si } z_i \leq z < z_{i+1} \\ 1 & \text{si } z_n \leq z \end{cases}$$

Entonces  $S_8(55) = 1/8 = 0.125$

Obtenidos estos valores, se completan las columnas restantes.

Los valores mayores de cada columna, son  $D^+$  y  $D^-$ , respectivamente.

1er. mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
55	0.1528	0.125	-0.0278	0.1528
112.5	0.3125	0.25	-0.0625	<b>0.1875</b>
122.5	0.3403	0.375	0.0347	0.0903
135	0.375	0.5	0.125	0
157.5	0.4375	0.625	<b>0.1875</b>	-0.0625
225	0.625	0.75	0.125	0
287.5	0.7986	0.875	0.0764	0.0486
310	0.8611	1	0.1389	-0.0139

$$n = 8 \quad D^+ = 0.1875 \quad D^- = 0.1875$$

$$V_n = D^+ + D^- = 0.3750$$

$$K = n^{\frac{1}{2}} V_n = 1.0607$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.608 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.493 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechaza la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

2o. mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
37.5	0.1042	0.1111	0.0069	0.1042
105	0.2917	0.2222	-0.0695	<b>0.1806</b>
135	0.375	0.3333	-0.0417	0.1528
142.5	0.3958	0.4444	0.0486	0.0625
210	0.5833	0.5556	-0.0277	0.1389
225	0.625	0.6667	0.0417	0.0694
240	0.6667	0.7778	0.1111	0
255	0.7083	0.8889	<b>0.1806</b>	-0.0695
345	0.9583	1	0.0417	0.0694

$$n = 9 \quad D^+ = 0.1806 \quad D^- = 0.1806$$

$$V_n = D^+ + D^- = 0.3612$$

$$K = n^{\frac{1}{2}} V_n = 1.0836$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.618 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.500 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de

horas.

3er. mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
27.5	0.0764	0.1429	<b>0.0665</b>	0.0764
135	0.375	0.2857	-0.0893	<b>0.2321</b>
172.5	0.4792	0.4286	-0.0506	0.1935
232.5	0.6458	0.5714	-0.0744	0.2172
250	0.6944	0.7143	0.0199	0.123
337.5	0.9375	0.8571	-0.0804	0.2232
360	1	1	0	0.1429

$$n = 7 \quad D^+ = 0.0665 \quad D^- = 0.2321$$

$$V_n = D^+ + D^- = 0.2986$$

$$K = n^{\frac{1}{2}} V_n = 0.7900$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.598 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.483 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

4o. mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
0	0	0.1429	0.1429	0
67.5	0.1875	0.2857	0.0982	0.0446
95	0.2639	0.4286	<b>0.1647</b>	-0.0218
150	0.4167	0.5714	0.1547	-0.0119
225	0.625	0.7143	0.0893	<b>0.0536</b>
255	0.7083	0.8671	0.1488	-0.006
322.5	0.8958	1	0.1042	0.0387

$$n = 7 \quad D^+ = 0.1647 \quad D^- = 0.0536$$

$$V_n = D^+ + D^- = 0.2183$$

$$K = n^{\frac{1}{2}} V_n = 1.0607$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.598 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.483 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

5o. mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
82.5	0.2292	0.1667	-0.0625	0.2292
195	0.5417	0.3333	-0.2084	<b>0.375</b>
232.5	0.6458	0.5	-0.1458	0.3125
267.5	0.7431	0.6667	-0.0764	0.2431
297.5	0.8264	0.8333	0.0069	0.1597
345	0.9583	1	<b>0.0417</b>	0.125

$$n = 6 \quad D^+ = 0.0417 \quad D^- = 0.375$$

$$V_n = D^+ + D^- = 0.4167$$

$$K = n^{\frac{1}{2}} V_n = 1.0207$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.582 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.471 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un *mínimo* número de horas.

60. MCS

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
97.5	0.2708	0.125	-0.1458	0.2708
165	0.4583	0.25	-0.2083	<b>0.3333</b>
190	0.5278	0.375	-0.1528	0.2778
206.25	0.5729	0.5	-0.0729	0.1979
225	0.625	0.625	0	0.125
245	0.6806	0.75	0.0694	0.0556
258.75	0.7188	0.875	0.1562	-0.0312
285	0.7917	1	<b>0.2083</b>	-0.0833

$$n = 8 \quad D^+ = 0.2083 \quad D^- = 0.3333$$

$$V_n = D^+ + D^- = 0.5416$$

$$K = n^{\frac{1}{2}} V_n = 1.5319$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.608 \Rightarrow K < K(\alpha)$$

∴ La hipótesis nula de uniformidad *no se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.493 \Rightarrow K > K(\alpha)$$

∴ La hipótesis nula de uniformidad *se rechaza*

En este mes con distintos valores de  $\alpha$ . obtenemos resultados distintos de la prueba.

Recuérdese que  $\alpha$  representa la probabilidad del error tipo I, que es de rechazar la hipótesis nula, cuando ésta fuere cierta, es decir la probabilidad de que el lactante aún no haya establecido su horario de alimentación, aunque la prueba indique lo contrario es de 0.1.

70.            mes

$z_i$	$F(z_i)$	$S_n(z_i)$	$S_n(z_i) - F(z_i)$	$F(z_i) - S_n(z_{i-1})$
140	0.3889	0.125	-0.2639	<b>0.3889</b>
155	0.4306	0.25	-0.1806	0.3056
183.75	0.5104	0.375	-0.1354	0.2604
217.5	0.6042	0.5	-0.1042	0.2292
225	0.625	0.625	0	0.125
240	0.6667	0.75	0.0833	0.0417
270	0.75	0.875	0.125	0
285	0.7917	1	<b>0.2083</b>	-0.0833

$$n = 8 \quad D^+ = 0.2083 \quad D^- = 0.3889$$

$$V_n = D^+ + D^- = 0.5972$$

$$K = n^{\frac{1}{2}} V_n = 1.6891$$

$$\text{Si } \alpha = 0.05, K(\alpha) = 1.608 \Rightarrow K > K(\alpha)$$

∴ La hipótesis nula de uniformidad *se rechaza*

$$\text{Si } \alpha = 0.10, K(\alpha) = 1.493 \Rightarrow K > K(\alpha)$$

∴ La hipótesis nula de uniformidad *se rechaza*

Ahora han cambiado los resultados que arroja la prueba, pues se tiene que se rechaza la hipótesis de uniformidad, por lo que hay evidencia estadística para afirmar que el bebé ya no se alimenta en tiempos aleatorios, lo que nos lleva a decir que ha establecido un horario de alimentación que abarque un mínimo número de horas.

## 4.6 PRUEBA DE WATSON

Se elaborará una tabla con los datos de cada mes con el objeto de sistematizar el

cálculo de la estadística de prueba  $U_n^2$ . Es necesario aclarar que:

$$c_i = 2i - 1$$

$\phi_i$  : grados correspondientes a la  $i$  - ésima observación

$$v_i = \phi_i/360$$

$$\bar{v} = \frac{\sum_{i=1}^n v_i}{n}$$

		1er.	mes			
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i/n$	
1	1	55	0.1528	0.0233	0.0191	
2	3	112.5	0.3125	0.0977	0.1172	
3	5	122.5	0.3403	0.1158	0.2127	
4	7	135	0.3750	0.1406	0.3281	
5	9	157.5	0.4375	0.1914	0.4922	
6	11	225	0.6250	0.3906	0.8594	
7	13	287.5	0.7986	0.6378	1.2977	
8	15	310	0.8611	0.7415	1.6146	
Total			3.9028	2.3387	4.9410	

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i/n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 2.3387 - 4.941 + 8 \left( \frac{1}{3} - 0.0001488 \right)
 \end{aligned}$$

$$= 0.063176$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.181 \Rightarrow U_n^2 < U(\alpha)$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.150 \Rightarrow U_n^2 < U(\alpha)$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

		2o. mes			
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i / n$
1	1	37.5	0.1042	0.0109	0.0116
2	3	105	0.2917	0.0851	0.0972
3	5	135	0.375	0.1406	0.2083
4	7	142.5	0.3958	0.1567	0.3078
5	9	210	0.5833	0.3402	0.5833
6	11	225	0.6250	0.3906	0.7639
7	13	240	0.6667	0.4445	0.9630
8	15	255	0.7083	0.5017	1.1805
9	17	345	0.9583	0.9183	1.8101
Total		4.7083	2.9886	5.9257	

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i / n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 2.9886 - 5.9257 + 9 \left( \frac{1}{3} - 0.0005357 \right) \\
 &= 0.0581
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.182 \Rightarrow U_n^2 < U(\alpha)$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.150 \Rightarrow U_n^2 < U(\alpha)$

$\therefore$  La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

3er mes					
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i / n$
1	1	27.5	0.0764	0.0058	0.0109
2	3	135	0.375	0.1406	0.1607
3	5	172.5	0.4792	0.2296	0.3423
4	7	232.5	0.6458	0.4171	0.6458
5	9	250	0.6944	0.4822	0.8928
6	11	337.5	0.9375	0.8789	1.4732
7	13	360	1	1	1.8571
Total			4.2083	3.1542	5.3828

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i / n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 3.1542 - 5.3828 + 7 \left( \frac{1}{3} - 0.010238 \right) \\
 &= 0.0330
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.180 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.149 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir

que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

		4o. mcs			
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i/n$
1	1	0	0	0	0
2	3	67.5	0.1875	0.0352	0.0804
3	5	95	0.2639	0.0696	0.1885
4	7	150	0.4167	0.1736	0.4167
5	9	225	0.625	0.3906	0.8036
6	11	255	0.7083	0.5017	1.1130
7	13	322.5	0.8958	0.8025	1.6636
Total			3.0972	1.9732	4.2658

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i/n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 1.9732 - 4.2658 + 7 \left( \frac{1}{3} - 0.003311 \right) \\
 &= 0.0176
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.180 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.149 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución

uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

		5o. mes			
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i/n$
1	1	82.5	0.2292	0.0525	0.0382
2	3	195	0.5417	0.2934	0.2709
3	5	232.5	0.6458	0.4171	0.5382
4	7	267.5	0.7431	0.5522	0.8670
5	9	297.5	0.8264	0.6829	1.2396
6	11	345	0.9583	0.9183	1.7569
Total			3.9445	2.9164	4.7108

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i/n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 2.9164 - 4.7108 + 6 \left( \frac{1}{3} - 0.0248 \right) \\
 &= 0.0566
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.179 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.149 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Lo que quiere decir que no se rechace la hipótesis de uniformidad es que hay evidencia estadística para afirmar que las observaciones se ajustan a una distribución

uniforme, esto es, el bebé se alimenta en tiempos aleatorios, lo que nos lleva a decir que no ha establecido un horario de alimentación que abarque un mínimo número de horas.

		6o.		mcs	
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i/n$
1	1	97.5	0.2708	0.0733	0.0339
2	3	165	0.4583	0.2100	0.1719
3	5	190	0.5278	0.2786	0.3299
4	7	206.25	0.5729	0.3282	0.5013
5	9	225	0.625	0.3906	0.7031
6	11	245	0.6806	0.4632	0.9358
7	13	258.75	0.7188	0.5167	1.1681
8	15	285	0.7917	0.6268	1.4844
Total			4.6459	2.8874	5.3284

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i/n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 2.8874 - 5.3284 + 8 \left( \frac{1}{3} - 0.0065 \right) \\
 &= 0.1734
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.181 \Rightarrow U_n^2 < U(\alpha)$

∴ La hipótesis nula de uniformidad *no se rechaza*

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.150 \Rightarrow U_n^2 > U(\alpha)$

∴ La hipótesis nula de uniformidad *se rechaza*

En este mes con distintos valores de  $\alpha$ , obtenemos resultados distintos de la prueba. Recuerdese que  $\alpha$  representa la probabilidad del error tipo I, que es de rechazar la hipótesis nula, cuando ésta fuere cierta, es decir la probabilidad de que el lactante aún no haya establecido su horario de alimentación, aunque la prueba indique lo contrario es de 0.1.

		7o. mes			
$i$	$c_i$	$\phi_i$	$v_i$	$v_i^2$	$c_i v_i/n$
1	1	140	0.3889	0.1512	0.0486
2	3	155	0.4306	0.1854	0.1615
3	5	183.75	0.5104	0.2605	0.319
4	7	217.5	0.6042	0.3651	0.5287
5	9	225	0.625	0.3906	0.7031
6	11	240	0.6667	0.4445	0.9167
7	13	270	0.75	0.5625	1.2188
8	15	285	0.7917	0.6268	1.4844
Total			4.7675	2.9866	5.3808

$$\begin{aligned}
 U_n^2 &= \sum v_i^2 - \sum (c_i v_i/n) + n \left[ \frac{1}{3} - \left( \bar{v} - \frac{1}{2} \right)^2 \right] \\
 &= 2.9866 - 5.3808 + 8 \left( \frac{1}{3} - 0.0092 \right) \\
 &= 0.1986
 \end{aligned}$$

Si  $\alpha = 0.05$ ,  $U_n^2(\alpha) = 0.181 \Rightarrow U_n^2 > U(\alpha)$

$\therefore$  La hipótesis nula de uniformidad se rechaza

Si  $\alpha = 0.10$ ,  $U_n^2(\alpha) = 0.150 \Rightarrow U_n^2 > U(\alpha)$

∴ La hipótesis nula de uniformidad *se rechaza*

Ahora han cambiado los resultados que arroja la prueba, pues se tiene que se rechaza la hipótesis de uniformidad, por lo que hay evidencia estadística para afirmar que el bebé ya no se alimenta en tiempos aleatorios, lo que nos lleva a decir que ha establecido un horario de alimentación que abarque un mínimo número de horas.

Se puede comprobar que, tal como lo esperábamos, los resultados de las pruebas son similares: es hasta el sexto o séptimo mes cuando la hipótesis nula de uniformidad se rechaza; es decir, en el contexto del ejemplo, el bebé logra establecer un horario de alimentación.

# CONCLUSIONES

La teoría de la estadística circular, representa una interesante interrelación de la estadística y la geometría que permite, entre otras cosas, replantear la propia metodología estadística con base en el cuerpo axiomático del círculo, en este caso, pero pudiendo abordar otra geometría como la esfera.

Las pruebas de aleatoriedad investigadas, pueden ser aplicadas en el análisis de cualquier fenómeno rítmico, en el que el periodo sea conocido o en aquéllos que involucren datos como direcciones en el plano u horas. La afirmación anterior es importante, ya que el uso de estas pruebas ha sido muy limitado hasta ahora, con excepción de la prueba de la Ji-Cuadrada, que se usa con mucha frecuencia en contextos distintos, es decir, con datos que no son representativos de un fenómeno cíclico. La aplicación desarrollada en esta tesis resultó ser un ejemplo interesante, sencillo y nuevo del alcance de las pruebas.

El análisis matemático más completo se pudo realizar con la prueba de Rayleigh y la prueba de Greenwood y Durand, que son la más usadas en la actualidad. Sin embargo, se comprobó que los resultados de la aplicación fueron muy parecidos en todas las pruebas.

Mediante el ejercicio de la aplicación de las pruebas, se pueden hacer las siguientes observaciones:

a) Como se señaló antes, la prueba de la Ji-Cuadrada es la más conocida y aplicada de las pruebas, aunque se pierda información al agrupar los datos, especialmente en muestras pequeñas. En el ejemplo estudiado se tenían pocos datos, por lo que no fue posible aplicarla.

b) La mayor ventaja de la prueba del Rango y la de Hodges y Ajne es la rapidez para la obtención de resultados, pues no se requieren de cálculos. La prueba de espacios de Rao requiere de más cálculos, pero son muy sencillos.

c) La prueba de Kuiper puede ser aplicada mediante algunos cálculos, pero también gráficamente.

No se pretende argumentar que la facilidad que nos ofrece la aplicación de las pruebas represente alguna ventaja frente a las demás, se menciona solo como una característica que no se quiso despreciar. La principal ventaja de las pruebas de aleatoriedad en estadística circular es, que en su conjunto, sirven como un método alternativo y complementario para el estudio de los fenómenos cíclicos y direccionales.

Debe resaltarse que el principal aporte de esta tesis es que se pudo abordar el estudio de dichos fenómenos de una manera distinta a la usual, rescatando la propia naturaleza de las observaciones, ya que su característica de ser periódicas les confiere un análisis especial.

Esta tesis es un trabajo motivacional, necesario para contribuir a rescatar estas ideas, que aunque en un principio no poseían todo el avance, tanto de formalismo matemático como en el campo de aplicación, ahora pueden servir para la propia ge-



# APÉNDICE

Tabla A Prueba de Rayleigh

r	n=5	6	7	8	9	10	11	12	13	14	15	16	17
.10		P > 0.900											
.12			.897	.884	.872	.859	.847	.835	.823	.811	.799	.788	
.14		.896	.880	.863	.846	.829	.813	.797	.782	.766	.752	.737	.723
.16	.894	.867	.845	.824	.803	.783	.763	.744	.725	.706	.688	.671	.654
.18	.865	.836	.808	.783	.757	.733	.710	.687	.665	.643	.623	.603	.583
.20	.833	.801	.768	.739	.709	.681	.654	.628	.603	.580	.557	.535	.513
.22	.800	.765	.727	.692	.659	.628	.598	.569	.542	.516	.492	.468	.446
.24	.764	.727	.684	.645	.609	.574	.541	.511	.482	.454	.429	.404	.381
.26	.729	.687	.640	.597	.558	.520	.486	.454	.423	.395	.369	.345	.322
.28	.693	.645	.595	.549	.507	.468	.432	.399	.368	.340	.314	.290	.267
.30	.656	.603	.550	.501	.457	.417	.380	.347	.316	.289	.263	.241	.219
.32	.620	.560	.506	.455	.409	.369	.332	.299	.269	.242	.218	.196	.177
.34	.584	.518	.462	.410	.364	.323	.287	.255	.226	.201	.178	.158	.140
.36	.547	.477	.419	.367	.321	.280	.245	.214	.188	.164	.143	.125	.110
.38	.511	.437	.378	.326	.280	.241	.208	.179	.154	.132	.114	.098	.084
.40	.475	.399	.339	.287	.243	.205	.174	.147	.125	.105	.089	.075	.064
.42	.440	.362	.301	.251	.209	.173	.144	.120	.100	.083	.069	.057	.048
.44	.405	.327	.265	.218	.178	.145	.118	.096	.079	.064	.052	.043	.035
.46	.371	.294	.234	.187	.150	.120	.096	.077	.061	.049	.039	.031	.025
.48	.338	.262	.204	.160	.125	.098	.077	.060	.047	.037	.029	.022	.018
.50	.305	.233	.177	.135	.104	.079	.061	.046	.035	.027	.021	.016	.012
.52	.274	.205	.152	.114	.085	.063	.047	.035	.026	.020	.015	.011	.008
.54	.244	.179	.130	.095	.069	.050	.037	.027	.019	.014	.010	.007	.005
.56	.215	.155	.110	.078	.055	.039	.028	.020	.014	.010	.007	.005	.004
.58	.189	.133	.092	.064	.044	.030	.021	.014	.010	.007	.005	.003	.002
.60	.165	.113	.076	.051	.034	.023	.015	.010	.007	.005	.003	.002	.001
.62	.144	.096	.063	.041	.027	.017	.011	.007	.005	.003	.002	.001	
.64	.126	.080	.051	.032	.020	.013	.008	.005	.003	.002	.001		
.66	.109	.067	.041	.025	.015	.009	.006	.003	.002	.001			
.68	.095	.055	.032	.019	.011	.007	.004	.002	.001				
.70	.082	.045	.025	.014	.008	.005	.003	.001					
.72	.070	.037	.020	.011	.006	.003	.002						
.74	.059	.030	.015	.008	.004	.002	.001						
.76	.049	.024	.012	.006	.003	.001							
.78	.040	.019	.009	.004	.002								
.80	.033	.014	.006	.003	.001								
.82	.026	.010	.004	.002									
.84	.020	.008	.003	.001									
.86	.015	.006	.002										
.88	.011	.004	.001										

P < 0.001

n= tamaño de la muestra, r=estadística de prueba, P=nivel crítico

Tabla B Prueba V

n	$\alpha=0.10$	0.05	0.01	0.005	0.001	0.0001
5	$u(\alpha)=1.3051$	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	3.3578
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	3.0380	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	3.6184
27	1.2856	1.6460	2.3132	2.5538	3.0421	3.6223
28	1.2855	1.6460	2.3136	2.5546	3.0439	3.6258
29	1.2853	1.6459	2.3141	2.5553	3.0455	3.6292
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	2.5752	3.0890	3.7165

n=tamaño de la muestra,  $\alpha$ =nivel de significancia,  $u(\alpha)$ =valor crítico de la estadística de prueba

Tabla C Prueba del Rango

n	$\alpha=0.005$	0.01	0.025	0.05	0.10
4	w( $\alpha$ )= 38.8	48.9	66.3	83.6	105.3
5	64.0	76.1	95.7	113.8	135.4
6	87.2	100.2	120.3	138.2	158.7
7	107.6	120.8	140.8	158.0	177.3
8	125.5	138.5	157.9	174.4	192.5
9	141.1	153.8	172.5	188.1	205.1
10	154.7	167.1	185.0	199.8	215.8
11	166.8	178.7	195.9	209.9	225.0
12	177.4	189.0	205.4	218.7	233.0
13	187.0	198.1	213.8	226.5	240.0
14	195.5	206.2	221.3	233.4	246.2
15	203.2	213.5	228.0	239.5	251.7
16	210.2	220.1	234.0	245.1	256.7
17	216.6	226.2	239.5	250.1	261.2
18	222.4	231.6	244.5	254.7	265.3
19	227.7	236.7	249.0	258.8	269.1
20	232.7	241.3	253.2	262.7	272.5
21	237.2	245.6	257.1	266.2	275.6
22	241.4	249.5	260.7	269.4	278.6
23	245.4	253.2	264.0	272.5	281.3
24	249.0	256.7	267.1	275.3	283.8
25	252.5	259.9	270.0	277.9	286.1
26	255.7	262.9	272.7	280.4	288.3
27	258.7	265.7	275.2	282.7	290.4
28	261.5	268.3	277.6	284.8	292.3
29	264.2	270.8	279.8	286.8	294.1
30	266.7	273.2	281.9	288.8	295.9

n=tamaño de la muestra,  $\alpha$ = nivel de significancia, w( $\alpha$ )=valor crítico de la estadística de prueba (grados)

Tabla D Prueba de espacios de Rao

n	$\alpha=0.01$	0.05	0.10
4	$U(\alpha) = 221.0$	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	185.8	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

n=tamaño de la muestra,  $\alpha$ =nivel de significancia,  $U(\alpha)$ =valor crítico de la estadística de prueba

Tabla E Prueba de Hodges y Ajne

n	K=0	1	2	3	4	5	6	7	8	9	10	11	12
5	P=.313												
6	.188	750											
7	.109	.547											
8	.063	.375	875										
9	.035	.246	.703										
10	.020	.156	.527										
11	.011	.097	.376	806									
12	.006	.059	.258	.645									
13	.003	.035	.171	.489	873				P >	0.900			
14	.002	.021	.111	.355	.733								
15		.012	.070	.250	.583								
16		.007	.044	.171	.444	800							
17		.004	.027	.114	.327	.661							
18		.002	.016	.075	.233	.523	850						
19		.001	.010	.048	.163	.399	.725						
20			.006	.030	.111	.296	.591	.887					
21			.003	.019	.074	.213	.466	.776					
22			.002	.012	.049	.151	.356	.651					
23			.001	.007	.032	.104	.265	.526	.818				
24				.004	.020	.071	.193	.413	.701				
25				.003	.013	.048	.137	.315	.580	852			
26				.002	.008	.031	.096	.235	.466	.745			
27					.005	.020	.066	.172	.364	.629	880		
28					.003	.013	.045	.124	.278	.515	.782		
29					.002	.008	.030	.087	.208	.410	.672		
30					.001	.005	.020	.061	.153	.320	.560	814	
31						.003	.013	.042	.110	.244	.454	.710	
32						.002	.008	.028	.078	.183	.360	.601	841
33						.001	.005	.019	.055	.135	.280	.496	.744
34							.003	.013	.038	.098	.214	.400	.638
35			P <	0.001			.002	.008	.026	.070	.160	.316	.534
36							.001	.005	.018	.049	.118	.245	.437
37								.003	.012	.034	.086	.187	.350
38								.002	.008	.024	.062	.140	.276
39								.001	.005	.016	.044	.104	.213
40									.003	.011	.031	.076	.163

n=tamaño de la muestra, K=estadística de prueba, P=nivel crítico

Tabla F Prueba de la Ji Cuadrada

$\chi^2$	v=1	2	3	4	5	6	7	8	9	10	11	12
1	P= 317	607	.801									
2	157	368	.572	736	849					P >	0 900	
3	083	.223	.392	.558	700	809	885					
4	046	135	261	408	549	677	780	857				
5	025	082	172	287	416	544	660	.758	834	891		
6	.014	050	112	199	.306	423	540	.647	.740	815	873	
7	008	.030	.072	.136	221	321	429	537	637	725	799	858
8	005	.018	.046	.092	.156	238	333	433	.534	629	713	785
9	003	011	029	061	109	174	253	342	437	532	622	703
10	.002	007	019	040	075	125	189	.265	350	440	530	616
11		.004	.012	.027	.051	.088	.139	202	.276	358	.443	529
12		002	.007	017	035	062	101	151	213	285	364	446
13		001	005	.011	023	043	072	112	.163	224	293	369
14			.003	.007	016	030	051	082	122	173	233	301
15			002	005	010	020	036	.059	091	132	183	241
16			001	003	007	014	025	.042	067	100	141	191
17				002	005	.009	017	.030	049	074	108	150
18				001	003	006	012	021	035	055	082	116
19					.002	004	.008	015	025	040	061	089
20					.001	.003	.006	010	.018	029	045	067
21						002	004	007	013	.021	033	050
22						001	003	005	009	015	024	038
23							.002	.003	.006	011	.018	028
24								002	004	008	013	020
25		P <	0 001					.002	003	005	009	015
26								001	.002	004	006	011
27									.001	003	005	008
28										002	.003	006
29										001	002	004
30											002	003

v=grados de libertad,  $\chi^2$  =estadística de prueba, P=nivel crítico

Tabla G Prueba de Kuiper

n	$\alpha=0.50$	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	$K(\alpha)=1.102$	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	2.007
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	2.125
12	1.138	1.378	1.517	1.637	1.776	1.870	1.957	2.062	2.137
13	1.141	1.382	1.522	1.642	1.782	1.876	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	1.825	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	2.230
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

n=tamaño de la muestra,  $\alpha$ =nivel de significancia,  $K(\alpha)$ =valor crítico de la estadística de prueba

Tabla H Prueba de Watson

N	$\alpha=0.10$	0.05	0.025	0.01	0.005
2	$U_n^2(\alpha) = 143$	.155	161	.164	.165
3	.145	.173	194	.213	.224
4	.146	.176	202	.233	.252
5	.148	.177	205	.238	.262
6	.149	.179	.208	.243	.269
7	.149	.180	.210	.247	.274
8	.150	.181	.211	.250	.278
9	.150	.182	.212	.252	.281
10	.150	.182	.213	.254	.283
12	.150	.183	.215	.256	.287
14	.151	.184	.216	.258	.290
16	.151	.184	.216	.259	.291
18	.151	.184	.217	.259	.292
20	.151	.185	.217	.261	.293
30	.152	.185	.219	.263	.296
40	.152	.186	.219	.264	.298
50	.152	.186	.220	.265	.299
100	.152	.186	.221	.266	.301
$\infty$	.152	.187	.221	.267	.302

n=tamaño de la muestra,  $\alpha$ =nivel de significancia,  $U_n^2(\alpha)$  = valor crítico de la estadística de prueba

## BIBLIOGRAFIA

- [1] Batschelet, E., *Circular Statistics in Biology*, Academic Press, London and New York, 1981, [1,4]
- [2] Bowman, F., *Introduction to Bessel Functions*, Dover Publications Inc., New York, 1985.
- [3] Lehmann, E.L., *Testing Statistical Hypotheses*, Wiley, London and New York, 1959, [4.1].
- [4] Mardia, K.V., *Statistics of Directional Data*, Academic Press, London and New York, 1972, . [5 4, 7].
- [5] Petiau, G., *La Théorie des fonctions de Bessel*, Centre National de la Recherche Scientifique, Paris, 1953, (329-335).
- [6] Sherman, B., *A random variable related to the spacing of sample values*, Ann. Math. Statist. **21** 339-361.
- [7] Stephens, M.A., *Test based on EDF Statistics*, Simon Fraser University, Burnaby, B. C., Canadá, 1969.
- [8] *Libro para el maestro, Matemáticas*, SEP, 1999, p. 306.