



**UEPB**

**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE BACHARELADO EM ESTATÍSTICA**

**DÉBORA DOS SANTOS FARIAS**

**UTILIZANDO ESTATÍSTICA CIRCULAR PARA OS CASOS DE DENGUE NA  
CIDADE DE CAMPINA GRANDE - PB**

**CAMPINA GRANDE - PB  
2023**

DÉBORA DOS SANTOS FARIAS

**UTILIZANDO ESTATÍSTICA CIRCULAR PARA OS CASOS DE DENGUE NA  
CIDADE DE CAMPINA GRANDE - PB**

Trabalho de Conclusão de Curso (Artigo) apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

**Orientadora:** Profa. Dra. Divanilda Maia Esteves

**CAMPINA GRANDE - PB  
2023**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

F224u Farias, Débora dos Santos.  
Utilizando Estatística Circular para os casos de dengue na cidade de Campina Grande - PB [manuscrito] / Débora dos Santos Farias. - 2023.  
47 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2023.

"Orientação : Profa. Dra. Divanilda Maia Esteves, Coordenação do Curso de Estatística - CCT. "

1. Estatística Circular. 2. Estatística Direcional. 3. Dados circulares. 4. Dengue. I. Título

21. ed. CDD 519.5

DÉBORA DOS SANTOS FARIAS

UTILIZANDO ESTATÍSTICA CIRCULAR PARA OS CASOS DE DENGUE NA CIDADE  
DE CAMPINA GRANDE - PB

Trabalho de Conclusão de Curso (Artigo) apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

Trabalho aprovado em 17 de Março de 2023.

**BANCA EXAMINADORA**

*DM Esteves*

---

Profa. Dra. Divanilda Maia Esteves  
(Orientadora)  
Universidade Estadual da Paraíba (UEPB)

*Ana Patricia Bastos Peixoto*

---

Profa. Dra. Ana Patricia Bastos Peixoto  
Universidade Estadual da Paraíba (UEPB)

*maria das vitórias A. serafim*

---

Profa. Me. Maria das Vitórias Alexandre  
Serafim  
Universidade Estadual da Paraíba (UEPB)

Dedico esse trabalho de pesquisa ao meu esposo Antonio, a minha mãe Dinalva, a meu avô Domingos (*in memoriam*) e as minhas irmãs Bruna, Sabrina, Aline e Sara.

“Ora, àquele que é poderoso para  
fazer tudo muito mais abundantemente  
além daquilo que pedimos ou pensamos,  
segundo o poder que em nós opera.”  
(Efésios 3:20)

## LISTA DE ILUSTRAÇÕES

Figura 1	–	Representação do ponto P em uma unidade circular. . . . .	14
Figura 2	–	Gráfico de dispersão para dados circulares. . . . .	15
Figura 3	–	Representação gráfica do histograma. . . . .	16
Figura 4	–	Representação gráfica do diagrama de rosa. . . . .	16
Figura 5	–	Representação gráfica da média aritmética em dados circulares. . . . .	17
Figura 6	–	Representação gráfica da direção média circular em dados circulares. . . . .	18
Figura 7	–	Gráfico da Função de Densidade de Distribuição Uniforme. . . . .	22
Figura 8	–	Gráfico da Função de Densidade de Distribuição de Von Mises com parâmetros $\mu = \pi$ e $k = 0,5$ . . . . .	23
Figura 9	–	Gráfico da Função de Densidade de Distribuição de Von Mises com parâmetros $\mu = \pi$ e $k = 5$ . . . . .	23
Figura 10	–	Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2018. . . . .	26
Figura 11	–	Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2019. . . . .	27
Figura 12	–	Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2020. . . . .	27
Figura 13	–	Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2021. . . . .	28
Figura 14	–	Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2018. . . . .	29
Figura 15	–	Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2019. . . . .	29
Figura 16	–	Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2020. . . . .	30
Figura 17	–	Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2021. . . . .	31

## LISTA DE TABELAS

Tabela 1 – Dados circulares de casos prováveis de dengue na cidade de Campina Grande	
- PB . . . . .	28
Tabela 2 – Teste de Uniformidade para os casos de dengue na cidade de Campina Grande	
- PB . . . . .	31

## **LISTA DE ABREVIATURAS E SIGLAS**

SINAN	Sistema de Informações de Agravos de Notificação
TABNET	Tabulador de informações de saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
AVC	Acidente Vascular Cerebral
PB	Estado da Paraíba

## LISTA DE SÍMBOLOS

$\mu$	Letra grega minúscula Mi
$\pi$	Letra grega minúscula Pi
$\theta$	Letra grega minúscula Teta
cos	Símbolo do cosseno
sin	Símbolo do seno
tan	Símbolo do tangente
$\Sigma$	Letra grega maiúscula Sigma
$\phi$	Letra grega Fi
log	Símbolo de logaritmo
$\delta$	Letra grega minúscula Delta
$\int$	Símbolo de integral
$\infty$	Símbolo de infinito

## SUMÁRIO

1	INTRODUÇÃO . . . . .	11
2	FUNDAMENTAÇÃO TEÓRICA . . . . .	12
2.1	Conceitos Básicos da Estatística Circular . . . . .	13
2.2	Representação Gráfica . . . . .	15
2.2.1	<i>Gráfico de dispersão</i> . . . . .	15
2.2.2	<i>Histogramas Circulares</i> . . . . .	15
2.2.3	<i>Diagrama de Rosa</i> . . . . .	16
2.3	Estatística Descritiva . . . . .	16
2.3.1	<i>Medidas de Posição</i> . . . . .	17
2.3.1.1	<i>Média Circular</i> . . . . .	17
2.3.1.2	<i>Mediana Circular</i> . . . . .	18
2.3.2	<i>Medidas de Concentração</i> . . . . .	18
2.3.2.1	<i>Comprimento Médio do Vetor Médio Resultante</i> . . . . .	18
2.3.3	<i>Medidas de Dispersão</i> . . . . .	19
2.3.3.1	<i>Variância Circular</i> . . . . .	19
2.3.3.2	<i>Desvio Padrão Circular</i> . . . . .	19
2.3.4	<i>Momentos Trigonométricos</i> . . . . .	19
2.4	Inferência . . . . .	21
2.4.1	<i>Distribuição Uniforme</i> . . . . .	22
2.4.2	<i>Distribuição Von Mises</i> . . . . .	22
2.5	Teste de Hipótese . . . . .	23
2.5.1	<i>Teste de Rayleigh</i> . . . . .	23
2.5.2	<i>Teste de Watson <math>U_n^2</math></i> . . . . .	24
2.5.3	<i>Teste de espaçamento de Rao</i> . . . . .	24
3	MATERIAL E MÉTODOS . . . . .	24
3.1	Base de dados e área do estudo . . . . .	24
3.2	Dados epidemiológicos . . . . .	25
3.3	Métodos e Pacotes Utilizados . . . . .	25
4	RESULTADOS E DISCUSSÃO . . . . .	26
5	CONCLUSÃO . . . . .	32
	REFERÊNCIAS . . . . .	33
	APÊNDICE A – DEMONSTRAÇÃO DAS EQUAÇÕES (15) E (16) . . . . .	34
A.1	Demonstração . . . . .	34
	APÊNDICE B – ROTINA NO R . . . . .	36

## UTILIZANDO ESTATÍSTICA CIRCULAR PARA OS CASOS DE DENGUE NA CIDADE DE CAMPINA GRANDE - PB

Débora dos Santos Farias\*  
Divanilda Maia Esteves†

### RESUMO

A Estatística Circular é um ramo particular da Estatística, o qual lida com dados que podem ser representados como pontos na circunferência do círculo unitário. O diferencial nesta abordagem é o fato de respeitar a natureza dos dados e fornecer técnicas e gráficos que destacam as características intrínsecas de variáveis direcionais. A principal propriedade dos dados circulares é o fato de apresentarem características periódicas, ou seja, não há um verdadeiro ponto inicial ou final, mas sim uma escala contínua que retorna em si. Por isso, os dados de variáveis circulares são representados por ângulos. A partir daí, são obtidos os parâmetros da Estatística Circular através da análise da estatística descritiva, da inferência e dos testes de hipóteses. Além dos dados direcionais se adequarem naturalmente a uma representação no círculo, qualquer fenômeno periódico com um período conhecido (dia, mês ou ano), pode ser representado no círculo onde a circunferência corresponde a esse período, agregando conforme necessário. Os dados utilizados neste trabalho se referem ao número de casos de dengue notificados mensalmente no município de Campina Grande e foram disponibilizados pelo SINAN por meio do TABENET<sup>1</sup> (TABENET, 2023). O período considerado vai de janeiro de 2018 a dezembro de 2021. Os procedimentos estatísticos foram executados no software R (R Core Team, 2021). Para as análises foram usados os pacotes *dplyr* e *Circular*. O objetivo do presente trabalho foi verificar a aplicabilidade da Estatística Circular para os casos de dengue. Para tal, foram utilizadas ferramentas gráficas, medidas de variabilidade e medidas descritivas circulares. Por fim, realizou-se o teste estatístico para verificar as observações por toda a circunferência, onde concluiu-se que a existência de casos prováveis de dengue na cidade de Campina Grande não se distribuem uniformemente durante os meses nos anos em estudo.

**Palavras-chaves:** Estatística Circular. Estatística Direcional. Dados Circulares. Dengue.

### ABSTRACT

Circular Statistics is a particular branch of Statistics which deals with data that can be represented as points on the circumference of the unit circle. The difference in this approach is the fact that it respects the nature of the data and provides techniques and graphs that highlight the intrinsic characteristics of directional variables. The main property of circular data is the fact that they have periodic characteristics, that is, there is no real starting or ending point, but a continuous scale that returns in itself. Therefore, data for circular variables are represented by angles. From there, the parameters of Circular Statistics are obtained through the analysis of descriptive statistics, inference and hypothesis tests. In addition to directional data naturally suited to a representation on the circle, any periodic phenomenon with a known period (day, month or year), can be represented on the circle where the circumference corresponds to that period, aggregating as necessary. The data used in this work refer to the number of cases of dengue

\* Aluna Débora dos Santos Farias do curso de Bacharelado de Estatística, Depto de Estatística, UEPB, Campina Grande, PB, debora.santos.farias@aluno.uepb.edu.br

† Profa. Dra. Divanilda Maia Esteves, Depto de Estatística, UEPB, Campina Grande, PB, di-ana.maia@servidor.uepb.edu.br

<sup>1</sup> Tabulador de domínio público desenvolvido pelo DATASUS que permite organizar dados para gerar informações a partir das bases de dados do sistema único de saúde, nele é possível gerar tabelas e produzir gráficos e mapas. links: <<http://tabnet.datasus.gov.br/cgi/defthtm.exe?sinannet/cnv/denguebpb.def>>

reported monthly in the city of Campina Grande and were made available by SINAN through TABENET (TABENET, 2023). The period considered ranges from January 2018 to December 2021. The statistical procedures were performed in the R software. For the analyzes, the *dplyr* and *Circular* packages were used. The aim of this study was to verify the applicability of Circular Statistics for dengue cases. For this, graphic tools, variability measures and circular descriptive measures were used. Finally, the statistical test was carried out to verify the observations throughout the circumference, where it was concluded that the existence of probable cases of dengue in the city of Campina Grande are not evenly distributed during the months in the years under study.

**Keywords:** Circular Statistics. Directional Statistics. Circular Data. Dengue.

## 1 INTRODUÇÃO

No século IV o Monge Macrobius iniciou as primeiras pesquisas usando métodos estatísticos para saber a inclinação das órbitas planetárias em função do tempo, comprovando que a Estatística Circular nasceu através de estudos voltados pela astronomia. Por volta de 1510 surgem as primeiras medições registradas de declinação magnética e em 1710, Sir Edmund Halley, publicou um mapa que mostrava as linhas de declinação magnética para todas as águas navegadas do mundo. Pelo menos em meados do século XVII surgiu as primeiras raízes da análise circular. O reverendo John Mitchell FRS em 1767 analisou separações angulares entre estrelas e descobriu que o número de pares próximos era muito grande para ser consistente com a hipótese de que as direções das estrelas estavam uniformemente distribuídas (FISHER, 1995).

Em 1858 Nightingale<sup>2</sup> usou uma variedade de métodos para argumentar seu caso, um deles era um gráfico que foi chamado de *diagrama de rosa*, ele e outros gráficos foram usados para ressaltar melhorias no saneamento sanitário. A partir das análises de dados circulares em Astronomia e Medicina, surgiram aplicações em várias áreas, tal como Biologia, Geografia, Geologia, Meteorologia e Oceanografia (FISHER, 1995). No século XIX, Lord Rayleigh (John William Strutt, Terceiro Barão Rayleigh) foi um dos pioneiros no estudo de modelagem de dados circulares, onde desenvolveu um teste estatístico formal para uniformidade contra um modelo alternativo de uma única direção preferida (RAYLEIGH, 1880).

Entre 1940 a 1942 Salvemini desenvolveu a variância circular para estudar os desvios da uniformidade das mortes mensais em Roma para vários grupos de idade. No entanto no início do século XX, surgiram outras distribuições, como a Von Mises (MISES, 1918) e a Wrapped normal, pois até aquele momento os estudos estavam concentrados na distribuição Uniforme.

A partir do paper de Fisher em 1953, surgiram os maiores avanços em análise de dados circulares (FISHER, 1953). Em 1950, os estatísticos começaram a ajustar distribuições circulares de probabilidade aos dados. E E.J. Gumbel investigou a adequação da distribuição Von Mises para a modelagem de conjuntos de dados. Depois disso, vieram vários estudos importantes para o desenvolvimento da teoria de análise estatística de dados circulares, como (WATSON; WILLIAMS, 1956), no qual foram introduzidos métodos de inferência sobre a direção média e dispersão para uma única amostra de dados circulares, e métodos de comparação entre duas ou mais amostras. Nos anos 60, surgiram outros nomes importantes para o desenvolvimento deste assunto, tais como G. S. Watson, M. A. Stephens e K. W. Mardia. (FISHER, 1995).

Consequentemente nas últimas décadas é notório que a análise de dados circulares vem

<sup>2</sup> Foi enfermeira-chefe do exército britânico durante a guerra da Crimeia e aprendera que milhares de vidas poderiam ser salvas simplesmente melhorando as condições sanitárias nos hospitais e quartéis militares. Além de promover a própria causa da reforma médica, ela ajudou a criar a noção revolucionária de que os fenômenos sociais podiam ser medidos objetivamente e submetidos à análise matemática.

crecendo como uma importante área da Estatística. Várias novas técnicas surgiram, principalmente para atender às necessidades de muitas áreas do conhecimento, tais como engenharia, medicina, geologia, biologia, geografia, meteorologia, psicologia, oceanografia e matemática, dentre outros.

As direções bidimensionais podem ser representadas como ângulos medidos com respeito a um ponto de partida e um sentido horário ou anti-horário. Como uma direção não tem magnitude, ela pode ser convenientemente representada como pontos na circunferência de um círculo unitário centrado em uma origem ou como vetores unitários conectando a origem a esses pontos. Por motivo desta representação circular, observações em tais direções bidimensionais também são chamados de dados circulares. De forma semelhante os dados direcionais em três dimensões também são referidos como dados esféricos (JAMMALAMADAKA; SENGUPTA, 2001).

Além dos dados direcionais (direção de migrações de animais, direção do vento, orientação da fratura de uma rocha, transporte de materiais geológicos, direção de correntes marinhas, etc.) se adequarem naturalmente a uma representação no círculo, fenômenos periódicos, com período conhecido, que também podem ser representados em uma circunferência, no qual o comprimento do círculo condiz ao período (NEVES, 2015). Assim, pode-se interessar em estudar dados circulares, por exemplo, em chegada de pacientes com COVID-19, com leucemia ou acidente vascular cerebral (AVC) em um hospital em certo período de tempo. Outro exemplo seria utilizar o círculo para representar os 365 dias do ano e fazer uma análise da ocorrência de acidentes de carros para avaliar se os acidentes são uniformemente distribuídos semanalmente.

Para trabalhar com dados circulares, deve-se utilizar uma teoria estatística específica e apropriada, diferente dos métodos convencionais. A razão desta afirmação está no fato de que, se medidas angulares fossem tratadas como lineares, seriam obtidos resultados errôneos. A Estatística usual apresenta uma distribuição linear, em que as variáveis estão em torno de uma reta numérica. No entanto, na Estatística Circular, os dados pontuais são ângulos, distribuídos geralmente em graus ou radianos, que substitui a reta numérica.

Um exemplo seria a hora do dia representada no relógio. À medida que o tempo passa, os ponteiros do relógio inevitavelmente passarão as mesmas 24 horas repetidamente. Para dados localizados na linha real, um valor de 360 é relativamente próximo de um valor de 355, mas relativamente distante da origem. Para variáveis circulares a situação é muito diferente. Enquanto um ângulo de  $355^\circ$  corresponde a um ponto da circunferência do círculo unitário que se aproxima do correspondente a  $360^\circ$ , os ângulos de  $0^\circ$  e  $360^\circ$  definem exatamente o mesmo ponto.

Sendo assim, a presente pesquisa tem como objetivo verificar a aplicabilidade da estatística circular para os casos de dengue no município de Campina Grande - PB e assim realizar um aprofundamento das técnicas utilizadas. Em vista disso deve-se identificar os parâmetros que regem a análise circular dos diversos eventos relacionados à direção ou sentido de fenômenos do cotidiano.

O artigo está estruturado em 5 seções, a primeira como uma breve revisão, na qual expõe uma ideia básica sobre o uso dos dados circulares. Em seguida a fundamentação teórica que foram abordadas os conceitos básicos da Estatística Circular, como também a representação gráfica e apresentação da estatística descritiva, distribuições circulares e testes de hipóteses. Da metodologia foi descrito a base de dados e área do estudo, com uma aplicação dos métodos circulares nos dados coletados para os casos de dengue, da análise de resultados e da conclusão.

## **2 FUNDAMENTAÇÃO TEÓRICA**

A Estatística é uma ciência que fornece métodos e princípios para coleta, organização e análise de dados oriundos de estudos ou experimentos, realizados nas mais diversas áreas

do conhecimento, com intuito de obtenção de resultados relevantes e assumindo um papel importante no que se refere à tomada de decisão em um acontecimento ou em situações de incertezas (MAGALHÃES; LIMA, 2015).

Quando as variáveis são circulares a Estatística tradicional não deve ser utilizada, desta forma, é fundamental um estudo relacionado a Estatística Circular. Na Estatística Circular cada observação pode ser especificada pelo ângulo formado entre a direção inicial do círculo e o ponto no círculo correspondente à observação, visto que uma direção inicial e uma orientação do círculo sejam determinadas (MARDIA; JUPP, 2000).

## 2.1 Conceitos Básicos da Estatística Circular

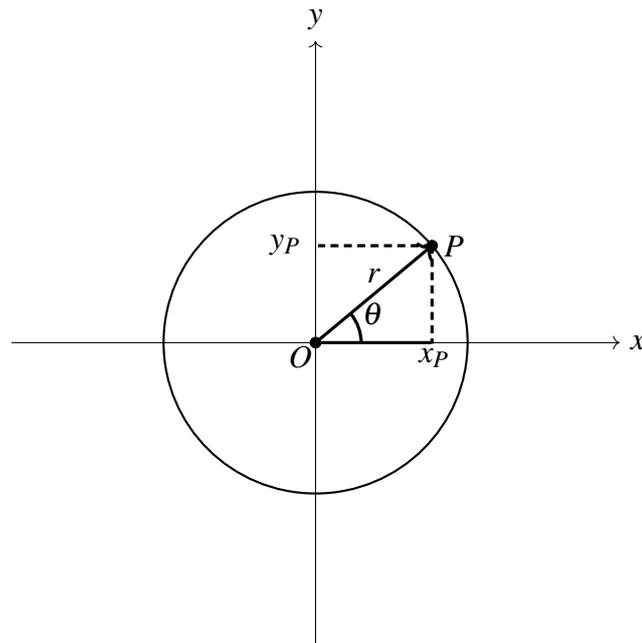
A Estatística Circular lida com dados que podem ser representados como pontos na circunferência do círculo unitário, onde cada observação de uma variável circular tem sua medida entre  $0^\circ$  e  $360^\circ$  ou de maneira análoga de 0 a  $2\pi$ .

A principal propriedade dos dados circulares é o fato de serem periódicos, ou seja, não há um verdadeiro ponto inicial ou final, mas sim uma escala contínua que retorna em si, ou seja  $0 = 2\pi$ . Geralmente os dados circulares são medidos em graus. Entretanto, por vezes é necessário medir em radianos. Lembrando que as medidas angulares são convertidas de graus para radianos multiplicando por  $\pi/180$  (MARDIA; JUPP, 2000).

Como as técnicas de análises para dados lineares mostram-se inapropriadas para o tratamento de dados circulares, existem técnicas estatísticas específicas para tais tipos de dados, sendo necessárias definições de medidas de tendência central e medidas de variabilidade, assim como modelos probabilísticos adequados que são usualmente adotados pela Estatística Circular.

Seja um ponto  $P \neq 0$ , onde no plano pode ser representado como  $(x_i, y_i)$  em termos de coordenadas cartesianas, no qual  $x_i$  é a projeção de  $P$  no eixo  $x$  e  $y_i$ , a projeção no eixo  $y$ . Pode-se descrever também esse ponto como coordenadas polares  $(r, \theta)$ , em que  $r$  é a distância do ponto até a origem e  $\theta$  o ângulo tomado no sentido anti-horário, da parte positiva do eixo  $Ox$  ao segmento  $OP$ , conforme está representado na Figura 1.

Figura 1 – Representação do ponto P em uma unidade circular.



Fonte: Elaborado pelo o autor, 2022.

Essa conversão entre o sistema de coordenadas polares com o sistema de coordenadas cartesianas é feito por meio das funções trigonométricas seno e cosseno (JAMMALAMADAKA; SENGUPTA, 2001), que são dadas por

$$x = r \cos \theta, \quad y = r \sin \theta \quad (1)$$

Para obter a conversão inversa com a Equação 1 utiliza-se:

$$x^2 = r^2 \cos^2 \theta, \quad (2)$$

$$y^2 = r^2 \sin^2 \theta, \quad (3)$$

$$\cos \theta = \frac{x}{r}, \quad (4)$$

$$\sin \theta = \frac{y}{r}, \quad (5)$$

$$\frac{y}{x} = \frac{\sin \theta}{\cos \theta} = \tan \theta. \quad (6)$$

Como o interesse é apenas na direção e não na magnitude do vetor, considera esses vetores como unidades de comprimento ( $r = 1$ ). (JAMMALAMADAKA; SENGUPTA, 2001). Então sabe-se que,

$$r = \sqrt{x^2 + y^2}, \quad (7)$$

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}} \quad (8)$$

$$\sin \theta = \frac{y}{\sqrt{x^2 + y^2}} \quad (9)$$

$$\tan \theta = \frac{y}{x}. \quad (10)$$

portanto, basta observar que,

$$x^2 + y^2 = r^2(\cos^2 \theta + \sin^2 \theta) = r^2, \quad (11)$$

o que implica  $r = |r| = \sqrt{x^2 + y^2}$ , pois  $r \geq 0$ .

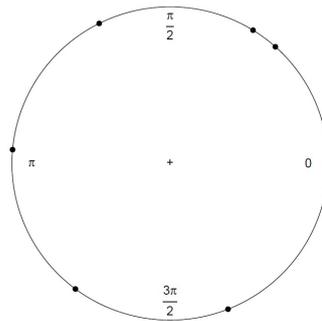
## 2.2 Representação Gráfica

Com o intuito de facilitar a análise e a interpretação, os dados circulares assim como os dados lineares podem ser representados por diagramas e gráficos.

### 2.2.1 Gráfico de dispersão

Geralmente utiliza-se o gráfico de dispersão no círculo, em que cada observação é representada por um ponto. Conforme está representado graficamente na Figura 2.

Figura 2 – Gráfico de dispersão para dados circulares.

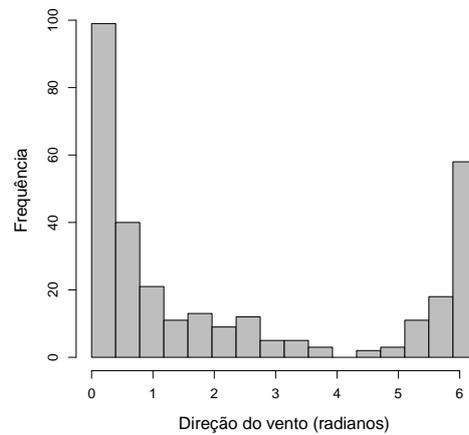


Fonte: Elaborado pelo o autor, 2022.

### 2.2.2 Histogramas Circulares

Os histogramas circulares são formados por um conjunto de barras que representam as frequências de cada classe, cujas bases estão sobre a circunferência do círculo unitário dividido de acordo com o tamanho de cada classe, com áreas proporcionais às frequências (FISHER, 1995) (Figura 3).

Figura 3 – Representação gráfica do histograma.

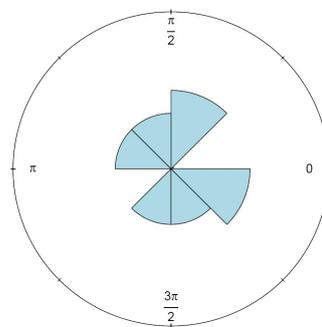


Fonte: Elaborado pelo o autor, 2022.

### 2.2.3 Diagrama de Rosa

O Diagrama de Rosa compreende o gráfico mais usual na representação de dados circulares, produzido por Florence Nightingale (1820-1910), foi um importante divisor de águas para a enfermagem. Em sua construção, os raios são proporcionais às frequências de cada classe e desta forma, há uma dilatação entre a origem e às barreiras dos intervalos (Figura 4).

Figura 4 – Representação gráfica do diagrama de rosa.



Fonte: Elaborado pelo o autor, 2022.

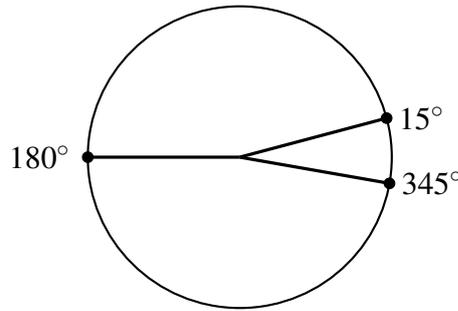
## 2.3 Estatística Descritiva

Na média circular não é possível pegar a soma de todos os valores e dividi-los pelo número de observações. Para definir uma direção média ou um vetor médio para um conjunto de dados circulares unimodais específico pode ser tentador calcular a média aritmética dos ângulos, porém não é correto utilizá-la. A média, como também o desvio padrão e outras medidas descritivas sofrem com sua forte dependência na escolha da direção zero e do sentido de rotação. Para ilustrar o cálculo médio do ângulo em dados circulares, suponha duas direções, dados pelos ângulos  $15^\circ$  e  $345^\circ$  que representa o voo<sup>3</sup> dos pássaros e deseja saber a direção média. De acordo

<sup>3</sup> As direções dos pássaros: Norte, Sul, Leste e Oeste (são as coordenadas geográficas); Leste como direção zero e sentido anti-horário como sentido de rotação positiva.

a Figura 5 o valor que corresponde a essa média é  $180^\circ$ , onde aponta para o oeste, enquanto as duas observações apontam essencialmente para o leste. Com esse exemplo fica visível que medidas usadas para variáveis lineares não são adequadas para representar as variáveis circulares. Portanto, é necessário alternativas tanto para medidas de posição como para medidas de dispersão ao lidar com dados circulares (JAMMALAMADAKA; SENGUPTA, 2001).

Figura 5 – Representação gráfica da média aritmética em dados circulares.



Fonte: Elaborado pelo o autor, 2022.

### 2.3.1 Medidas de Posição

#### 2.3.1.1 Média Circular

Suponha  $n$  vetores unitários  $\vec{OP}_1, \vec{OP}_2, \dots, \vec{OP}_n$ , com ângulos correspondentes  $\theta_i = 1, \dots, n$ . Define-se a média circular,  $\bar{\theta}$  de  $\theta_1, \dots, \theta_n$ , correspondente ao vetor resultante da soma  $\vec{OP}_1 + \vec{OP}_2 + \dots + \vec{OP}_n$ . Tal vetor é o centro de massa dos pontos  $(\cos \theta_1, \sin \theta_1), \dots, (\cos \theta_n, \sin \theta_n)$ , e as coordenadas cartesianas do centro de massa são  $(\bar{C}, \bar{S})$  dada por (MARDIA; JUPP, 2000):

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i, \quad \bar{S} = \frac{1}{n} \sum_{i=1}^n \sin \theta_i \quad (12)$$

Portanto, tem-se que  $\bar{\theta}$  é a solução das equações

$$\bar{C} = \bar{R} \cos \bar{\theta}, \quad \bar{S} = \bar{R} \sin \bar{\theta} \quad (13)$$

tem-se o comprimento médio resultante  $R$  (desde que  $R > 0$ ), onde é dado por:

$$\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2} \quad (14)$$

Observe que  $\bar{\theta}$  não é definido quando  $R = 0$ . Onde  $R > 0$ , então, a direção média é dada explicitamente por:

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S}/\bar{C}), & \text{se } \bar{S} \geq 0 \text{ e } \bar{C} > 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + \pi, & \text{se } \bar{C} < 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + 2\pi, & \text{se e } \bar{C} < 0, \\ \text{indefinido}, & \bar{S} = 0, \bar{C} = 0. \end{cases}$$

Da Equação 13 pode-se mostrar (vide Apêndice A) que,

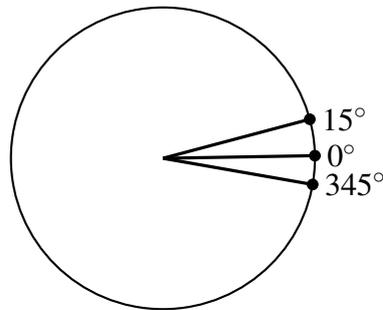
$$\frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \bar{\theta}) = \bar{R} \quad (15)$$

e

$$\frac{1}{n} \sum_{i=1}^n \sin(\theta_i - \bar{\theta}) = 0 \quad (16)$$

Ao utilizar a direção média circular do exemplo dos voos dos pássaros, realizou-se de forma correta os cálculos e encontrou o valor no ângulo de  $0^\circ$ . Portanto pela Figura 6, percebe-se que o valor está coerente com a interpretação dessa medida de posição.

Figura 6 – Representação gráfica da direção média circular em dados circulares.



Fonte: Elaborado pelo o autor, 2022.

### 2.3.1.2 Mediana Circular

A direção dessa medida de posição é definida  $\tilde{\theta}$  de ângulos  $\theta_1, \dots, \theta_n$ , em qualquer ângulo  $\phi$  tal que,

- (i) metade dos pontos de dados estejam no arco  $[\phi, \phi + \pi)$ ;
- (ii) a maioria dos pontos de dados são mais próximos de  $\phi$  do que para  $\phi + \pi$ .

Análoga a mediana para dados lineares, quando o tamanho da amostra  $n$  é ímpar, a mediana da amostra é um dos pontos dos dados, se  $n$  é par, é apropriado tomar a mediana da amostra como o ponto médio de dois pontos de dados adjacentes (MARDIA; JUPP, 2000).

## 2.3.2 Medidas de Concentração

### 2.3.2.1 Comprimento Médio do Vetor Médio Resultante

O comprimento médio resultante  $\bar{R}$  associado à direção média  $\bar{\theta}$ , e é dado por:

$$\bar{R} = \frac{R}{n} \quad (17)$$

A quantidade  $\bar{R}$  é utilizado como medida de concentração dos dados e encontra-se no intervalo  $(0, 1)$ . Seus valores extremos têm algumas propriedades interessantes. Se as direções  $\theta_1, \dots, \theta_n$  estão fortemente agrupadas, tem-se  $\bar{R} = 1$ , representando uma concentração máxima dos dados, e corresponde a uma variabilidade mínima entre eles. Por outro lado, se  $\theta_1, \dots, \theta_n$  estejam dispersos no círculo, nesta situação tem-se uma concentração mínima  $\bar{R} = 0$  e uma variabilidade máxima. No entanto,  $\bar{R} \simeq 0$  não significa que as observações estão dispersas uniformemente ao redor do círculo (MARDIA; JUPP, 2000).

### 2.3.3 Medidas de Dispersão

#### 2.3.3.1 Variância Circular

Para propósitos mais descritivos e inferenciais, o comprimento médio resultante  $\bar{R}$  é mais importante do que qualquer medida de dispersão, no entanto a sua utilização não é usual. Todavia, para fins de comparação com dados na reta, às vezes é útil considerar medidas de dispersão de dados circulares denotada de variância circular amostral, que pode ser definida por (MARDIA; JUPP, 2000):

$$V = 1 - \bar{R}. \quad (18)$$

A definição é análoga à variância de dados lineares, no sentido de quanto menor o valor da variância circular, maior a concentração nos dados. Sendo que,

$$0 \leq V \leq 1 \text{ e } 0 \leq \bar{R} \leq 1$$

Onde a variância mínima ocorre quando  $V = 0$  ( $\bar{R} = 1$ ), e corresponde as observações estarem concentradas no mesmo local. No caso para o máximo limite, ocorre quando  $V = 1$  ( $\bar{R} = 0$ ), ou seja, as observações estarem dispersas na circunferência.

#### 2.3.3.2 Desvio Padrão Circular

O conceito de desvio padrão é diferente do conceito para dados lineares, que equivale à raiz quadrada da variância. O desvio padrão circular amostral é definido por (MARDIA; JUPP, 2000):

$$v = \{-2 \log(1 - V)\}^{1/2} = \{-2 \log \bar{R}\}^{1/2}. \quad (19)$$

Observe que  $v$  assume valores em  $[0, \infty]$ , enquanto  $V$  assume valores em  $[0, 1]$ . Uma boa aproximação para o desvio padrão circular é dada por:

$$v \simeq \begin{cases} (2V)^{1/2}, & \text{para } V \text{ pequeno } (V < 0,18), \\ [2(1 - \bar{R})]^{1/2} & \text{para } \bar{R} \text{ grande } (\bar{R} > 0,82). \end{cases}$$

### 2.3.4 Momentos Trigonométricos

Conforme visto anteriormente, as expressões

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i, \quad \bar{S} = \frac{1}{n} \sum_{i=1}^n \sin \theta_i$$

desempenham papéis importantes para determinação da média direcional e variância amostral. Definimos o primeiro momento trigonométrico em torno de  $0^\circ$  por:

$$m'_1 = \bar{C} + i\bar{S}.$$

Observe que,

$$\bar{C} = \bar{R} \cos \bar{\theta}, \quad \bar{S} = \bar{R} \sin \bar{\theta}$$

Logo,

$$\begin{aligned} m'_1 &= \bar{R} \cos \bar{\theta} + i\bar{S} \sin \bar{\theta} \\ &= \bar{R}(\cos \bar{\theta} + i \sin \bar{\theta}) \end{aligned}$$

Portanto,

$$m'_1 = \bar{R}e^{i\bar{\theta}}. \quad (20)$$

Estendendo esta noção, definimos o  $p$ -ésimo momento trigonométrico em relação à direção  $0^\circ$  para  $p = 1, 2, \dots$  como

$$m'_p = a_p + ib_p, \quad (21)$$

Onde,

$$a_p = \frac{1}{n} \sum_{i=1}^n \cos p\theta_i, \quad b_p = \frac{1}{n} \sum_{i=1}^n \sin p\theta_i. \quad (22)$$

Sendo assim,

$$m'_p = \bar{R}_p e^{i\bar{\theta}_p}. \quad (23)$$

no qual  $\bar{\theta}_p$  e  $\bar{R}_p$  denotam a média circular e o comprimento médio do vetor médio resultante amostral de  $\theta_1, \dots, \theta_n$ . O  $p$ -ésimo momento trigonométrico em relação à média circular  $\bar{\theta}$  é

$$m_p = \bar{a}_p + i\bar{b}_p, \quad (24)$$

Em que,

$$\bar{a}_p = \frac{1}{n} \sum_{i=1}^n \cos [p(\theta_i - \bar{\theta})], \quad \bar{b}_p = \frac{1}{n} \sum_{i=1}^n \sin [p(\theta_i - \bar{\theta})]. \quad (25)$$

Em particular,

$$m_1 = \bar{a}_1 + i\bar{b}_1 = \frac{1}{n} \sum_{i=1}^n \cos [p(\theta_i - \bar{\theta})] + i \frac{1}{n} \sum_{i=1}^n \sin [p(\theta_i - \bar{\theta})]$$

Portanto os resultados das Equações (15) e (16) temos,

$$m_1 = \bar{R} + i \cdot 0 = \bar{R}.$$

## 2.4 Inferência

Em muitas pesquisas deseja-se estudar um determinado aspecto de um conjunto de indivíduos que tem determinada característica em comum (população), porém não é possível avaliar todo o grupo. Neste caso, realiza-se o estudo com os resultados obtidos a partir da análise de uma amostra e assim generaliza esses resultados para toda a população.

A Inferência Estatística apresenta técnicas adequadas que possibilitam fazer tais extrapolações e os principais ramos são: testes de hipóteses e estimação. Para que os resultados da modelagem estatística sejam confiáveis, as características de interesse da população podem ser representadas por uma variável aleatória e, de acordo com a natureza dos dados, buscar modelos probabilísticos adequados. Desta maneira, além da visão descritiva da Estatística básica, é viável abordar uma perspectiva probabilística da Estatística Circular através do ajuste de alguns modelos de distribuições (GÓIS, 2002).

Um meio de especificar uma distribuição circular é através de sua função de distribuição, no qual a sua probabilidade total está concentrada na circunferência do ciclo unitário. Como cada ponto da circunferência representa uma direção, essa distribuição é uma forma de atribuir probabilidades a diferentes direções ou definir uma distribuição circular. Uma variável aleatória circular  $\theta$ , medida em radianos, pode ser tomada como  $[0, 2\pi)$  ou  $[-\pi, \pi)$ . (MARDIA; JUPP, 2000)

As distribuições circulares são essencialmente de dois tipos: podem ser discretas ou podem ser absolutamente contínuas. No caso contínuo, tem-se que:

$$F(x) = Pr(0 < \theta \leq x), \quad 0 \leq x \leq 2\pi,$$

e

$$F(x + 2\pi) - F(x) = 1, \quad -\infty < x < \infty, \quad (26)$$

Então, uma função de distribuição  $F$  é definida como a função de densidade de probabilidade  $f(\theta)$  e possui as seguintes propriedades (FISHER, 1995):

- (i)  $f(\theta) \geq 0, -\infty < \theta < \infty$ ;
- (ii)  $\int_0^{2\pi} f(\theta) d\theta = 1$ ;
- (iii)  $f(\theta) = f(\theta + 2\pi)$  ( $f$  é periódica com período  $2\pi$ ).

Logo após, de forma breve serão percorridos alguns modelos probabilísticos usuais para dados circulares. Mais especificamente, a distribuição Uniforme e a distribuição de Von Mises.

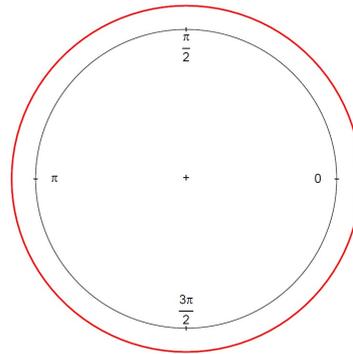
### 2.4.1 Distribuição Uniforme

A distribuição Uniforme é a distribuição mais básica sobre o círculo, todas as direções são igualmente prováveis, por isso essa distribuição caracteriza isotropia ou aleatoriedade circular. Observe-se que esta distribuição apresenta como características fundamentais não possuir uma média circular bem definida e o comprimento do vetor resultante ser igual a zero. Assim, a probabilidade total está espalhada uniformemente sobre a circunferência conforme é representada graficamente na Figura 7, cuja a função densidade de probabilidade é dada por (JAMMALAMADAKA; SENGUPTA, 2001):

$$f(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta \leq 2\pi \quad (27)$$

Quando uma distribuição não é Uniforme, haverá concentração em uma ou mais direções preferenciais. Mais comumente, se for apenas em uma direção a distribuição é chamada de unimodal.

Figura 7 – Gráfico da Função de Densidade de Distribuição Uniforme.



Fonte: Elaborado pelo o autor, 2022.

### 2.4.2 Distribuição Von Mises

Do ponto de vista da inferência estatística, esta distribuição é uma das mais importantes para dados circulares. Esta é uma distribuição unimodal simétrica que é o modelo mais comum para amostras unimodais de dados circulares. Ela é tão importante para dados circulares quanto o modelo Gaussiano é para dados lineares em geral. Esse modelo foi proposto pelo físico alemão R. Von Mises, em 1918, ao investigar desvio de pesos atômicos de valores inteiros (MARDIA; JUPP, 2000).

Portanto, uma variável aleatória circular  $\theta$  tem distribuição Von Mises com parâmetros  $\mu$  e  $k$  se sua função densidade de probabilidade for dada por:

$$f(\theta; \mu, k) = \frac{1}{2\pi I_0(k)} e^{k \cos(\theta - \mu)} \quad 0 \leq \theta < 2\pi, \quad 0 \leq \mu < 2\pi \text{ e } k > 0, \quad (28)$$

onde  $I_0(k)$  é uma função de Bessel modificada de ordem 0,

$$I_0(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{k \cos \theta} d\theta \quad (29)$$

O parâmetro  $\mu$  refere-se à média circular da distribuição e o  $k$  corresponde ao parâmetro de concentração e ele é similar a variância em uma distribuição normal. Quando  $k$  se aproxima de 0 a distribuição de Von Mises se aproxima da distribuição Uniforme. Em contrapartida, quando  $k$  se afasta de 0, concentra-se em  $\mu$ . A Figura 8 e a Figura 9 a seguir exemplifica essa distribuição com parâmetros distintos.

Figura 8 – Gráfico da Função de Densidade de Distribuição de Von Mises com parâmetros  $\mu = \pi$  e  $k = 0,5$

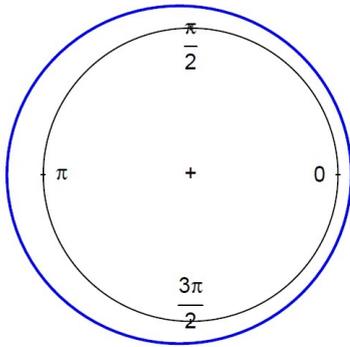
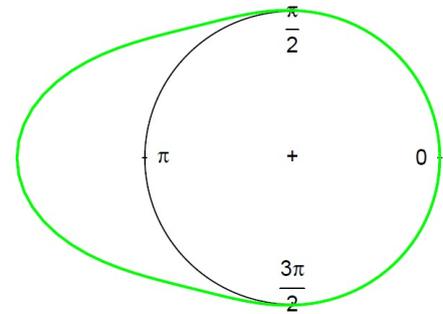


Figura 9 – Gráfico da Função de Densidade de Distribuição de Von Mises com parâmetros  $\mu = \pi$  e  $k = 5$



Fonte: Elaborado pelo o autor, 2022.

## 2.5 Teste de Hipótese

De acordo com (GÓIS, 2002), pode-se verificar que a Estatística Circular dispõe de uma série de testes que avaliam a qualidade do ajustamento das distribuições teóricas referentes aos dados resultantes das observações. A princípio deve-se supor a distribuição principal, no entanto, a busca por métodos mais robustos para uma grande classe de modelos possíveis levam naturalmente para os testes não-paramétricos (PEWSEY, 2013).

Em muitas circunstâncias há interesse em tomar a decisão de aceitar ou rejeitar determinada afirmação baseando-se em evidências fornecidos pela amostra. Assim, o pesquisador formula duas hipóteses, onde a hipótese nula é confrontada a uma hipótese alternativa. Caso haja alguma evidência na informação da amostra que possibilita a rejeição da hipótese nula, a hipótese alternativa passa a ser considerada.

### 2.5.1 Teste de Rayleigh

O teste de Rayleigh é usado para verificar a distribuição das observações da amostra estudada por toda circunferência. Este método baseia-se no fato do comprimento do vetor médio  $\bar{R}$  apresentar um grande valor, significa uma grande concentração dos dados, assim existirá uma menor probabilidade dos dados serem distribuídos de maneira uniforme em torno da média, ou seja, se p-valor for inferior ao nível de significância escolhido (normalmente 0,05), indicando que

os dados não apresentam distribuição Uniforme, mostrando indícios de uma direção preferencial. A estatística de teste é dada por (MARDIA; JUPP, 2000):

$$Z = n\bar{R}^2 \quad (30)$$

onde,  $n$  é número de observações e  $\bar{R}$  é o comprimento do vetor médio.

Os valores da estatística de teste são usados para encontrar a função do p-valor qual o mesmo é definido através de:

$$p = \begin{cases} e^{-z} \left( 1 + \frac{(2z-z^2)}{4n} - \frac{24z-132z^2+76z-9z^4}{288n^2} \right), & \text{se } n < 50 \\ e^{-z}, & \text{se } n \geq 50 \end{cases} \quad (31)$$

Quando não rejeita-se a hipótese nula, compreende-se que a distribuição circular é uniforme e quando a hipótese alternativa não é descartada, compreende-se que a distribuição circular não é uniformemente distribuída por toda circunferência.

### 2.5.2 Teste de Watson $U_n^2$

Para este teste a hipótese nula defende a ideia que a direção média das amostras são iguais e para hipótese alternativa que as médias diferem. De acordo com (MARDIA; JUPP, 2000)  $U^2$  é análogo para dados circulares da estatística Cramér-von Mises  $W^2$  para dados na linha real. Assim, tem-se:

$$U^2 = \sum_{i=1}^n \left[ \left( U_i - \frac{(i-1/2)}{n} \right) - \left( \bar{U} - \frac{1}{2} \right) \right]^2 + \frac{1}{12n} \quad (32)$$

onde,

- $n$ : o número de observações consideradas por ordem crescente;
- $U_i$ : refere-se a função de distribuição;
- $\bar{U}$ : a média dos  $U_i$ .

### 2.5.3 Teste de espaçamento de Rao

Este teste analisa os espaçamentos entre os pontos adjacentes de modo que estejam distribuídos à volta de um círculo. Conforme (GÓIS, 2002) em uma distribuição Uniforme o espaçamento entre os dados deverá ser aproximadamente  $\frac{360^\circ}{n}$ , caso os espaçamentos verificados sejam muito diferentes deste valor, então é reduzida a probabilidade dos dados estarem uniformemente distribuídos.

## 3 MATERIAL E MÉTODOS

### 3.1 Base de dados e área do estudo

Para o desenvolvimento deste trabalho foram utilizados dados reais provenientes do SINAN para obtenção de informações sobre os casos de dengue no município de Campina

Grande. Os dados analisados compreendem mensalmente entre janeiro de 2018 e dezembro de 2021.

A cidade de Campina Grande, está localizada na Mesorregião do Agreste Paraibano, possuindo uma área territorial de  $591,658 \text{ km}^2$ , e com população estimada em 413.830 habitantes (IBGE, 2021), onde tem um clima caracterizado como tropical semiúmido, com chuvas geralmente de fevereiro a setembro, sendo que o período mais chuvoso ocorre entre maio e julho.

### 3.2 Dados epidemiológicos

Através do Sistema de Informações de Agravos de Notificação (SINAN) foram obtidos os dados epidemiológicos, no qual, tem como objetivo coletar, transmitir e disseminar dados gerados pelo Sistema de Vigilância Epidemiológica. O sistema é alimentado pela notificação e investigação de casos de doenças e agravos que compõem a lista nacional de doenças de notificação compulsória.

Segundo o Ministério da Saúde, o SINAN foi implantado, a partir de 1993 e a utilização dos dados epidemiológicos permitem que os profissionais de saúde consigam identificar a realidade epidemiológica de determinadas áreas geográficas, e baseado nisso auxiliar no planejamento da saúde e definir prioridades de intervenção.

### 3.3 Métodos e Pacotes Utilizados

Antes de analisar os resultados foi preciso realizar o tratamento das bases disponibilizada do SINAN por meio do TABENET (TABENET, 2023). Sendo assim, o primeiro passo do tratamento da base retirada do SINAN foi efetuar a filtragem referente ao município de notificação, em seguida ao mês de notificação e períodos disponíveis, onde os períodos para análise são relacionados aos anos de 2018 a 2021, a partir dessas manipulações foram criados novos bancos de dados no software Excel. Com isso os novos bancos de dados contêm informações de:

- **mes:** Janeiro a Dezembro;
- **casosCG:** Casos prováveis segundo município de Campina Grande - PB referentes a cada mês;
- **angulo:** Referentes aos meses.

Como a Estatística Circular se encarrega de estudar o direcionamento dos dados, os valores dos casos prováveis de dengue e os ângulos são números inteiros, e trabalhar com os dados corretamente é fundamental, desta maneira foi necessário transformá-los em radianos e posteriormente em dados circulares. Neste contexto, o tratamento estatístico foi iniciado a partir da transformação direcional do mês em medida angular, tendo o mês de janeiro como referência, sendo assim, utilizou-se 12 intervalos de  $30^\circ$ . Desta forma, cada mês correspondeu a um valor distribuído ao longo da circunferência.

Fundamentado nisso, encontrou-se o ângulo médio por meio do comando *mean.circular(x)*, onde  $x$  refere-se ao valor do dado em circular. Como também, foram calculadas as medidas referentes a mediana *median.circular(x)* e o comprimento do vetor médio *rho.circular(x)*, medidas de dispersão como a variância circular *var.circular(x)* e o desvio padrão *sd.circular(x)*. Através desses parâmetros encontrados, dando ênfase a média circular e a variância circular, observou-se a distribuição dos pontos por meio do gráfico de dispersão obtido pelo o comando *plot.circular(x)* e o diagrama de rosa *rose.diag(x)*, no qual representam a distribuição das observações colhidas em função da circunferência. Além do mais, para saber o comportamento da distribuição em toda

a circunferência das amostras foi utilizado o teste de Rayleigh  $rayleigh.test(x)$ , seguido pelos os testes de Watson  $watson.test(x)$  e de Rao  $rao.spacing.test(x)$ , ao nível de 5% de significância.

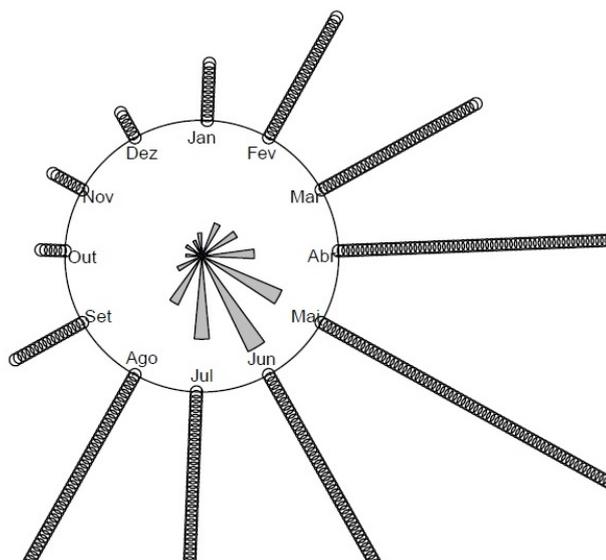
Os procedimentos estatísticos realizados neste trabalho foram executado no software R (R Core Team, 2021). Para as análises foram utilizados os seguintes pacotes: *dplyr* que realiza manipulações de dados, onde especificamente usou a função *glimpse* para exibir informações do dataset, como quantidade de linhas e o tipo das variáveis e, *Circular* para as análises de dados circulares.

#### 4 RESULTADOS E DISCUSSÃO

Nesta seção apresenta os resultados das análises da Estatística Circular, onde realizou-se análise descritiva, com média circular, mediana circular, comprimento médio resultante, variância circular, desvio circular, além de fazer o uso dos testes para saber o comportamento da distribuição em toda a circunferência das amostras em estudo. Para uma melhor visualização das informações obtidas pelas estatísticas, foram criados gráficos para observar como os pontos distribuem-se ao longo da circunferência.

Na Figura 10, está representado o diagrama de rosa e o gráfico de dispersão circular para os casos de dengue no ano de 2018, observa-se que os dados estão mais concentrados nos meses de maio a julho. Pela distribuição dos pontos, parece trata-se de dados com distribuição unimodal, todavia só pode afirmar acerca disso através de teste específico.

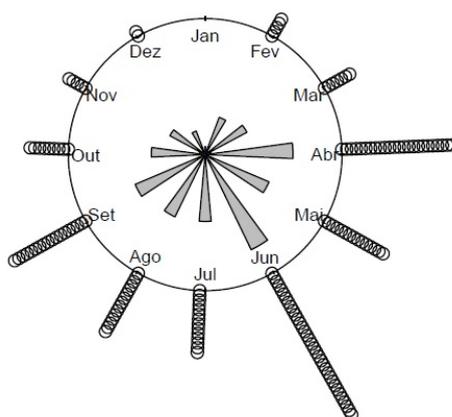
Figura 10 – Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2018.



Fonte: Elaborado pelo o autor, 2022.

Observando o diagrama circular e o gráfico de dispersão circular para os casos de dengue no ano de 2019 que está representado na Figura 11, percebe-se uma maior concentração dos dados entre os meses de abril e setembro, porém não é possível concluir sobre a uniformidade dos dados graficamente.

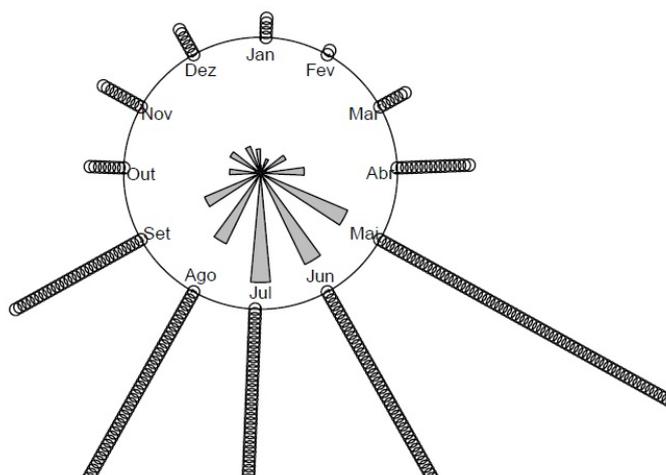
Figura 11 – Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2019.



Fonte: Elaborado pelo o autor, 2022.

Na Figura 12, observa-se que o diagrama circular e o gráfico de dispersão circular para os casos de dengue no ano de 2020 indicam que os dados se concentram nos meses de maio a setembro e a priori não há indícios de uniformidade dos dados.

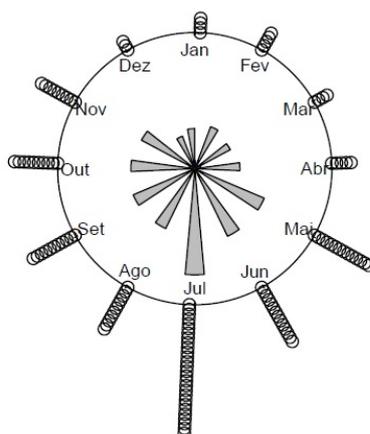
Figura 12 – Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2020.



Fonte: Elaborado pelo o autor, 2022.

Para os casos de dengue no ano de 2021 pode-se observar que o diagrama circular e o gráfico de dispersão circular representado na Figura 13 apresenta uma maior concentração dos dados nos meses de maio a novembro, no entanto não há evidências de uniformidade dos dados.

Figura 13 – Representação do gráfico de dispersão circular e do diagrama de rosa dos casos prováveis de dengue no município de Campina Grande - PB no ano 2021.



Fonte: Elaborado pelo o autor, 2022.

Para caracterizar os casos prováveis de dengue na cidade de Campina Grande, usou-se as medidas de posição e as medidas de dispersão que estão apresentadas na Tabela 1.

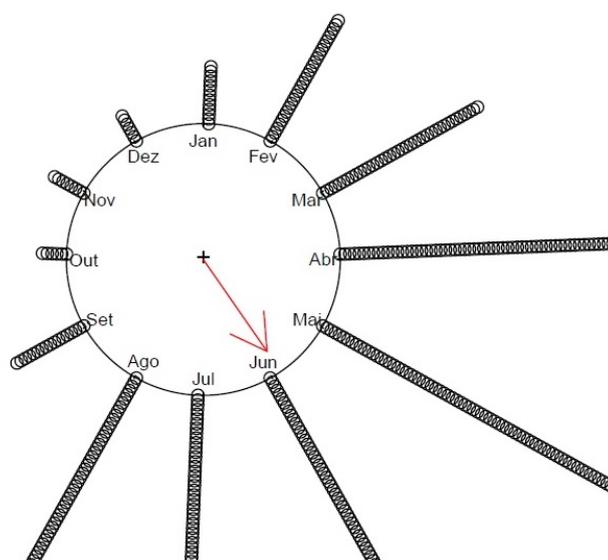
Tabela 1 – Dados circulares de casos prováveis de dengue na cidade de Campina Grande - PB.

Ano	$n$	$\bar{\mu}(\text{°})$	$\bar{\mu}(\text{rad})$	$\tilde{\theta}(\text{°})$	$\tilde{\theta}(\text{rad})$	$\bar{R}$	$V$	$\nu$
2018	867	145,4592°	2,5387	150°	2,6180	0,6841	0,3159	0,8714
2019	149	157,4519°	2,7481	150°	2,6180	0,4487	0,5513	1,2661
2020	401	166,9832°	2,9144	180°	3,1420	0,6478	0,3522	0,9319
2021	118	185,4058°	-3,0472	180°	3,1420	0,3807	0,6193	1,3898

Fonte: Elaborado pelo o autor, 2022.

Ao analisar os casos prováveis de dengue na cidade de Campina Grande para o ano de 2018, a direção média foi igual a 2,5387 radianos que corresponde a 145,4592°, para a direção mediana dos dados, o valor foi igual a 2,6180 radianos, correspondendo a direção ao mês de junho. Para esses dados então, há pouca diferença entre as direções média e mediana. Essa semelhança entre seus valores é consistente com a distribuição dos dados, sendo próxima da simetria. A variabilidade em torno da direção média foi igual a 0,3159, sendo o ano com a menor variabilidade, nota-se que seu valor não foi próximo de 1, o que denota que os dados não estão dispersos na circunferência, também pode ser evidenciado concentração dos dados ao longo da direção média, de acordo os dados da Tabela 1. Desta forma, pode-se constatar ao longo dos 12 meses do ano a direção média dos casos prováveis de dengue foi aproximadamente no mês de junho, conforme está representado na Figura 14.

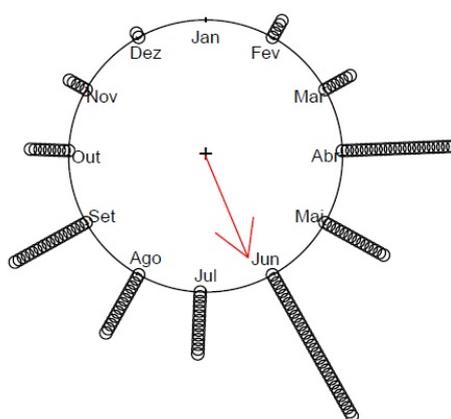
Figura 14 – Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2018.



Fonte: Elaborado pelo o autor, 2022.

Em relação aos casos prováveis de dengue na cidade de Campina Grande para o ano de 2019, tem como direção média  $2,7481$  radianos que corresponde a  $157,4519^\circ$ , a direção mediana foi igual a  $2,6180$  radianos correspondendo a direção dos dados ao mês de junho, semelhante ao ano anterior. A variabilidade em torno da direção média foi igual  $0,5513$ , observa-se que seu valor não foi próximo de 1, o que evidencia que os dados não estão dispersos na circunferência, também pode ser demonstrado que há uma razoável concentração dos dados ao longo da direção média, em conformidade com os dados da Tabela 1. Assim, pode-se notar ao longo dos 12 meses do ano a direção média dos casos prováveis de dengue em torno do mês de junho, conforme está representado na Figura 15.

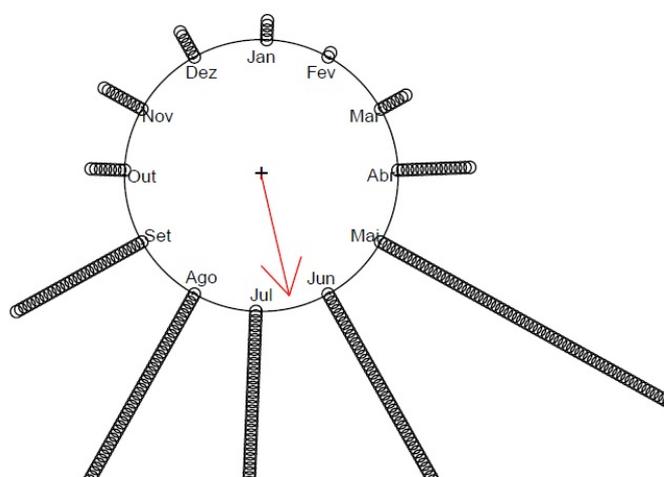
Figura 15 – Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2019.



Fonte: Elaborado pelo o autor, 2022.

Para o ano de 2020, os casos prováveis de dengue na cidade de Campina Grande teve como direção média de 2,9144 radianos que condiz a  $166,9832^\circ$ , para direção mediana um valor de 3,1420 radianos, representando uma direção ao mês de julho, onde percebe-se pouca diferença para esses dados entre a média e a mediana. A variabilidade em torno da direção média igual 0,3522, pode-se observar que seu valor não foi próximo de 1, o que demonstra que os dados não estão dispersos na circunferência, nota-se que houve aglomerações em torno da média, de acordo com os dados da Tabela 1. Sendo assim, constata-se que ao longo dos 12 meses do ano a direção média dos casos prováveis de dengue foi próximo ao mês de julho conforme representado na Figura 16.

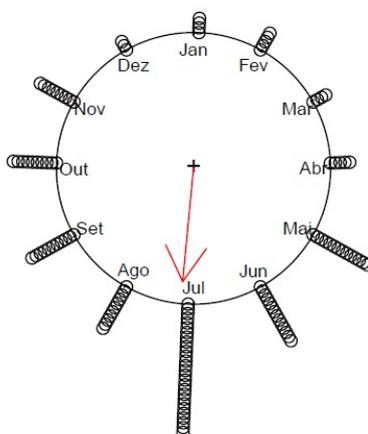
Figura 16 – Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2020.



Fonte: Elaborado pelo o autor, 2022.

A direção média dos casos prováveis de dengue na cidade de Campina Grande no ano de 2021 foi igual a  $-3,0472$  radianos que corresponde a  $185,4058^\circ$ , a direção mediana igual a 3,1420 radianos, o que corresponde uma direção ao mês de julho, e a variabilidade em torno da direção média igual 0,6193, onde representa uma maior variabilidade dos dados, observa-se que seu valor não foi próximo de 1, o que confirma que os dados não estão dispersos na circunferência, nota-se que houve uma baixa concentração dos dados em torno da média, conforme os dados da Tabela 1. Em vista disso, percebe-se que ao longo dos 12 meses do ano a direção média dos casos prováveis de dengue foi próximo ao mês de julho de acordo a Figura 17.

Figura 17 – Representação do gráfico de dispersão circular dos casos prováveis de dengue no município de Campina Grande - PB no ano 2021.



Fonte: Elaborado pelo o autor, 2022.

Para realização dos testes de Uniformidade aplicou-se o teste Rayleigh, Watson e de espaçamento de Rao, como estão representados na Tabela 2.

Tabela 2 – Teste de uniformidade para os casos de dengue na cidade de Campina Grande - PB.

Ano	Teste de Rayleigh ( $H_0$ )	Teste de Watson	Teste de Rao
2018	< 0,001	< 0,001	< 0,001
2019	< 0,001	< 0,001	< 0,001
2020	< 0,001	< 0,001	< 0,001
2021	< 0,001	< 0,001	< 0,001

Fonte: Elaborado pelo o autor, 2022.

Como apenas a avaliação gráfica não é possível concluir sobre a uniformidade dos dados, desta forma aplicou-se o teste de Rayleigh de uniformidade para os 4 anos (2018 a 2021) referentes os casos prováveis de dengue na cidade de Campina Grande, onde a hipótese nula ( $H_0$ ) defendeu a ideia de que os casos prováveis de dengue na cidade de Campina Grande acontecem de forma aleatória, ou seja, se distribuem uniformemente nos 12 meses para os anos de 2018 a 2021. Já em relação a hipótese alternativa ( $H_1$ ), testou se existe não aleatoriedade dos casos prováveis de dengue, isto é, se houve uma direção específica dos dados. De acordo com os valores descritos na Tabela 2, pode-se observar que na ( $H_0$ ) o p-valor é baixo, evidenciando assim a rejeição dessa hipótese, ou seja, indicando que há evidências que os dados estão agregados. Assim como no teste de Rayleigh, os dois testes enfaticamente apontaram que os dados de 2018, 2019, 2020 e 2021 não seguem uma distribuição uniforme, apontando assim para uma distribuição unimodal dos dados.

## 5 CONCLUSÃO

Diante do exposto, os métodos pertencentes a Estatística Circular permitem a avaliação da concentração dos dados em uma circunferência, como também proporciona o cálculo da média circular, mediana circular, da variância e do desvio padrão circular, levando em consideração a periodicidade natural desse tipo de dados, o que é bem diferente aos dados estudados pela Estatística tradicional.

Também verificou-se tanto por meio do Teste de Rayleigh, de Watson e de espaçamento de Rao a existência de casos prováveis de dengue na cidade de Campina Grande em direções não aleatórias, garantido uma direção específica (Sazonalidade dos dados) dos casos de dengue para o período de 4 anos.

Em vista disto, a proposta deste trabalho foi apresentar e desenvolver a metodologia para medidas circulares e avaliar os casos de dengue no município de Campina Grande. Além do mais, os resultados demonstraram que os casos de dengue aumentam no período chuvoso (maio a julho) e que isso deve-se a ocorrência de fatores climáticos, acúmulo de água, falta de apoio de gestão, que contribuem para formação de ambientes favoráveis para o desenvolvimento e a proliferação do vetor.

Sendo assim, vale salientar que a disseminação dessas técnicas é de suma importância para a obtenção adequada dos casos que envolvem as medidas circulares, proporcionando interpretações coerentes dos resultados, quer sejam para representação gráfica, análises descritivas ou mesmo para ajustar um determinado modelo.

## REFERÊNCIAS

- FISHER. Dispersion on a sphere, proc. r. In: *Soc. London*. [S.l.: s.n.], 1953. v. 17. Citado na página 11.
- FISHER, N. I. *Statistical analysis of circular data*. [S.l.]: cambridge university press, 1995. Citado 3 vezes nas páginas 11, 15 e 21.
- GÓIS, J. E. S. Contribuição dos modelos estocásticos para o estudo da climatologia urbana. 2002. Citado 3 vezes nas páginas 21, 23 e 24.
- IBGE. *IBGE - Instituto Brasileiro de Geografia e Estatística*. 2021. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/pb/campina-grande.html>>. Citado na página 25.
- JAMMALAMADAKA, S. R.; SENGUPTA, A. *Topics in circular statistics*. [S.l.]: world scientific, 2001. v. 5. Citado 4 vezes nas páginas 12, 14, 17 e 22.
- MAGALHÃES, M. N.; LIMA, A. C. P. D. *Noções de probabilidade e estatística*. [S.l.]: Editora da Universidade de São Paulo, 2015. v. 7. Citado na página 13.
- MARDIA, K. V.; JUPP. *Directional statistics*. [S.l.]: Wiley Online Library, 2000. v. 2. Citado 7 vezes nas páginas 13, 17, 18, 19, 21, 22 e 24.
- MISES, R. von. Uber die "ganzzahligkeit" der atomgewicht und verwandte fragen. *Physikal. Z.*, v. 19, p. 490–500, 1918. Citado na página 11.
- NEVES, M. R. *Modelos para séries temporais de dados circulares*. Tese (Doutorado) — Dissertação de Mestrado, Instituto de Matemática, Universidade Federal do . . . , 2015. Citado na página 12.
- PEWSEY. *Circular statistics in R*. [S.l.]: Oxford University Press, 2013. Citado na página 23.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<http://www.R-project.org/>>. Citado 2 vezes nas páginas 10 e 26.
- RAYLEIGH, L. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Taylor & Francis, v. 10, n. 60, p. 73–78, 1880. Citado na página 11.
- TABENET. 2023. Disponível em: <[http://tabnet.datasus.gov.br/Tutorial/Tutorial\\_tabNet\\_FINAL.pdf](http://tabnet.datasus.gov.br/Tutorial/Tutorial_tabNet_FINAL.pdf)>. Acesso em: 20/03/2023. Citado 3 vezes nas páginas 10, 11 e 25.
- WATSON, G. S.; WILLIAMS, E. J. On the construction of significance tests on the circle and the sphere. *Biometrika*, JSTOR, v. 43, n. 3/4, p. 344–352, 1956. Citado na página 11.

## APÊNDICE A – DEMONSTRAÇÃO DAS EQUAÇÕES (15) E (16)

Neste apêndice apresentaremos a demonstrações das equações (15) e (16)

### A.1 Demonstração

Propriedade do cosseno da diferença de dois arcos trigonométricos, utiliza-se,

$$\cos(a - b) = \cos a \cdot \cos b + \sin a \cdot \sin b$$

Aplicando na expressão (15), observa-se que

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \bar{\theta}) &= \frac{1}{n} \sum_{i=1}^n (\cos \theta_i \cdot \cos \bar{\theta} + \sin \theta_i \cdot \sin \bar{\theta}) \\ &= \frac{1}{n} \left[ \cos \bar{\theta} \sum_{i=1}^n \cos \theta_i + \sin \bar{\theta} \sum_{i=1}^n \sin \theta_i \right] \end{aligned}$$

Temos,

$$\frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \cos \bar{\theta}) = \bar{C} \cos \bar{\theta} + \bar{S} \sin \bar{\theta}$$

Substituindo a Equação (13) na expressão acima temos

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \cos \bar{\theta}) &= \bar{R} \cos^2 \bar{\theta} + \bar{R} \sin^2 \bar{\theta} \\ &= \bar{R} (\cos^2 \bar{\theta} + \sin^2 \bar{\theta}) \\ &= \bar{R}. \end{aligned}$$

Logo,

$$\frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \cos \bar{\theta}) = \bar{R}.$$

Para a propriedade do seno da diferença de dois arcos trigonométricos, temos,

$$\sin(a - b) = \sin a \cdot \cos b - \sin b \cdot \cos a$$

Fazendo a aplicação

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \sin(\theta_i - \bar{\theta}) &= \frac{1}{n} \sum_{i=1}^n (\sin \theta_i \cdot \cos \bar{\theta} - \sin \bar{\theta} \cdot \cos \theta_i) \\ &= \frac{1}{n} \left[ \cos \bar{\theta} \sum_{i=1}^n \sin \theta_i - \sin \bar{\theta} \sum_{i=1}^n \cos \theta_i \right] \end{aligned}$$

Fazendo as substituições,

$$\begin{aligned}\frac{1}{n} \sum_{i=1}^n \sin(\theta_i - \bar{\theta}) &= \bar{S} \cos \bar{\theta} - \bar{C} \sin \bar{\theta} \\ &= \frac{\bar{S} \cdot \bar{C}}{\bar{R}} - \frac{\bar{C} \cdot \bar{S}}{\bar{R}} \\ &= 0.\end{aligned}$$

Portanto,

$$\frac{1}{n} \sum_{i=1}^n \sin(\theta_i - \bar{\theta}) = 0.$$

**APÊNDICE B – ROTINA NO R**

```
#-----  
#  
## Dengue - Notificações registradas no Sistema de informação de agravos  
## de notificação - PB  
## Casos Prováveis por Mês notificação segundo Município de notificação  
## (OBS: Município de Campina Grande)  
## Período: 2018 a 2021  
## Fonte: Ministério da Saúde/SVS - Sistema de Informação de  
## Agravos de Notificação - Sinan Net  
#  
#-----  
#  
## Análise de Dados Circulares  
#  
# Pacotes necessários  
  
library(readxl)  
library(tidyverse)  
library(circular)  
  
#-----  
# Dados - CG 2018  
#-----  
  
CG_2018= read_excel("denguePB.xlsx", sheet = 8)  
glimpse(CG_2018)  
  
## Manipulando e visualizando os dados  
  
# Informando os valores dos angulos  
  
CG2018 = rep(CG_2018$angulo, CG_2018$casosCG)  
  
# Transformando em radiano  
  
CG2018.rad = rad(CG2018)  
  
# Transformando os dados em circular  
  
CG2018.circ = as.circular(CG2018.rad)  
  
## Analises descritivas
```

```
summary(CG2018.circ)

# Media circular

mean.circular(CG2018.circ)

# Para converter a media em angulo: angulo medio

deg(mean.circular(CG2018.circ))

# Mediana circular

median.circular(CG2018.circ)

# Para converter a mediana em angulo

deg(median.circular(CG2018.circ))

# Comprimento do vetor medio r

rho.circular(CG2018.circ)

## Medidas de dispersao

# Variancia circular

var.circular(CG2018.circ)

# Para converter a varianca em angulo

deg(var.circular(CG2018.circ))

# Desvio padrao angular

sd.circular(CG2018.circ)

# Para converter o desvio padrao em angulo:

deg(sd.circular(CG2018.circ))
```

```

## Testes de Rayleigh, Watson e Rao

# H0: Nao ha agregacao dos dados
# H1: Ha agregacao dos dados

rayleigh.test(CG2018.circ)
watson.test(CG2018.circ)
rao.spacing.test(CG2018.circ)

## Visualizacao grafica

plot.circular(CG2018.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
              zero = pi/2, stack = T, units = "rads", axes = F,
              col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                          "Nov", "Dez", "Jan", "Fev", "Mar","Abr")),
              cex =0.7, col = "black")

# Diagrama de rosa

rose.diag(cbind(CG2018.circ), rotation = "clock", zero = pi/2,
           units = "rads", axes = F, border = c("black"),
           col = "grey", prop = 1.4, bins = 35, ticks = F,
           shrink=2, add = T) #prop muda o tamanho das barras

# Media

plot.circular(CG2018.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
              zero = pi/2, stack = T, units = "rads", axes = F,
              col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                          "Nov", "Dez", "Jan", "Fev", "Mar","Abr")),
              cex =0.7, col = "black")

arrows.circular(mean.circular(CG2018.circ), rho.circular(CG2018.circ),

```

```
zero = pi/2, rotation = "clock", col = "red", shrink = 1.2)

#-----
# Dados - CG 2019
#-----

CG_2019= read_excel("denguePB.xlsx", sheet = 6)
glimpse(CG_2019)

## Manipulando e visualizando os dados

# Informando os valores dos angulos

CG2019 = rep(CG_2019$angulo, CG_2019$casosCG)

# Transformando em radiano

CG2019.rad = rad(CG2019)

# Transformando os dados em circular

CG2019.circ = as.circular(CG2019.rad)

## Analises descritivas

summary(CG2019.circ)

# Media circular

mean.circular(CG2019.circ)

# Para converter a media em angulo: angulo medio

deg(mean.circular(CG2019.circ))

# Mediana circular

median.circular(CG2019.circ)
```

```
# Para converter a mediana em angulo
deg(median.circular(CG2019.circ))

# Comprimento do vetor medio r
rho.circular(CG2019.circ)

## Medidas de dispersao

# Variancia circular
var.circular(CG2019.circ)

# Para converter a varianca em angulo
deg(var.circular(CG2019.circ))

# Desvio padrao angular
sd.circular(CG2019.circ)

# Para converter o desvio padrao em angulo:
deg(sd.circular(CG2019.circ))

## Teste de Rayleigh, Watson e Rao

# H0: Nao ha agregacao dos dados
# H1: Ha agregacao dos dados

rayleigh.test(CG2019.circ)
watson.test(CG2019.circ)
rao.spacing.test(CG2019.circ)

## Visualizacao grafica

plot.circular(CG2019.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
             zero = pi/2, stack = T, units = "rads", axes = F,
             col = c("black"), ticks = F, shrink=2, pch = 1)
```

```

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                           "Nov", "Dez", "Jan", "Fev", "Mar","Abr")),
              cex =0.7, col = "black")

# Diagrama de rosa

rose.diag(cbind(CG2019.circ), rotation = "clock", zero = pi/2,
          units = "rads", axes = F, border = c("black"),
          col = "grey", prop = 1.6, bins = 35, ticks = F,
          shrink=2, add = T) #prop muda o tamanho das barras

# Media

plot.circular(CG2019.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
             zero = pi/2, stack = T, units = "rads", axes = F,
             col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                           "Nov", "Dez", "Jan", "Fev", "Mar","Abr")),
              cex =0.7, col = "black")

arrows.circular(mean.circular(CG2019.circ), rho.circular(CG2019.circ),
                zero = pi/2, rotation = "clock", col = "red", shrink = 1.8)

#-----
# Dados - CG 2020
#-----

CG_2020= read_excel("denguePB.xlsx", sheet = 4)
glimpse(CG_2020)

## Manipulando e visualizando os dados

# Informando os valores dos angulos:

```

```
CG2020 = rep(CG_2020$angulo, CG_2020$casosCG)

# Transformando em radiano

CG2020.rad = rad(CG2020)

# Transformando os dados em circular

CG2020.circ = as.circular(CG2020.rad)

## Analises descritivas

summary(CG2020.circ)

# Media circular

mean.circular(CG2020.circ)

# Para converter a media em angulo: angulo medio

deg(mean(CG2020.circ))

# Mediana circular

median.circular(CG2020.circ)

# Para converter a mediana em angulo

deg(median.circular(CG2020.circ))

# Comprimento do vetor medio r

rho.circular(CG2020.circ)

## Medidas de dispersao

# Variancia circular

var.circular(CG2020.circ)
```

```

# Para converter a varianca em angulo

deg(var.circular(CG2020.circ))

# Desvio padrao angular

sd.circular(CG2020.circ)

# Para converter o desvio padrao em angulo:

deg(sd.circular(CG2020.circ))

## Teste de Rayleigh, Watson e Rao

# H0: Nao ha agregacao dos dados
# H1: Ha agregacao dos dados

rayleigh.test(CG2020.circ)
watson.test(CG2020.circ)
rao.spacing.test(CG2020.circ)

## Visualizacao grafica

plot.circular(CG2020.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
             zero = pi/2, stack = T, units = "rads", axes = F,
             col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
             c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                         "Nov", "Dez", "Jan", "Fev", "Mar", "Abr")),
             cex =0.7, col = "black")

# Diagrama de rosa

rose.diag(cbind(CG2020.circ), rotation = "clock", zero = pi/2,
          units = "rads", axes = F, border = c("black"),
          col = "grey", prop = 1.6, bins = 35, ticks = F,
          shrink=2, add = T) #prop muda o tamanho das barras

```

```

# Media

plot.circular(CG2020.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
             zero = pi/2, stack = T, units = "rads", axes = F,
             col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
             c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                         "Nov", "Dez", "Jan", "Fev", "Mar", "Abr")),
             cex =0.7, col = "black")

arrows.circular(mean.circular(CG2020.circ), rho.circular(CG2020.circ),
              zero = pi/2, rotation = "clock", col = "red", shrink = 1.4)

#-----
# Dados - CG 2021
#-----

CG_2021= read_excel("denguePB.xlsx", sheet = 2)
glimpse(CG_2021)

## Manipulando e visualizando os dados

# Informando os valores dos angulos

CG2021 = rep(CG_2021$angulo, CG_2021$casosCG)

# Transformando em radiano

CG2021.rad = rad(CG2021)

# Transformando os dados em circular

CG2021.circ = as.circular(CG2021.rad)

## Analises descritivas

summary(CG2021.circ)

```

```
# Media circular
mean.circular(CG2021.circ)

# Para converter a media em angulo: angulo medio
deg(mean(CG2021.circ))+360

# Mediana circular
median.circular(CG2021.circ)

# Para converter a mediana em angulo
deg(median.circular(CG2021.circ))

# Comprimento do vetor medio r
rho.circular(CG2021.circ)

## Medidas de dispersao

# Variancia circular
var.circular(CG2021.circ)

# Para converter a varianca em angulo
deg(var.circular(CG2021.circ))

# Desvio padrao angular
sd.circular(CG2021.circ)

# Para converter o desvio padrao em angulo
deg(sd.circular(CG2021.circ))
```

```

## Teste de Rayleigh, Watson e Rao

# H0: Nao ha agregacao dos dados
# H1: Ha agregacao dos dados

rayleigh.test(CG2021.circ)
watson.test(CG2021.circ)
rao.spacing.test(CG2021.circ)

## Visualizacao grafica

plot.circular(CG2021.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
              zero = pi/2, stack = T, units = "rads", axes = F,
              col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                          "Nov", "Dez", "Jan", "Fev", "Mar", "Abr")),
              cex =0.7, col = "black")

# Diagrama de rosa

rose.diag(cbind(CG2021.circ), rotation = "clock", zero = pi/2,
           units = "rads", axes = F, border = c("black"),
           col = "grey", prop = 1.6, bins = 35, ticks = F,
           shrink=2, add = T) #prop muda o tamanho das barras

# Media

plot.circular(CG2021.circ, cex=1, bin=80, sep=0.035, rotation = "clock",
              zero = pi/2, stack = T, units = "rads", axes = F,
              col = c("black"), ticks = F, shrink=2, pch = 1)

# Para colocar os nomes dos meses

axis.circular(at=circular(sort(seq(0, 11/6*pi, pi/6), decreasing = T)),
              c(labels = c("Mai", "Jun", "Jul", "Ago", "Set", "Out",
                          "Nov", "Dez", "Jan", "Fev", "Mar", "Abr")),
              cex =0.7, col = "black")

arrows.circular(mean.circular(CG2021.circ), rho.circular(CG2021.circ),
                zero = pi/2, rotation = "clock", col = "red", shrink = 2.2)

```

## AGRADECIMENTOS

Primeiramente agradeço a Deus por sua infinita bondade, que ao longo desses anos de graduação tem me dado saúde e forças para superar as dificuldades, sem sua presença seria impossível todas as minhas conquistas.

Agradeço em especial a meu esposo Antonio Farias, pelo amor, apoio e incentivo. Obrigada por sua compreensão e paciência demonstrada. E ao nosso filho (a), que estatisticamente não viria agora, mas divinamente foi providenciado.

Gostaria de agradecer a minha mãe e as minhas irmãs, pelo amor, apoio e incentivo durante toda a minha vida.

A minha orientadora Profa. Dra. Divanilda Maia Esteves, com ela pude desenvolver 2 PIBICs (Projeto de Iniciação Científica), e que desde de 2018 vem passando seus conhecimentos de forma tão generosa, uma profissional excelente em quem me espelho diariamente.

Quero agradecer aos colegas e amigos que a universidade juntou, especialmente Damião Everton Gomes Maia, Maria Fagna Felix da Silva, Sóstenes Jerônimo e Patrick da Silva Oliveira, que em todo o decorrer do curso compartilhamos saberes e conhecimentos.

Quero agradecer também a todos os professores do Departamento de Estatística que me ajudaram de alguma forma a chegar até aqui.