

Estatística circular: um estudo direcional, via coordenadas geográficas, dos alunos aprovados na Escola Preparatória de Cadetes do Ar

Circular statistics: a directional study, via geographic coordinates, of students approved at the Air Cadet Preparatory School

Estadística circular: un estudio direccional, vía coordenadas geográficas, de los estudiantes aprobados en la Escuela Preparatoria de Cadetes Aéreos

Elton Weiglas de Paula¹

Universidade Federal de São João del-Rei (UFSJ), São João del-Rei, MG, Brasil

 <https://orcid.org/0000-0003-4848-6903>,  <http://lattes.cnpq.br/5040402898478825>

Carla Regina Guimarães Brighenti²

Universidade Federal de São João del-Rei (UFSJ), São João del-Rei, MG, Brasil

 <https://orcid.org/0000-0002-7822-3744>,  <http://lattes.cnpq.br/9456505192458557>

Resumo: A estatística circular tem aplicações em diversas áreas cujos dados possuam direções, ângulos ou tenham comportamento periódico, tais como: direção de migração de animais, dos ventos ou de materiais geológicos, horários de pico de trânsito ou de nascimento de inseto, além de coordenadas geográficas, entre outros. No entanto, em nosso cotidiano podemos avaliar instintivamente, de forma equivocada, dados com características angulares de forma clássica, utilizando a estatística descritiva linear. Assim, o objetivo deste trabalho foi estudar a teoria da estatística circular e comparar os resultados com a estatística descritiva linear, esta utilizada para dados na reta real e contemplada no ensino regular. Além disso, foi evidenciada a importância do estudo da estatística circular para dados angulares, juntamente com a apresentação de gráficos circulares, com o auxílio do software R. Para aplicação da teoria foi utilizado um conjunto de dados reais das coordenadas geográficas das cidades de origem de alunos da Escola Preparatória de Cadetes do Ar (EPCAR) comparando as medidas estatísticas de três diferentes anos de ingresso.

Palavras-chave: média direcional; desvio padrão circular; funções trigonométricas; boxplot circular.

Abstract: Circular statistics has applications in several areas whose data have directions, angles or periodic behavior, such as: direction of migration of animals, winds or geological materials, times of peak traffic or the

¹**Currículo sucinto:** Licenciado em Matemática pela Universidade Federal de São João del-Rei, especialista em Matemática Financeira e Estatística pela Faculdade Futura – SP e mestre em Ensino de Matemática pela Universidade Federal de São João del-Rei. **Contribuição de autoria:** Escrita – Primeira Redação, Escrita – Revisão e Edição, Investigação, Metodologia. **Contato:** eltonweiglas@yahoo.com.br.

²**Currículo sucinto:** Licenciatura em Matemática pela Universidade Tricordiana de Educação, licenciatura em Química pela Universidade Federal de São João del-Rei, mestrado em Agroquímica pela Universidade Federal de Viçosa, especialização em Matemática pela Universidade Federal de Viçosa, mestrado, doutorado e pós-doutorado em Estatística e Experimentação Agropecuária pela Universidade Federal de Lavras. **Contribuição de autoria:** Escrita – Revisão e Edição, Supervisão. **Contato:** carlabrighenti@ufsj.edu.br.



birth of an insect, in addition to geographic coordinates among others. However, in our daily lives, we can instinctively evaluate, in a wrong way, data with angular characteristics in a classical way, using linear descriptive statistics. Thus, the objective of this work was to study the theory of circular statistics and compare the results with linear descriptive statistics, which is used for data on the real line and contemplated in regular education. In addition, the importance of the study of circular statistics for angular data was highlighted, along with the presentation of circular graphs, with the aid of the software R. For the application of the theory, a set of real data of the geographical coordinates of the cities of origin of students of the Air Cadet Preparatory School (EPCAR) was used, comparing the statistical measures of three different years of entry.

Keywords: directional average; circular standard deviation; trigonometric functions; circular boxplot.

Resumen: La estadística circular tiene aplicaciones en varias áreas cuyos datos tienen direcciones, ángulos o tienen comportamiento periódico, tales como: dirección de migración de animales, vientos o materiales geológicos, horas pico de tránsito o nacimiento de insectos, así como coordenadas geográficas entre otras. Sin embargo, en nuestra vida cotidiana podemos evaluar intuitivamente, de manera incorrecta, datos con características angulares de una manera clásica, utilizando estadística descriptiva lineal. Por lo tanto, el objetivo de este trabajo fue estudiar la teoría de la estadística circular y comparar los resultados con la estadística descriptiva lineal, que se utiliza para los datos en la línea real y contemplada en la educación regular. Además, se destacó la importancia del estudio de la estadística circular para los datos angulares, junto con la presentación de gráficos circulares, con la ayuda del software R. Para la aplicación de la teoría, se utilizó un conjunto de datos reales de las coordenadas geográficas de las ciudades de origen de los estudiantes de la Escuela Preparatoria de Cadetes del Aire (EPCAR), comparando las medidas estadísticas de tres años diferentes de ingreso.

Palabras clave: media direccional; desviación estándar circular; funciones trigonométricas; diagrama de caja circular.

Data de submissão: 27 de março de 2023.

Data de aprovação: 12 de junho de 2023.

1 Introdução

Observa-se em diversas áreas do conhecimento um enorme interesse em estudar eventos que fornecem dados direcionais ou medidas angulares. De fato, quando estamos interessados em estudar fenômenos cujos dados possuem direções, tais como direção dos ventos, transporte de materiais geológicos, direção de migrações de animais, entre outros, a ferramenta ideal é denominada estatística circular (Jammalamadaka; SenGupta, 2001).



Os dados direcionais podem ser representados geometricamente por pontos em uma circunferência centrada na origem do sistema do plano cartesiano e pode-se também utilizar a representação em coordenadas polares. Observe que quando se fixa um determinado ponto P no plano cartesiano cujas coordenadas são (x,y) , esse ponto possui uma distância r em relação à origem. Ao assumir um vetor que esteja no eixo das abscissas e, ao rotacionar este vetor no sentido anti-horário até que o mesmo passe por P, o deslocamento formado entre o vetor e o eixo x definirá um ângulo θ . Assim, P poderá ser retratado como $(r \cos \theta, r \sin \theta)$, sendo possível analisar cada dado disponível de forma angular.

Constata-se ainda que, ao analisar dados direcionais nos quais o interesse está apenas em estudar a direção e não a magnitude do vetor, a análise é facilitada pela representação por meio de vetores de comprimento unitário. Em especial, um conjunto de dados que apresenta coordenadas geográficas pode ser transformado e analisado como medidas angulares. Pode-se notar, por exemplo, a estatística circular aplicada a dados geográficos nos trabalhos de Silva, Brighenti e Resende (2022) e Menezes(2018). Ressalta-se que pode existir uma predisposição em desejar manipular esses dados circulares com as ferramentas da estatística descritiva linear, o que resultaria em soluções não confiáveis.

O ensino da estatística descritiva linear está previsto na Base Nacional Curricular Comum (BNCC) no qual uma de suas competências para o Ensino Médio é: “Resolver e elaborar problemas, em diferentes contextos, que envolvem cálculo e interpretação das medidas de tendência central (média, moda, mediana) e das medidas de dispersão (amplitude, variância e desvio padrão)” (Brasil, 2018, p. 529). As medidas descritivas que são discutidas usualmente nas salas de aulas são as médias aritmética, geométrica e harmônica. No entanto, não há menção sobre as médias de dados angulares, apesar de sua extensa aplicabilidade. Diante do exposto, este trabalho visa apresentar o uso da estatística circular como ferramenta para análise de dados angulares e seu uso aplicado em um conjunto de dados reais dos alunos ingressos da Escola Preparatória de Cadetes do Ar (EPCAR).

Inicialmente, será apresentada uma revisão sobre os conceitos básicos no âmbito da estatística circular, tais como medidas descritivas, distribuição de probabilidade e teste de hipóteses, juntamente com a exibição de medidas de posição e dispersão advindas da estatística linear para fins de comparação entre as duas áreas de estudo. Posteriormente, esses conceitos serão utilizados em um estudo direcional sobre a procedência dos estudantes que ingressaram na EPCAR entre os anos de 2020 a 2022.



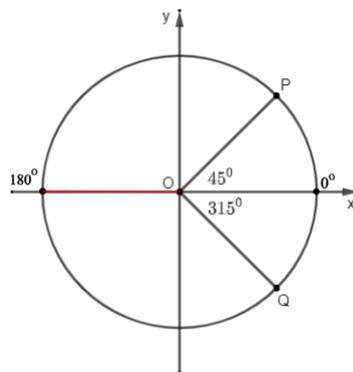
2 Estatística Descritiva Circular

A princípio, suponha dois dados angulares, como por exemplo $P = 45^\circ$ e $Q = 315^\circ$. Ao calcular a média aritmética usual desses dados obtém-se o seguinte resultado:

$$\bar{x} = \frac{45^\circ + 315^\circ}{2} = 180^\circ \quad (1)$$

Contudo, ao plotar tais dados na circunferência, nota-se que a média aritmética 180° não condiz com o esperado da Figura 1, sendo que o valor mais intuitivo seria a média igual a 0° . Assim percebe-se a necessidade de utilizar outro recurso estatístico para obter medidas mais condizentes.

Figura 1: Representação dos pontos P e Q



Fonte: Elaboração dos autores.

Seja (x, y) coordenadas de um ponto P no plano cartesiano; observe que este mesmo ponto $P = (x, y)$ pode ser representado sobre uma circunferência de centro $O = (0, 0)$ e raio r , pois a distância do ponto P ao centro O , denotada por $d(O, P) = r$, satisfaz as condições de P pertencer à circunferência. Note que $\overline{OP} = r$ e que o menor ângulo formado por \overline{OP} com o eixo x é θ , então aplicando as relações métricas de um triângulo retângulo obtemos

$$\cos(\theta) = \frac{x}{r} \Leftrightarrow x = r \cos \theta \quad \text{e} \quad \sin(\theta) = \frac{y}{r} \Leftrightarrow y = r \sin \theta \quad (2)$$

e assim o ponto $P = (x, y)$ pode ser representado por $P = (r \cos \theta, r \sin \theta)$. Para $r = 1$, tem-se que o ponto P pertence a um círculo unitário caracterizado por $P = (\cos \theta, \sin \theta)$.



2.1 Média Direcional

Considere uma amostra aleatória de tamanho n dada pelos pontos $P_1, P_2, P_3, \dots, P_n$ em um círculo unitário, e seus respectivos ângulos $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ em relação ao eixo x . Cada ponto P_i pode ser representado por um vetor $P_i = (\cos \theta_i, \sin \theta_i)$ em que o comprimento R do vetor resultante da soma dos n vetores unitários, será dado por

$$R = \sqrt{\left(\sum_{i=1}^n \cos \theta_i\right)^2 + \left(\sum_{i=1}^n \sin \theta_i\right)^2} \tag{3}$$

Assumindo $C = \sum_{i=1}^n \cos \theta_i$ e $S = \sum_{i=1}^n \sin \theta_i$, o comprimento do vetor resultante será descrito pela seguinte igualdade:

$$R = \sqrt{C^2 + S^2} \tag{4}$$

Em seguida, será apresentado o conceito da **Média Direcional** ($\bar{\theta}$), sendo que esta média tem a mesma direção que o centro de massa do círculo unitário (Mardia, 1972). O centro de massa será dado por $M = (\bar{C}, \bar{S})$, tal que

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i \tag{5}$$

$$\bar{S} = \frac{1}{n} \sum_{i=1}^n \sin \theta_i \tag{6}$$

Define-se também o **comprimento médio resultante** (\bar{R}) como sendo o tamanho médio do vetor resultante e este é obtido pela seguinte equação:

$$\bar{R}^2 = \bar{C}^2 + \bar{S}^2 \implies \bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}; \quad 0 \leq \bar{R} \leq 1, \text{ pois, } \bar{R} = \frac{R}{n} \tag{7}$$

O comprimento médio resultante é dito como uma medida de dispersão, pois ele fornece um grau de concentração dos ângulos observados e, quanto mais próximo de 1 o valor de \bar{R} está, menos dispersos estão os pontos. Por outro lado, quanto mais próximo de 0 for este valor \bar{R} , menos concentrados estão os dados (Silva; Brighenti; Resende, 2022). Em seguida, temos que a **Média Direcional** é definida como o ângulo ($\bar{\theta}$) e deve satisfazer as seguintes equações:

$$\begin{cases} \cos \bar{\theta} = \frac{\bar{C}}{\bar{R}} \\ \sin \bar{\theta} = \frac{\bar{S}}{\bar{R}} \end{cases} \tag{8}$$



Logo, o valor da média direcional será obtido pelas seguintes expressões:

$$\bar{\theta} = \begin{cases} \arctan \frac{\bar{S}}{\bar{C}}, & \text{se } \bar{S} > 0 \text{ e } \bar{C} > 0 \\ \arctan \frac{\bar{S}}{\bar{C}} + \pi, & \text{se } \bar{C} < 0 \\ \arctan \frac{\bar{S}}{\bar{C}} + 2\pi, & \text{se } \bar{S} < 0 \text{ e } \bar{C} > 0 \end{cases} \quad (9)$$

Para exemplificar essa parte teórica e retomando o exemplo fictício citado anteriormente com dois dados angulares, segue que sua média circular pode ser encontrada por meio dos seguintes passos: primeiramente, calcula-se o centro de massa M dos pontos $P = 45^\circ$ e $Q = 315^\circ$ segundo as equações (5) e (6).

$$\bar{C} = \frac{1}{2} \sum_{i=1}^2 \cos \theta_i = \frac{1}{2} (\cos (45^\circ) + \cos (315^\circ)) = \frac{1}{2} \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \right) = \frac{\sqrt{2}}{2} \quad (10)$$

$$\bar{S} = \frac{1}{2} \sum_{i=1}^2 \sin \theta_i = \frac{1}{2} (\sin (45^\circ) + \sin (315^\circ)) = \frac{1}{2} \left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \right) = 0 \quad (11)$$

Desse modo, o comprimento médio do vetor resultante, conforme (7), é

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} = \sqrt{\left(\frac{\sqrt{2}}{2}\right)^2 + (0)^2} = \frac{\sqrt{2}}{2} \quad (12)$$

e, assim, a média direcional (8) será

$$\begin{cases} \cos \bar{\theta} = \frac{\bar{C}}{\bar{R}} = \frac{\frac{\sqrt{2}}{2}}{\frac{\sqrt{2}}{2}} = 1 \\ \sin \bar{\theta} = \frac{\bar{S}}{\bar{R}} = \frac{0}{\frac{\sqrt{2}}{2}} = 0 \end{cases} \implies \bar{\theta} = \arctan \frac{\bar{S}}{\bar{C}} = \frac{0}{1} = 0^\circ \quad (13)$$

Logo, a média circular $\bar{\theta}$ dos ângulos 45° e 315° é igual a 0° . Para ampliar o entendimento do assunto, apresenta-se a seguir um exemplo adaptado de Ayres *et al.* (2007), em que os dados possuem como característica um comportamento angular. *Exemplo:* Em colméias de abelhas foi realizado um estudo sobre a direção do voo nupcial de 11 rainhas, em relação à entrada da colméia, obtendo-se os resultados conforme Tabela 1. Em seguida, estão plotados a direção do voo de cada uma dessas rainhas bem como a direção da média aritmética linear (Figura 2); e também o gráfico de pontos juntamente com a apresentação do vetor indicando a direção média circular dos dados (Figura 3).

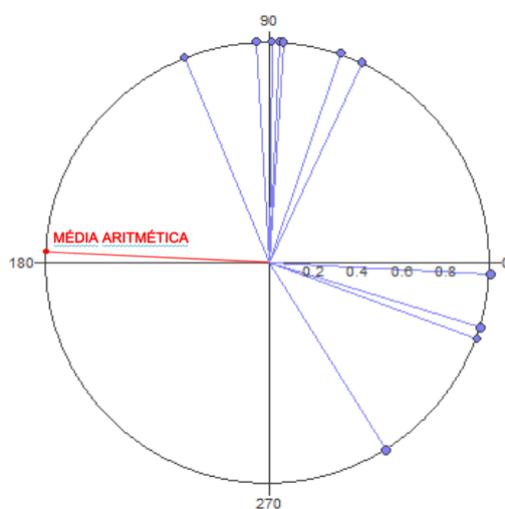


Tabela 1: Direção do voo nupcial, em graus, das 11 rainhas

| Rainha | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---------|----|-----|----|-----|----|-----|-----|----|----|-----|----|
| Ângulos | 87 | 343 | 71 | 112 | 86 | 340 | 357 | 65 | 93 | 302 | 89 |

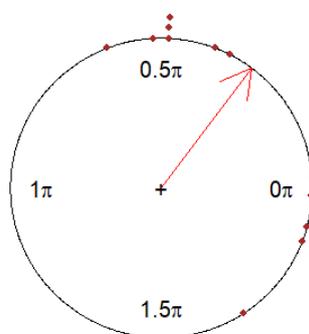
Fonte: Adaptado de Ayres *et al.* (2007).

Figura 2: Direção do voo das 11 rainhas na circunferência e da direção da média aritmética linear calculada



Fonte: Elaboração dos autores.

Figura 3: Gráfico de pontos relacionados a direção do voo das 11 rainhas, bem como a representação da direção média angular calculada



Fonte: Elaboração dos autores.



A média aritmética neste caso será dada por:

$$\bar{x} = \frac{87^\circ + 343^\circ + 71^\circ + 112^\circ + 86^\circ + 340^\circ + 357^\circ + 65^\circ + 93^\circ + 302^\circ + 89^\circ}{11} \approx 176,82^\circ \quad (14)$$

o que não reflete adequadamente a situação, pois o valor encontrado de $176,82^\circ$ está bem distante da realidade plotada dos voos das abelhas apresentados na Figura 2. No entanto, com o uso da estatística circular, tem-se como resultado os seguintes valores:

1 – Média dos cossenos obtida por (5):

$$\bar{C} = \frac{1}{11}(\cos 87^\circ + \cos 343^\circ + \cos 71^\circ + \dots + \cos 89^\circ) = 0,353212261 \quad (15)$$

2 – Média dos senos dada por (6):

$$\bar{S} = \frac{1}{11}(\sin 87^\circ + \sin 343^\circ + \sin 71^\circ + \dots + \sin 89^\circ) = 0,476264102 \quad (16)$$

3 – Média angular de (9):

$$\bar{\theta} = \arctan\left(\frac{0,476264102}{0,353212261}\right) = 53,4382^\circ \quad (17)$$

4 – Comprimento do vetor médio resultante em (7):

$$\bar{R} = \sqrt{(0,353212261)^2 + (0,476264102)^2} \approx 0,5929 \quad (18)$$

Nota-se neste exemplo que a média direcional ou angular do voo das 11 rainhas é de aproximadamente $53,44^\circ$, graficamente representada na Figura 3, e uma possível análise preliminar desses dados com apenas esta parte teórica apresentada, mostra que existe uma certa dispersão das direções, pois o comprimento médio do vetor resultante foi de $\bar{R} = 0,5929$, um valor que exprime relativamente uma heterogeneidade das direções dos voos.

2.2 Mediana Circular

A mediana direcional β dos ângulos $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ é definida como um ângulo Φ que satisfaz as seguintes condições (Mardia, 1972): a metade dos dados estejam no arco $[\Phi, \Phi + \pi)$ e a maioria dos dados são mais próximos de Φ do que $\Phi + \pi$.

Quando o número de elementos do espaço amostral n é ímpar, a direção da mediana é um dos dados, ou seja, será um único ângulo, mais precisamente, o dado da $\frac{n+1}{2}$ – ésima posição.

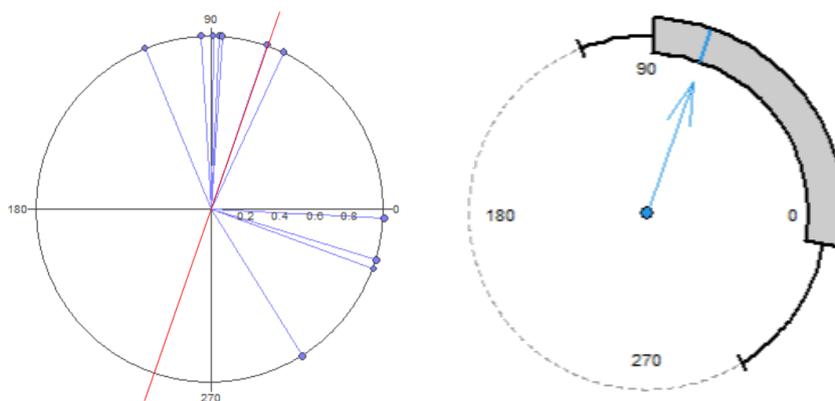


Quando n é par, a direção da mediana é a média aritmética angular entre os dois ângulos centrais do espaço amostral. Com relação ao exemplo das abelhas da Tabela 1, se a análise da mediana for feita usando a estatística linear, tem-se que o conjunto de dados possui 11 elementos, ou seja, n é ímpar e, ordenando-os por ordem crescente, obtem-se como resultado o valor que ocupa a 6ª posição desta amostra. De fato,

$$(65^\circ \ 71^\circ \ 86^\circ \ 87^\circ \ 89^\circ \ 93^\circ \ 112^\circ \ 302^\circ \ 340^\circ \ 343^\circ \ 357^\circ)$$

Portanto, a mediana (Me) é igual a 93° . O que não condiz com a direção esperada considerando a Figura 2. Em contrapartida, embasado na estatística circular, a direção da mediana Φ tem que satisfazer as exigências, de que a metade dos dados estejam no arco $[\Phi, \Phi + \pi)$ e a maioria dos dados estejam mais próximos de Φ do que $\Phi + \pi$. Desse modo, a mediana direcional será $\Phi = 71^\circ$, pois, a metade dos dados estão contidos no intervalo $[71^\circ, 251^\circ)$ e a maioria dos dados estão mais próximos de 71° do que do ângulo 251° . Graficamente, será representado um gráfico de pontos angulares das abelhas rainhas, de modo que a semirreta em vermelho contém o valor da mediana angular, sem contar que esta semirreta divide a amostra ao meio; ao lado deste, o denominado Boxplot circular, sendo este último um gráfico que tem como característica transmitir a ideia de posição, dispersão e assimetria dos dados, bem como a posição da mediana que está representada direcionalmente pelo vetor em azul. Vale ressaltar que a análise do Boxplot circular é semelhante ao Boxplot contemplado no Ensino Médio; de fato a caixa em cinza, conhecida como intervalo interquartil, contempla 50% dos dados sem contar que é uma ferramenta para avaliar o grau da dispersão dos dados e as caldas representam os limites superiores e inferiores do gráfico.

Figura 4: Mediana direcional do voo nupcial, em graus, das 11 rainhas, representada pelo gráfico dos pontos angulares e pelo Boxplot circular



Fonte: Elaboração dos autores.



2.3 Medidas de Dispersão

A **variância circular** amostral (V) é uma medida de dispersão circular definida por:

$$V = 1 - \bar{R}; \quad 0 \leq V \leq 1 \quad (19)$$

Nesse sentido, quanto menor for o valor de V , maior será a concentração dos dados. Análise análoga à variância utilizada em dados lineares (Silva; Brighenti; Resende, 2022). Novamente, considerando o exemplo das abelhas (1), observe que ao calcular a variância de forma linear tem-se um valor bem distante dos dados estudados:

$$s^2 = \sum_{i=1}^{11} \frac{(x_i - \bar{x})^2}{10} = \frac{(87^\circ - 176,82^\circ)^2 + \dots + (357^\circ - 176,82^\circ)^2}{10} \approx 4175,6 \quad (20)$$

De acordo com a definição (19) de variância para dados angulares e utilizando do resultado do comprimento do vetor resultante em (18), tem-se o seguinte:

$$V = 1 - \bar{R} \Rightarrow V = 1 - 0,5929 = 0,4071 \quad (21)$$

Segundo Mardia (1972), define-se como **Desvio Padrão Circular** a seguinte relação:

$$v = [-2 \ln(1 - V)]^{\frac{1}{2}} \Leftrightarrow v = [-2 \ln(\bar{R})]^{\frac{1}{2}} \in [0, \infty) \quad (22)$$

Assim, os cálculos do desvio para o Exemplo 1, são dados a seguir. Para a estatística linear tem-se:

$$s = \sqrt{s^2} \Rightarrow s = \sqrt{(4175,6)} \approx 64,62^\circ \quad (23)$$

Com base na definição de desvio padrão angular (22), observa-se o seguinte:

$$v = [-2 \ln(1 - V)]^{\frac{1}{2}} \Rightarrow v = \sqrt{-2 \ln(1 - 0,4071)} \approx 1,0225 \text{ rad} \approx 58,59^\circ \quad (24)$$

Dessa maneira, para fins de comparação e análise dos dados até aqui calculados, pode-se ver, conforme a Tabela 2, que os dados angulares e lineares das respectivas medidas analisadas são bem distintas.



Tabela 2: Comparação dos resultados do exemplo do voo das 11 abelhas rainhas obtidos por meio da estatística linear e angular

| Estadística | Média | Mediana | Desvio padrão |
|-------------|----------|---------|---------------|
| Angular | 53,4382° | 71° | 58,59° |
| Linear | 176,82° | 93° | 64,62° |

Fonte: Elaboração dos autores.

Portanto, verifica-se que quando se quer estudar dados angulares é necessário efetuar um estudo mais cauteloso e embasado na estatística descritiva circular, visto que a estatística usual fornece valores totalmente distintos dos valores reais e bem distantes da interpretação exigida em um estudo estatístico.

3 Distribuição de Probabilidade Circular

As propriedades das funções de distribuição circular e as funções de densidade de probabilidade possuem muitos aspectos em comum com as funções que são discutidas na estatística linear (Pewsey; Neuhauser; Ruxton, 2013). No entanto, há também importantes diferenças que decorrem na necessidade de se definir as distribuições advindas de dados circulares.

Uma distribuição circular é uma distribuição de probabilidade cuja probabilidade total está concentrada na circunferência em um círculo unitário. Uma vez que cada ponto da circunferência representa uma direção, tal distribuição é uma forma de atribuir probabilidades para direções diferentes ou definindo uma distribuição direcional (Jammalamadaka; SenGupta, 2001, p. 38).

Em especial, as funções de densidade de probabilidade $f(\lambda)$ existem se, e só se, as seguintes propriedades forem satisfeitas (Mardia, 1972):

1. $f(\lambda) \geq 0$;
2. $f(\lambda + 2k\pi) = f(\lambda)$ para $k \in \mathbb{Z}$;
3. $\int_0^{2\pi} f(\lambda) d\lambda = 1$.

As distribuições circulares, essencialmente as distribuições que relacionam as variáveis contínuas e as massas de probabilidades, são chamadas de distribuições contínuas. Para as distri-



buições contínuas existem vários modelos probabilísticos circulares, dos quais pode-se destacar a distribuição uniforme, a von Mises, a Cardioide, a Arqueadas (Wrapped), entre outras. Neste tópico, será discutido, de forma bem sucinta, dentre os vários modelos probabilísticos circulares, a distribuição circular uniforme, conhecida como a mais simples, e a distribuição de von Mises, dita como a mais importante. Para melhores detalhes, consultar Mardia (1972) e Jammalamadaka e SenGupta (2001).

3.1 Distribuição Uniforme Circular

A função de densidade de probabilidade uniforme circular é definida como:

$$f(\lambda) = \frac{1}{2\pi}; 0 \leq \lambda \leq 2\pi \tag{25}$$

A distribuição uniforme é dita como a mais básica no círculo e é a única distribuição no círculo que é invariante sob rotação e reflexão. Haja vista que na distribuição uniforme todas as direções são igualmente prováveis, ou seja, não existe uma direção preferencial e, além disso, o comprimento do vetor resultante é sempre igual a zero (Mardia, 1972).

3.2 Distribuição de von Mises

Diante da semelhança com a distribuição normal (distribuição de Gauss) da estatística linear, a distribuição de von Mises é considerada como uma das mais importantes distribuições de probabilidade circular, visto que muitas de suas propriedades analisadas, tais como, simetria, parâmetros e a própria regularidade, são examinadas de forma análoga à estatística linear; assim, a distribuição de von Mises – $vM(\bar{\theta}, \kappa)$ será definida com a seguinte função de densidade de probabilidade:

$$f(\lambda; \bar{\theta}, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\lambda - \bar{\theta})} \tag{26}$$

onde $\bar{\theta}$ é a direção média, $0 \leq \bar{\theta} < 2\pi$ e κ é o parâmetro de concentração, com $\kappa \geq 0$. Tem-se, ainda, que $I_0(\kappa)$ representa a função de Bessel ¹ modificada de primeiro tipo e ordem zero, dada pela seguinte igualdade:

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos(\lambda)} d\lambda = \sum_{r=0}^{\infty} \left(\frac{\kappa}{2}\right)^{2r} \left(\frac{1}{r!}\right)^2 \tag{27}$$

¹A função de Bessel foi definida primeiramente por Daniel Bernoulli, sendo obtida a partir da solução de uma equação diferencial ordinária de segunda ordem.



Uma observação importante em relação ao parâmetro de concentração κ e à função (26) é que quando $\kappa \rightarrow 0$, a função densidade de von Mises se aproxima da distribuição circular uniforme; por outro lado, quando $\kappa \rightarrow \infty$, menor será a variabilidade dos dados (Jammalamadaka; SenGupta, 2001).

4 Testes de Hipóteses

Em paralelo à estatística linear, a estatística circular possui alguns testes de hipóteses como, por exemplo, o teste de Rao, o teste de Rayleigh, o teste de Watson e o teste de Kuiper. A principal finalidade desses testes é estabelecer a qualidade de ajustamento dentre as distribuições de probabilidade já mencionadas neste trabalho (Góis, 2002). Mas, especificamente, o teste que será evidenciado nesta seção será o teste de Rayleigh. O teste de Rayleigh é usado para verificar se uma população está distribuída uniformemente em torno da circunferência. A hipótese nula (H_0) afirma que a população apresenta uma distribuição circular uniforme; já a hipótese alternativa (H_1), por sua vez, afirma que a população não é uniformemente distribuída em torno da circunferência e, portanto, existe uma direção preferida. O cálculo aplicado neste teste tem como base um valor Z cuja sua determinação é dada pela seguinte expressão:

$$Z = n\bar{R}^2 \quad (28)$$

onde n é o número de observações dos dados.

Sabe-se que \bar{R} , de certa forma, transmite a ideia de concentração dos dados. Quando a amostra segue uma distribuição uniforme circular, normalmente o valor do \bar{R} é pequeno; e quando o valor do vetor médio resultante assume valores próximos a 1, a hipótese nula é rejeitada. Assim, quanto maior for o vetor médio resultante, maior será o valor de Z , conforme (28), o que significa uma maior concentração dos dados em torno da média, resultando em uma menor probabilidade destes dados estarem uniformemente distribuídos (Góis, 2002). A equação que determina o cálculo da probabilidade Z é descrita pela seguinte igualdade:

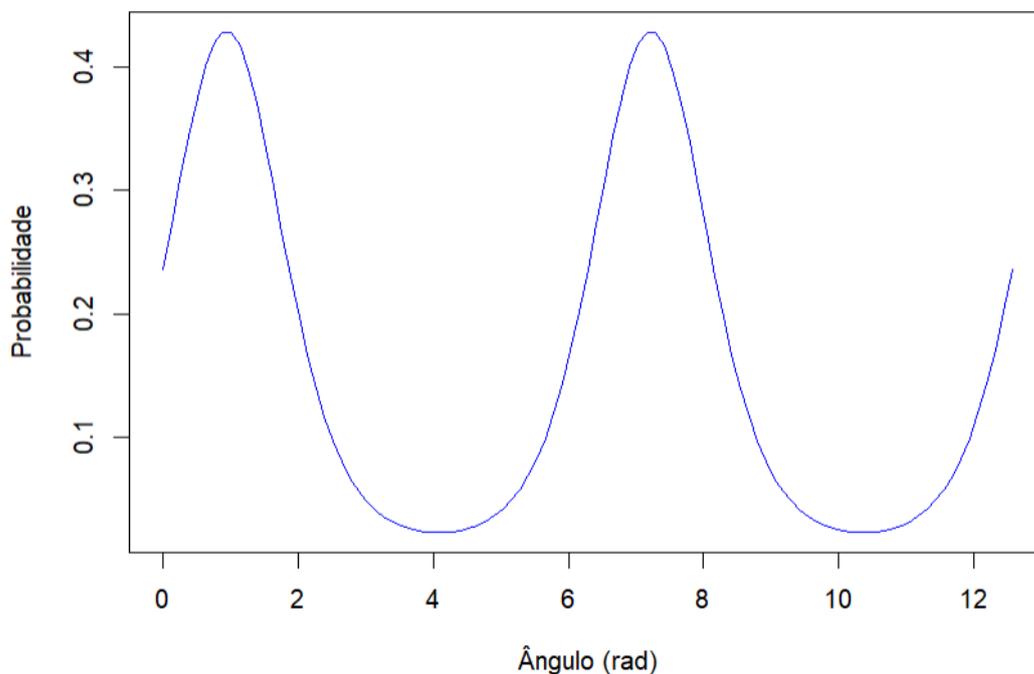
$$P(Z \geq z) = e^{\left[\sqrt{1+4n+4(n^2-nz)}-(1+2n)\right]} \quad (29)$$

Normalmente, o nível de significância escolhido é 0,05; com isso uma probabilidade menor a este valor indica que os dados não estão distribuídos uniformemente, evidenciando uma direção tendenciosa dos dados. Considerando os resultados obtidos da Tabela (1), relacionados ao exemplo



do voo nupcial das 11 rainhas, embasados na estatística circular, tem-se que, de acordo com o teste de Rayleigh, o valor Z calculado é de aproximadamente 3,87. Além disso, a probabilidade calculada, segundo a equação (29), resulta em um valor-p igual a 0,0172, um valor inferior ao nível de significância adotado de 0,05; portanto, os voos das rainhas não estão distribuídos uniformemente em torno do círculo. Logo, a distribuição de probabilidade que modela este exemplo é von Mises $(0,9326729; 1,480571)$. Graficamente, nota-se a distribuição que consta na Figura 5.

Figura 5: Gráfico da função de probabilidade von Mises das 11 rainhas



Fonte: Elaboração dos autores.

Em relação à curva de densidade de von Mises das 11 abelhas rainhas (Figura 5), juntamente com o Boxplot circular (Figura 4), ao analisar o gráfico de forma semelhante à forma realizada na curva de distribuição normal gaussiana, nota-se uma assimetria em torno das medidas de posição, $\bar{\theta} = 0,9326 \text{ rad} \approx 53,44^\circ$ e $\Phi = 1,2391 \text{ rad} = 71^\circ$, que representam, principalmente, a direção média dos voos nupciais das abelhas. Mas, percebe-se ainda, que existe uma leve dispersão dos dados, pois o valor do desvio padrão calculado foi de $v = 1,0225 \text{ rad} = 58,59^\circ$, o que reafirma o que foi citado nas conclusões da equação (18) relacionadas ao vetor médio resultante ($\bar{R} = 0,5929$) do exemplo das abelhas.



5 Uso da Estatística Circular para Dados Geográficos

Dados angulares podem ser obtidos de dados geográficos. Segundo Falconi (2009), para se obter as distâncias entre dois pontos geográficos, é necessário que se obtenha a diferença entre as medidas de latitude (DLA) e a diferença de longitude (DLO) das coordenadas geográficas desses pontos; a partir daí, é preciso que os valores angulares obtidos da DLA e DLO sejam transformados em distâncias. As distâncias relacionadas às diferenças de latitude e de longitude são dadas pela seguinte expressão:

$$\text{DLA ou DLO} = \left[g(60) + \text{min} + \frac{s}{60} \right] (1.852) \quad (30)$$

tal que as constantes presentes g , min e s são as medidas de graus, minutos e segundos, respectivamente. É importante frisar que 1 minuto da circunferência terrestre equivale a uma milha náutica (NM), e uma milha náutica equivale a 1.852 metros. Com o uso das relações trigonométricas entre as localidades, o ângulo (a) será encontrado segundo a Equação (31):

$$\tan(a) = \frac{\text{DLA}}{\text{DLO}} \Rightarrow a = \arctan\left(\frac{\text{DLA}}{\text{DLO}}\right) \quad (31)$$

A partir deste princípio, podem ser calculados os ângulos e distâncias entre duas cidades de interesse. Na cidade de Barbacena, Estado de Minas Gerais, está localizada a Escola Preparatória de Cadetes do Ar (EPCAR) que é uma organização de ensino do comando da Aeronáutica cuja missão é preparar seus alunos para o ingresso no Curso de Formação de Oficiais Aviadores (CFOAV) na Academia da Força Aérea (AFA). Neste artigo, serão abordados dados reais relacionados ao perfil dos alunos ingressantes no Curso Preparatório de Cadetes do Ar (CPCAR) desta instituição. Os dados referem-se aos estados de origem dos alunos aprovados nas turmas denominadas Anhur, Hércules e Cavok. Este conjunto de dados será utilizado para verificar a dispersão dos ingressantes nesta escola, haja vista que este estabelecimento possui um concurso a nível federal e uma relação candidato/vaga que girou em torno de 75 a 135 candidatos por vaga (2020 a 2022), ou seja, um concurso bem disputado.

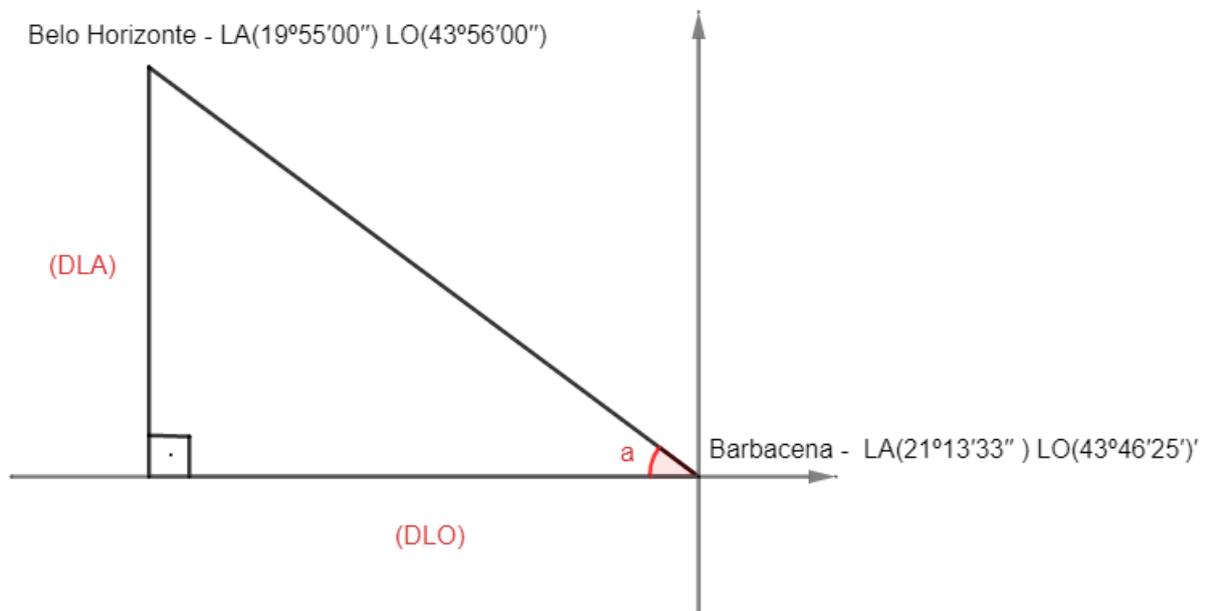
Inicialmente, foi realizado o levantamento dos estados de origem de cada aluno dos três esquadrões e assumiu-se como coordenadas geográficas de localização, latitude (LA) e longitude (LO), as coordenadas relacionadas as suas respectivas capitais, disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE)². Para determinar as distâncias e as direções angulares, foi fixada a cidade de Barbacena/MG, sede da EPCAR, como ponto referencial, ou seja, Barbacena

²Banco de Dados Geodésicos do IBGE. Disponível em: <http://www.bdg.ibge.gov.br/appbdg/>.



foi considerada como o ponto de origem de um sistema de eixos cartesianos, e as demais capitais como sendo as coordenadas geográficas de origem dos estados que possuem ao menos um aprovado no concurso. Para exemplificar essa situação, serão expostos os cálculos referentes aos alunos oriundos do estado de Minas Gerais, considerando as coordenadas geográficas da cidade de Belo Horizonte.

Figura 6: Coordenadas geográficas de Barbacena e Belo Horizonte em um sistema de eixos cartesianos



Fonte: Elaboração dos autores.

De fato, analisando o exemplo da Figura 6, tem-se:

- Barbacena

$$LA = 21^{\circ}13'33'' = 1.273,55 \text{ min} \text{ e } LO = 43^{\circ}46'25'' \approx 2.626,42 \text{ min}$$

- Belo Horizonte

$$LA = 19^{\circ}55'00'' = 1.195 \text{ min} \text{ e } LO = 43^{\circ}56'00'' = 2.636 \text{ min}$$

- $DLA = 1.273,55 - 1.195 = 78,55 \Rightarrow (78,55)(1.852) = 145.474,6$

- $DLO = 2.626,42 - 2.636 = -9,58 \Rightarrow (-9,58)(1.852) = -17.742,16$



Conforme equação (31),

$$\tan(\mathbf{a}) = \frac{\text{DLA}}{\text{DLO}} = \frac{145.474,6}{-17.742,16} \Rightarrow \mathbf{a} = \arctan\left(\frac{145.474,6}{-17.742,16}\right) \approx -83,0465 \quad (32)$$

O ângulo pertence ao 2º quadrante, logo, de (9),

$$a = -83,0465^\circ + 180^\circ \approx 97^\circ$$

Portanto, segue que a direção angular de Belo horizonte em relação à cidade de Barbacena será de 97° . Para o cálculo de distância entre as localidades, leva-se em consideração a distância em linha reta das cidades. Deste modo, com suporte nas relações trigonométricas, ao aplicar o teorema de Pitágoras, segue o valor da distância entre Barbacena e Belo Horizonte:

$$\begin{aligned} d^2 &= \text{DLA}^2 + \text{DLO}^2 \\ &= (145.474,6)^2 + (-17.742,16)^2 \implies \\ d &\approx 147.000 \text{ metros} \approx 147 \text{ km} \end{aligned} \quad (33)$$

Repetindo esse procedimento para as demais localidades, juntamente com o auxílio do software R (2021) e seus pacotes Circular e CircStats em busca de suporte matemático para a exibição dos resultados, foram calculados os demais ângulos e distâncias entre Barbacena e as capitais dos vários estados contemplados com pelo menos um aluno aprovado no CPCAR, como se pode observar na Tabela 3. Posteriormente a isso, serão apresentados resultados relacionados às medidas de posição e dispersão embasados de acordo com a teoria da estatística circular e linear, a fins de análise e comparação entre resultados obtidos (Tabela 4). E, também, serão mostrados na Figura 7 os gráficos denominados como diagrama de rosas e os Boxplot angular e linear, respectivamente.



Tabela 3: Estados de origem dos alunos da EPCAR dos anos de 2020 a 2022 (Anhur, Hércules e Cavok) e suas respectivas coordenadas geográficas considerando as distâncias das capitais em relação à Barbacena/ MG

| Estado | Anhur | Hércules | Cavok | Latitude | Longitude | Direção (°) | Distância(km) |
|--------------|-------|----------|-------|-----------|-----------|-------------|---------------|
| Barbacena/MG | | | | 21°13'33" | 43°46'25" | 0 | 0 |
| AL | - | 1 | - | 9°39'57" | 35°44'06" | 55,2 | 1565 |
| AM | - | - | 1 | 3°4'25" | 60°0'20" | 130,7 | 2662 |
| BA | - | 2 | 4 | 12°58'13" | 38°30'45" | 57,5 | 1088 |
| CE | 1 | 3 | 1 | 3°43'06" | 38°32'34" | 73,4 | 2031 |
| DF | 6 | 5 | 4 | 15°47'38" | 47°52'58" | 127,1 | 757 |
| ES | 1 | 1 | 2 | 20°19'08" | 40°20'16" | 14,8 | 395 |
| GO | 1 | 3 | 3 | 16°40'00" | 49°15'00" | 140,3 | 792 |
| MG | 27 | 25 | 22 | 19°55'00" | 43°56'00" | 97,0 | 147 |
| MS | 2 | 2 | 1 | 20°28'53" | 54°36'58" | 176,1 | 1208 |
| MT | 1 | - | - | 15°35'45" | 56°05'49" | 155,5 | 1506 |
| PA | 1 | - | 2 | 1°27'21" | 48°30'14" | 103,5 | 2259 |
| PB | - | 4 | 1 | 7°07'08" | 34°52'53" | 57,8 | 1853 |
| PE | 6 | 5 | 1 | 8°03'14" | 34°52'51" | 56,0 | 1766 |
| PI | - | 1 | - | 5°05'20" | 42°48'07" | 86,6 | 1797 |
| PR | 5 | 4 | 2 | 25°25'47" | 49°16'19" | 217,4 | 769 |
| RJ | 82 | 108 | 50 | 22°54'10" | 43°12'28" | 288,7 | 197 |
| RN | 6 | 8 | 4 | 5°47'42" | 35°12'32" | 61,0 | 1961 |
| RO | - | 1 | 1 | 8°45'43" | 63°54'14" | 148,2 | 2631 |
| RR | - | 1 | 2 | 2°49'12" | 60°40'19" | 132,6 | 2777 |
| RS | 3 | 5 | 5 | 30°01'58" | 51°13'48" | 229,8 | 1283 |
| SC | 2 | 2 | - | 27°35'49" | 48°32'56" | 233,2 | 885 |
| SE | - | 1 | 1 | 10°54'36" | 37°04'12" | 57,0 | 1367 |
| SP | 24 | 33 | 19 | 23°33'01" | 46°38'02" | 219,1 | 410 |
| Total | 168 | 215 | 126 | - | - | - | |

Fonte: Elaboração dos autores.



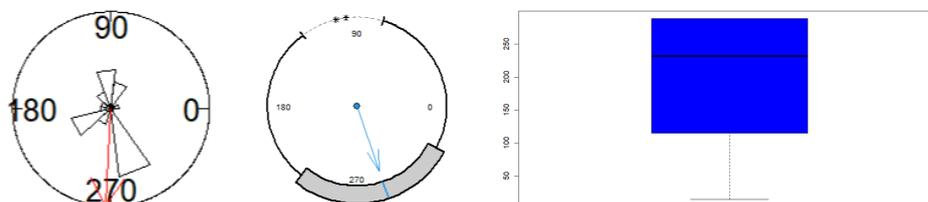
Tabela 4: Parâmetros dos dados circulares e lineares das turmas Anhur(A), Hércules(H) e Cavok(C)

| Turma | Circulares - (radianos) | | | | | Lineares | | | |
|-------|-------------------------|-------------|------------------|-------------|------------------|-----------|--------|-------|-------|
| | $\bar{\theta}$ | \bar{R} | Φ | V | v | \bar{x} | Me | s^2 | s |
| A | 267,2° (-1,62) | - (0,33) | 288,7° (5,04) | - (0,68) | 85,9° (1,50) | 214,9° | 231,5° | 319 | 17,9° |
| H | 271,8° (-1,54) | - (0,36) | 288,7° (5,04) | - (0,64) | 82,2° (1,43) | 216,5° | 288,7° | 253,5 | 15,9° |
| C | 251,5° (-1,89) | - (0,18) | 288,7° (5,04) | - (0,82) | 105,7° (1,84) | 197,8° | 219,1° | 374,1 | 19,3° |

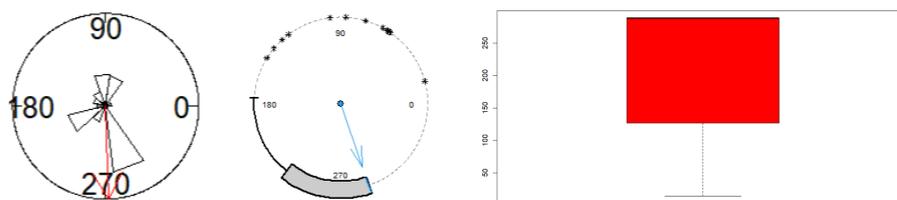
Fonte: Elaboração dos autores

Figura 7: Diagrama de rosas, Boxplots circular e linear das turmas ingressantes nos anos de 2020, 2021 e 2022, respectivamente

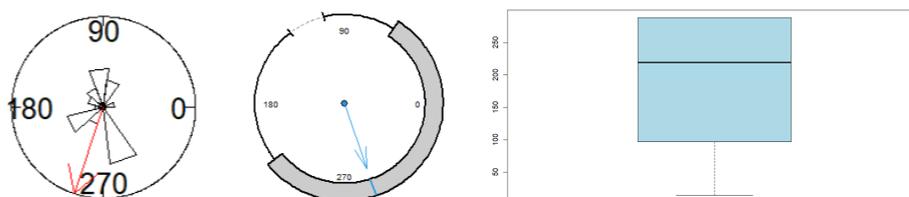
Ingressantes em 2020 - 3º ano.



Ingressantes em 2021 - 2º ano.



Ingressantes em 2022 - 1º ano.



Fonte: Elaboração dos autores.



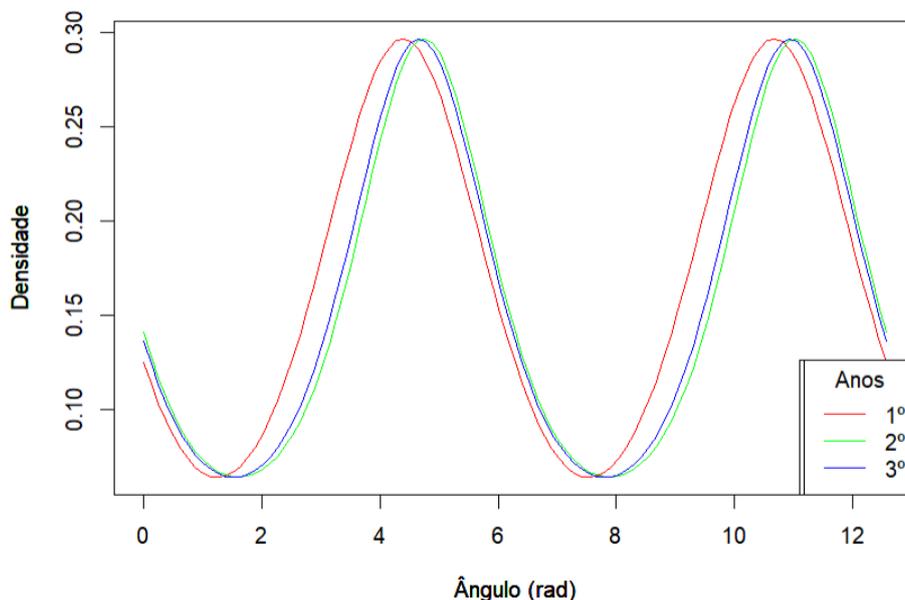
Em relação à Figura 7, estão listados os gráficos das respectivas turmas: Anhur, Hércules e Cavok. Dentre estes gráficos, está o diagrama de rosas, um gráfico que se assemelha ao histograma da estatística linear; em especial, o diagrama de rosas apresenta em sua essência as frequências e a concentração das vagas ocupadas em relação às direções angulares. E, a fins de comparação entre a estatística circular e a estatística linear, estão listados os gráficos denominados Boxplot angular e linear; uma análise que se nota entre estes gráficos é o fato que, se o estudo fosse embasado pela estatística linear, os dados apresentariam uma dispersão muito mais alta que de fato ocorre na distribuição das vagas, ou seja, o Boxplot linear aparenta uma dispersão das vagas muito maior que a realidade, sem contar que, nesta análise, não existe nenhum ponto discrepante nos dados das turmas Hércules e Cavok, o que não é verdade segundo os Boxplot circulares dos respectivos anos.

Para que se possa ter condições de argumentação em relação ao ingresso dos alunos no concurso do CPCAR nos diferentes anos analisados até aqui, foi aplicado o teste de Rayleigh a fim de verificar se de fato os dados analisados possuem uma distribuição uniforme em torno de uma circunferência. A partir deste teste, verificou-se a rejeição da hipótese nula (H_0), pois, segundo os resultados dos valores-p calculados das turmas Anhur, Hércules e Cavok, respectivamente, todos foram inferiores ao nível de significância adotado de 0,05; conseqüentemente, considerou-se a hipótese alternativa (H_1), que afirma que os dados possuem uma tendência direcional, ou seja, os dados possuem uma distribuição von Mises, o que nos dá suporte para um melhor ajuste das distribuições de probabilidade.

Diante dos fatos, foi realizado o ajuste dos dados conforme as definições de distribuição de probabilidade de von Mises, o que confirmou uma periodicidade direcional em relação aos alunos aprovados no concurso de admissão da EPCAR. Percebe-se claramente no gráfico da Figura 8 que nos anos de 2020, 2021 e 2022, ingressantes avaliados neste trabalho, tanto a média direcional quanto a mediana para dados circulares são valores entre 4 e 5 radianos, algo em torno de 4,7 rad aproximadamente, valores que remetem à direção angular do estado do Rio de Janeiro, o que confirma os cálculos apresentados na Tabela 4; observa-se, ainda, uma predominância das vagas ocupadas no concurso da região sudeste, algo em torno de 70% e 80%, valores muito significativos e dominantes.



Figura 8: Gráfico da função de probabilidade de von Mises das 3 turmas



Fonte: Elaboração dos autores.

Um possível destaque que vale a pena enfatizar neste momento, é o fato que neste período de 2020 a 2022 ocorreu a pandemia causada pelo vírus SARS-CoV-2 (COVID-19). O vírus foi identificado pela primeira vez em Wuhan, na China, em 31 de dezembro de 2019, e, desde então, se espalhou globalmente, causando uma pandemia que afetou milhões de pessoas em todo o mundo (Sanar, 2020).

Com base nisso, o terceiro ano, turma Anhur incorporada em 2020, realizou o exame de admissão em um período sem pandemia, em condições normais, visto que a prova foi realizada em 07/07/2019. Já a turma Hércules, ingressante em 2021, realizou a prova no auge da pandemia; de fato, em 11/03/2020 a Organização Mundial de Saúde (OMS) declarou o início da pandemia e em 20/03/2020 o Ministério da Saúde declarou distanciamento social, o que impactou totalmente a rotina de estudos nas escolas, ou seja, foi necessário a implementação de modelos de aulas até então distantes da rotina e com realidades bem diferentes em relação aos anos anteriores; observa-se ainda que, normalmente, as provas da EPCAR são aplicadas nas últimas semanas de junho, ou nas primeiras semanas de julho, porém, neste período, o Brasil já registrava um total de 64.365 mortes, o que ocasionou adiamento do exame para 18/10/2020.



Em relação ao primeiro ano, a turma Cavok, ingressante em 2022, fez a prova do concurso em 04/07/2021, ano em que já havia uma esperança de controle da doença devido ao início da campanha de vacinação em 17/01/2021, mas mesmo assim a pandemia ainda estava causando muitas mortes; por exemplo, em 26/04/2021, o número de óbitos já ultrapassava o total de mortes do ano de 2020 e, um pouco antes do dia do exame, o Brasil já se aproximava de 510 mil mortes pela doença.

Assim, com base nessas informações e, principalmente, nos resultados dos Boxplots circulares (7) e no gráfico de função de densidade de probabilidade von Mises (8) relacionados às três turmas, tem-se que estes gráficos além de apresentarem informações relacionadas à existência de uma assimetria nas vagas ocupadas, mostram uma concentração regional nas distribuições dos ingressantes em relação ao país. O principal destaque vai para a turma que ingressou em 2021, em se tratando de um concurso nacional, nesta turma a pandemia influenciou negativamente no preenchimento das vagas disponibilizadas; as vagas preenchidas ficaram ainda mais concentradas e com muito mais pontos discrepantes. Porém, a turma que entrou em 2022, também na pandemia, apresentou uma maior dispersão, o que nos remete em uma entrada que contemplou alunos oriundos de outros estados que até então não possuíam um alto índice de aprovação. Mas, mesmo assim, nota-se que os exames realizados em anos com contextos nacionais e sanitários totalmente distintos, ainda sim, manteve-se a periodicidade e a média direcional deste concurso nacional.

6 Considerações Finais

Embasados na estatística circular, viu-se que os métodos até aqui citados permitiram que os dados estudados, oriundos de coordenadas geográficas, fossem avaliados de forma consistente, o que a estatística linear não proporcionou. Além do mais, foi possível uma avaliação sobre a concentração dos dados bem como sobre sua distribuição de probabilidade. Verificou-se por meio do teste de Rayleigh uma tendência direcional das coordenadas das capitais dos estados de onde vêm os alunos que são aprovados no concurso, confirmando a rejeição da hipótese nula, que garantia a aleatoriedade dos aprovados. A ideia deste trabalho foi apresentar, principalmente, as medidas de tendência central, haja vista, a má aplicação da teoria da estatística linear em dados cuja abordagem envolve dados periódicos, angulares ou direcionais. Além do mais foram citadas, a fim de conhecimento, diferentes ferramentas de distribuição de probabilidade e testes de hipóteses, claro, dando ênfase e apresentando a parte teórica da distribuição de von Mises e do teste de Rayleigh,



que mostraram-se ser ferramentas essenciais na análise direcional dos ingressantes na EPCAR. Obteve-se, também, a apresentação de gráficos pouco conhecidos por estudantes do Ensino Médio, como exemplo o diagrama de rosas e o Boxplot circular, o que podem vir a estimular um olhar crítico dos alunos do Ensino Médio para dados angulares e até mesmo uma breve introdução conceitual sobre a importância em se ter técnicas coerentes diante de dados que possuam direções. No mais, uma reflexão diante dos aprovados da própria instituição, indicando se não seria viável melhorar a divulgação nas regiões com menos incidência de aprovados, ou até mesmo uma melhoria nas políticas de acesso, a fim de amenizar o regionalismo das vagas ocupadas.

Referências

AYRES, Manuel; Jr. AYRES, Manuel; AYRES, Daniel Lima; SANTOS, Alex de Assis Santos dos. **BioEstat**: aplicações estatísticas nas áreas das ciências bio-médicas. Belém, PA: Instituto Mamirauá, 2007.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, DF, 2018.

Disponível em:

<http://portal.mec.gov.br/docman/abril-2018-pdf/85121-bncc-ensino-medio/file>. Acesso em: 15 set. 2022.

FALCONI, Carlos Eduardo. Calculando distâncias e direções utilizando coordenadas geográficas.

Piloto Policial. 23 set. 2009. Disponível em: <https://www.pilotopolicial.com.br/calculando-distancias-e-direcoes-utilizando-coordenadas-geograficas/>. Acesso em: 3 nov. 2022.

GÓIS, Joaquim Eduardo Souza. **Contribuição dos modelos estocásticos para o estudo da climatologia urbana**. Orientadores: Henrique José de Figueiredo Garcia Pereira e Carlos Manuel Novais Madureira. 2002. 685 f. Tese (Doutorado) – Faculdade de Engenharia, Universidade do Porto, Porto, Portugal, 2002.

JAMMALAMADAKA, Sreenivasa Rao; SENGUPTA, Ashis. **Topics in circular statistics**. Series on Multivariate Analysis. v. 5. New Jersey: World Scientific, 2001.

MARDIA, Kantil. **Statistics of directional data**. London: Academic Press, 1972.

MENEZES, Eliane Barbosa. **Um estudo direcional dos estudantes que chegam à UFRB**: uma aplicação da estatística circular. Orientadores: Silvia Patrícia Barreto Santana e Celso Luiz Borges de Oliveira. 2018. 50 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências Exatas e Tecnológicas) – Universidade Federal do Recôncavo da Bahia, Cruz das Almas, Bahia, 2018.

PEWSEY, Arthur; NEUHÄUSER, Markus; RUXTON, Graeme D. **Circular statistics in R**. Oxford: Oxford University Press, 2013.



R Development Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2021. ISBN 3-900051-07-0. Disponível em: <https://www.r-project.org/>. Acesso em: 3 ago. 2022.

SANAR. **Linha do tempo do coronavírus no Brasil**. 19 mar. 2020. Disponível em: <https://www.sanarmed.com/linha-do-tempo-do-coronavirus-no-brasil>. Acesso em: 23 abr. 2023.

SILVA, Clodoaldo Teodosio Santana da; BRIGHENTI, Carla Regina Guimarães; RESENDE, Luiz Fernando Silva. Estatística circular aplicada aos dados de localização dos municípios de origem dos alunos do PROFMAT-UFSJ - Campus Santo Antonio. **Revista de Matemática de Ouro Preto**, Minas Gerais, v. 1, p. 1-10, nov. 2022. Disponível em: <https://periodicos.ufop.br/rmat/article/view/5216>. Acesso em: 26 dez. 2023.

