ISSN: ISSN: 0974-5823 Vol. 7 No. 2 February, 2022

International Journal of Mechanical Engineering

Energy-Like Statistic Properties Under Von-Mises Samples and Circular Distance

Ahmed Jebur Ali[1, a)] and Samira Faisal Abushilah[1, b)]

[1]*University of Kufa/ College of Education for Girls/ Department of Mathematics/ Najaf/ Iraq*

Abstract. In this article, distance-based statistic properties under Von-Mises samples and angular separation circular distance have been theoretically computed. This statistical test could be used as a two-sample test to detect the homogeneity between the distributions of two groups, where the data of these groups could be circular or linear data. Additionally, a simulation study is performed to validate the obtained properties using the R software version 3.6.2.

Keywords: Circular data, Von-Mises distribution, Energy statistic, Circular distance, Non-parametric two-sample test.

1. INTRODUCTION

In different scientific disciplines such as biology and medicine many variables are circular, where a cyclical scale is used to record this data, for examples compass direction or time of a day or dihedral angles on proteins. Such type of data is called circular data and it cannot be analysed using traditional statistical methods. Therefore, we seek circular statistical methods to deal with this type of data, for more information about circular data see for example [1-3].

For differences between two circular distributions, the most commonly-used test is Watson's U2 test [4], which is a non-parametric rank-based test. This test is implemented in most software applications such as that MATLAB [5] and R [6]. Wheeler and Watson [7] proposed a statistical test to test the difference between two distributions of circular data.

For multivariate circular data it is important to distinguish two multivariate samples but can be difficult. in 1984-1985, Gabor Székely introduced novel statistic depends on distance, energy statistic ($\mathcal{E}$-statistic), which has been introduced in a series of lectures [8]. This statistic is a function of distances between observations in metric space and it is very useful, and powerful than traditional statistics (non-energy type) such as F-statistic and correlation [9]. There are many applications for energy statistics in the recent years such as [10-13]. Let $\{x_1, x_2, \ldots, x_n\}$and $\{y_1, y_2, \ldots, y_m\}$ be independent samples from two independent random variables X and Y then the energy statistic is defined by

$$\mathcal{E}(X,Y) = \frac{nm}{n+m}\left(\frac{2}{nm}A_1 - \frac{1}{n^2}A_2 - \frac{1}{m^2}A_3\right) \qquad (1)$$

where

$$\mathrm{A}_1 = \sum_{i=1}^{n}\sum_{j=1}^{m} D(x_i, y_j), \qquad \mathrm{A}_2 = \sum_{i=1}^{n}\sum_{j=1}^{n} D(x_i, x_j), \qquad \mathrm{A}_3 = \sum_{i=1}^{m}\sum_{j=1}^{m} D(y_i, y_j),$$

and $D(x_i, y_j)$ represents the distance measure between the observations x_i and y_j.

However, the energy statistic could be used with a randomization test (which is very time consuming) to detect the homogeneity between the distributions of two groups because the distribution of the energy statistic is unknown and it is not easy to compute. Therefore, in 2021, Ali and Abushilah [14] proposed a novel test statistic to detect the homogeneity between two groups of circular data $S_1 = \{\phi_1, \phi_2, \ldots, \phi_n\}$ and $\mathrm{S}_2 = \{\psi_1, \psi_2, \ldots, \psi_m\}$, to test the null hypothesis

$$H_0: G_\Phi = G_\Psi \qquad \text{against} \qquad H_a: G_\Phi \neq G_\Psi.$$

Energy-like statistic is defined by the following form:

$$\mathcal{T}_d(S_1, S_2) = \left(\frac{\mathrm{nm}}{\mathrm{n+m}}\right)^{\frac{1}{2}} \sum_{i=1}^{n}\sum_{j=1}^{m} D(\phi_i, \psi_j), \qquad (2)$$

Where $D(\phi_i, \psi_j)$ is a distance function between the observations ϕ_i and ψ_j.

In this paper, the theoretical properties of the energy-like statistic are theoretically calculated under circular distance and von-Mises samples. This statistic could be used as a nonparametric two-sample test with a permutation test [15] to detect the homogeneity between the distributions of two groups of circular data.

2. BASIC CONCEPTS

This section introduces basic concepts.

Definition (2.1) [3]:

Circular statistics is a field of statistics that deals with data that may be represented as points on the unit circle's circumference. The phrase "circular data" is used to differentiate them from the more common "linear data" that we are more familiar with. More precisely, we refer to the unit circle as the basis for circular data.

Definition (2.2) [1]:

The von-Mises distribution is a symmetric unimodal distribution and is the most common model for unimodal samples of circular data. The following is the definition of the probability density function (PDF):

$$f(\vartheta|\mu,\kappa) = \frac{e^{\kappa cos(\vartheta-\mu)}}{2\pi I_0(\kappa)}$$

Where

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

is the zero-order modified Bessel function. The concentration parameter is $.\kappa$, and the mean direction is μ.

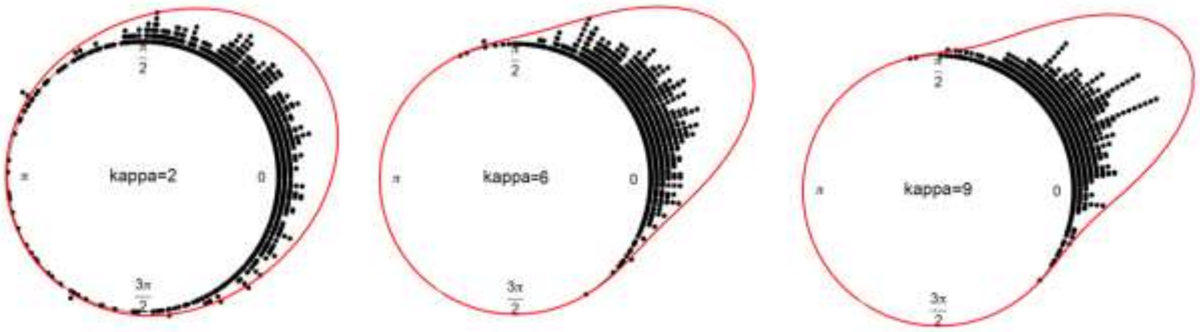

Figure 1. The von-Mises distribution which has been used in this analysis over a range of concentration parameters (values given in the plots, with sample size 1000 for each).

3. PROPERTIES OF THE STATISTIC $\mathcal{T}_d$ UNDER ONE-DIMENSIONAL CASE

Let Φ and Ψ be two independent random variables such that $\Phi \sim vM(\mu_1,\kappa_1)$ and $\Psi \sim vM(\mu_2,\kappa_2)$ and let $S_1 = \{\phi_1,\phi_2,\dots,\phi_n\}$and $S_2 = \{\psi_1,\ \psi_2,\dots,\psi_m\}$ be random samples from these random variables, then the expected value for the statistic $\mathcal{T}_d$ is calculated as follows:

$$E(\mathcal{T}_d(S_1,S_2)) = E\left(\left(\frac{nm}{n+m}\right)^{\frac{1}{2}} \sum_{i=1}^{n}\sum_{j=1}^{m} D(\phi_i,\psi_j)\right)$$

Since $D(\phi_i,\psi_j) = 1-\cos(\phi-\psi_j), i = 1,2,\dots,n; j = 1,2,\dots,m$, then we get

$$E(\mathcal{T}_d(S_1,S_2)) = \Big(\frac{nm}{n+m}\Big)^{\frac{1}{2}}\left(\sum_{i=1}^{n}\sum_{j=1}^{m} E\big[1-cos(\phi_i-\psi_j)\big]\right)$$

$$= \Big(\frac{nm}{n+m}\Big)^{\frac{1}{2}}\left(nm - \sum_{i=1}^{n}\sum_{j=1}^{m} E\big[cos(\phi_i-\psi_j)\big]\right) \qquad (3)$$

The quantity $E\big[\cos(\phi_i - \psi_j)\big]$ in Equation (1) is computed under the assumption of the independent and identically distributed (i.i.d.) Von Mises samples as follows:

$$E[cos(\phi-\psi)] = \int_0^{2\pi}\int_0^{2\pi} cos(\phi-\psi)\, f(\phi,\psi)\, d\phi\, d\psi.$$

$$= \int_0^{2\pi}\int_0^{2\pi} \big(cos(\phi)cos(\psi) + sin(\phi)sin(\psi)\big)\, f(\phi)f(\psi)\, d\phi\, d\psi.$$

$$= \frac{1}{I_0^2(\kappa)}\int_0^{2\pi}\int_0^{2\pi} \big(cos(\phi)cos(\psi) + sin(\phi)sin(\psi)\big)\, \frac{e^{\kappa cos(\phi)}e^{\kappa cos(\psi)}}{(2\pi)(2\pi)}\, d\phi\, d\psi.$$

$$= \frac{I_1(\kappa)}{I_0^2(\kappa)}\left[\int_0^{2\pi} cos(\psi)\frac{e^{\kappa\, cos(\Psi)}}{2\pi} d\psi\right]$$

As a result, we get the following

$$\mathrm{E}[co\, s(\phi-\psi)] = \frac{\mathrm{I}_1^2(\kappa)}{\mathrm{I}_0^2(\kappa)}. \qquad (4)$$

So, the expected value of the statistic $\mathcal{T}_d$ is equal to

$$\mathrm{E}(\mathcal{T}_d(S_1,S_2)) = \left(\frac{(\mathrm{mn})^3}{\mathrm{m}+\mathrm{n}}\right)^{\frac{1}{2}}\left(\frac{\mathrm{I}_0^2 - \mathrm{I}_1^2(\kappa)}{\mathrm{I}_0^2(\kappa)}\right) \qquad (5)$$

The variance of the statistic $\mathcal{T}_d$ is calculated as follows:

$$\mathrm{var}(\mathcal{T}_d(S_1,S_2)) = \mathrm{var}\left(\Big(\frac{\mathrm{nm}}{\mathrm{n}+\mathrm{m}}\Big)^{\frac{1}{2}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{D}(\phi_\mathrm{i},\psi_\mathrm{j})\right)$$

$$= \Big(\frac{\mathrm{nm}}{\mathrm{n}+\mathrm{m}}\Big)\,\mathrm{var}\left(\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{D}(\phi_\mathrm{i},\psi_\mathrm{j})\right) \qquad (6)$$

The quantity in the right hand side of Equation (6) is computed as follows:

$$var(D(\phi,\psi)) = var(1-cos(\phi-\psi))$$

$$= var(cos(\phi-\psi))$$

$$= \mathrm{E}(cos^2(\phi-\psi)) - \big(\mathrm{E}(\cos(\phi-\psi))\big)^2$$

$$= 1 - E(sin^2(\phi - \psi)) - \frac{I_1^4(\kappa)}{I_0^4(\kappa)} \quad \text{[using Equation (2)]}$$

$$= 1 - \frac{1}{2}\big(1 - E(cos(2\phi - 2\psi))\big) - \frac{I_1^4(\kappa)}{I_0^4(\kappa)} \qquad (7)$$

and,

$$E(cos(2\phi - 2\psi)) = \int_0^{2\pi}\int_0^{2\pi} cos(2\phi - 2\psi)\, f(\phi,\psi)\, d\phi\, d\psi.$$

$$= \int_0^{2\pi}\int_0^{2\pi} \big(cos(2\phi)cos(2\psi) + sin(2\phi)sin(2\psi)\big)\, f(\phi)f(\psi)\, d\phi\, d\psi.$$

$$= \frac{I_2(\kappa)}{I_0^2(\kappa)}\left[\int_0^{2\pi} cos(2\psi)\frac{e^{\kappa\, cos(\psi)}}{2\pi} d\psi + 0\right]$$

As a result, we get the following:

$$E[cos(2\phi, 2\psi)] = \frac{I_2^2(\kappa)}{I_0^2(\kappa)} \qquad (8)$$

Substituting Equation (8) in Equation (7), we get the result

$$var\big(D(\phi,\psi)\big) = \frac{I_0^4(\kappa) + 2I_0^2(\kappa)I_2^2(\kappa) - I_1^4(\kappa)}{2I_0^4(\kappa)} \qquad (9)$$

The formula $cov\big(\mathrm{D}(\Phi_1,\psi_1)\big),\big(\mathrm{D}(\Phi_1,\psi_2)\big)$ is calculated as follows:

$$cov\big(D(\phi_1,\psi_1)\big),\big(D(\phi_1,\psi_2)\big) = E\big(D(\phi_1,\psi_1)\big).\big(D(\phi_1,\psi_2)\big) - E\big(D(\phi_1,\psi_1)\big)E\big(D(\phi_1,\psi_2)\big)$$

$$= E\big(D(\phi_1,\psi_1)\big).\big(D(\phi_1,\psi_2)\big) - \left(1 - \frac{I_1^2(\kappa)}{I_0^2(\kappa)}\right)^2 \qquad (10)$$

The first quantity in Equation (10) is computed as follows:

$$E\big(D(\phi_1,\psi_1)\big).\big(D(\phi_1,\psi_2)\big) = E\big((1 - cos(\phi_1 - \psi_1))(1 - cos(\phi_1 - \psi\Psi_2))\big)$$

$$= E(1 - \cos(\phi_1 - \psi_2) - \cos(\phi_1 - \psi_1) + \cos(\phi_1 - \psi_1)\cos(\phi_1 - \psi_2))$$

$$= 1 - E(\cos(\phi_1 - \psi_2)) - E(\cos(\phi_1 - \psi_1)) + E(\cos(\phi_1 - \psi_1)\cos(\phi_1 - \psi_2))$$

$$= 1 - \frac{2\mathrm{I}_1^2(\kappa)}{\mathrm{I}_0^2(\kappa)} + E(\cos(\phi_1 - \psi_1)\cos(\phi_1 - \psi_2)) \qquad (11)$$

In Equation (8) the second quantity $E(\cos(\Phi_1 - \Psi_1)\cos(\Phi_1 - \Psi_2))$ is equal to

$$= E((\cos(\phi_1)\cos(\psi_1) + \sin(\phi_1)\sin(\psi_1))(\cos(\phi_1)\cos(\psi_2) + \sin(\phi_1)\sin(\psi_2))$$

$$= E(\cos^2(\phi_1)\cos(\psi_1)\cos(\psi_2)) + E(\cos(\phi_1)\cos(\psi_1)\sin(\phi_1)\sin(\psi_2)$$
$$+ E(\cos(\phi_1)\cos(\psi_2)\sin(\phi_1)\sin(\psi_1)) + E(\sin^2(\phi_1)\sin(\psi_1)\sin(\psi_2)) \qquad (12)$$

where

$$E(cos^2(\phi_1)\, cos(\psi_1)\, cos(\psi_2)) = \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} cos^2(\phi_1)\, cos(\psi_1)\, cos(\psi_2)\, f(\phi_1,\psi_1,\psi_2)\, d\phi_{\,a}\, d\psi_1\ d\psi_2$$

$$=\frac{1}{I_0^3(\kappa)}\int_0^{2\pi}\int_0^{2\pi}\left(\int_0^{2\pi} cos(\phi_1)^2\frac{e^{\kappa\, cos(\phi_1)}}{2\pi}d\phi_1\right)cos(\psi_1)cos(\psi_2)\frac{e^{\kappa cos(\psi_1)}e^{\kappa cos(\psi_2)}}{(2\pi)(2\pi)}d\psi_1 d\psi_2$$

$$=\frac{1}{I_0^3(\kappa)}\int_0^{2\pi}\int_0^{2\pi}\left(\frac{1}{2}\int_0^{2\pi}\left(\frac{e^{\kappa\, cos(\phi_1)}}{2\pi}+\frac{cos(2\phi_1)\, e^{\kappa\, cos(\phi_1)}}{2\pi}\right)d\phi_1\right)cos(\psi_1)cos(\psi_2)\frac{e^{\kappa cos(\psi_1)}e^{\kappa cos(\psi_2)}}{(2\pi)(2\pi)}d\psi_1 d\psi_2$$

$$=\frac{\left(I_0(\kappa)+I_2(\kappa)\right)}{2I_0^3(\kappa)}\int_0^{2\pi}\int_0^{2\pi} cos(\psi_1)cos(\psi_2)\frac{e^{\kappa cos(\psi_1)}e^{\kappa cos(\Psi_2)}}{(2\pi)(2\pi)}d\psi_1 d\psi_2$$

As a result, we get the following

$$E(\cos^2(\phi_1)\cos(\psi_1)\cos(\psi_2))=\frac{\left(I_0(\kappa)+I_2(\kappa)\right)I_1^2(\kappa)}{2I_0^3(\kappa)} \qquad (13)$$

We noticed that the following quantities in Equations (12) is equal to zero, so substitution Equation (13) in Equation (12) we obtain the result

$$E(cos(\phi_1-\psi_1)\,cos(\phi_1-\psi_2))=\frac{\left(I_0(\kappa)+I_2(\kappa)\right)I_1^2}{2I_0^3(\kappa)} \qquad (14)$$

Substitution Equation (14) in Equation (11) we get

$$E\left(D(\phi_1,\psi_1)\right).\left(D(\phi_1,\psi_2)\right)=1-\frac{2I_1^2(\kappa)}{I_0^2(\kappa)}+\frac{\left(I_0(\kappa)+I_2(\kappa)\right)I_1^2(\kappa)}{2I_0^3(\kappa)}$$

and

$$cov\left(D(\phi_1,\psi_1)\right),\left(D(\phi_1,\psi_2)\right)=\frac{I_1^2(\kappa)}{2I_0^2(\kappa)}+\frac{I_1^2(\kappa)I_2(\kappa)}{2I_0^2(\kappa)}-\frac{I_1^4(\kappa)}{I_0^4(\kappa)} \qquad (15)$$

$$var\left(\sum_{i=1}^{n}\sum_{j=1}^{m}D\left(\phi_i,\psi_j\right)\right)=nm\left[\left(\frac{I_0^4(\kappa)+2I_0^2(\kappa)I_2^2-2I_1^4(\kappa)}{2I_0^4(\kappa)}\right)+(n+m-2)\left(\frac{I_0^2I_1^2(\kappa)+I_0^2(\kappa)I_1^2(\kappa)I_2(\kappa)-2I_1^4(\kappa)}{2I_0^4(\kappa)}\right)\right] \qquad (16)$$

Applying Equation (16) in Equation (6) we get the result

$$var(\mathcal{T}_d(S_1,S_2))=\frac{(nm)^2}{n+m}\left[\frac{I_0^4(\kappa)+2I_0^2(\kappa)I_2^2-2I_1^4(\kappa)}{2I_0^4(\kappa)}+\frac{(n+m-2)(I_0^2I_1^2(\kappa)+I_0^2(\kappa)I_1^2(\kappa)I_2(\kappa)-2I_1^4(\kappa))}{2I_0^4(\kappa)}\right] \qquad (17)$$

4. PROPERTIES OF THE STATISTIC $\mathcal{T}_d$ UNDER P-DIMENSIONAL CASE ($p>1$)

Let X and Y in $R^p, p>1$ be Let $\{(\phi_{11},\dots,\phi_{p1}),\dots,(\phi_{1n},\dots,\phi_{pn})\}$, and $\{(\psi_{11},\dots,\psi_{p1}),\dots,(\psi_{1m},\dots,\psi_{pm})\}$ be independent circular random variables from $\mathrm{BvM}(\mu,\Lambda)$, where $\mu=\binom{0}{0}$ and $\Lambda=\begin{pmatrix}\kappa & 0\\ 0 & \kappa\end{pmatrix}$, and let the distance measure be torus distance, then the expected value and the variance of the $\mathcal{T}_d$ statistic are defined by:

$$E(\mathcal{T}_d(S_1,S_2))=E\left(\left(\frac{nm}{n+m}\right)^{\frac{1}{2}}\sum_{i=1}^{n}\sum_{j=1}^{m}D\left(\phi_{1i},\psi_{1j}\right)+\cdots+D\left(\phi_{ni},\psi_{mj}\right)\right)$$

$$=\left(\frac{nm}{n+m}\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}\sum_{j=1}^{m}E\left[D\left(\phi_{1i},\psi_{1j}\right)+\cdots+D\left(\phi_{ni},\psi_{mj}\right)\right]\right)$$

Since $D(\phi,\psi)=1-\cos(\phi-\psi)$, then we get

$$E(\mathcal{T}_d(S_1,S_2)) = \left(\frac{nm}{n+m}\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} \sum_{j=1}^{m} E\left[\left(1-cos(\phi_{1i}-\psi_{1j})\right)+\cdots+\left(1-cos(\phi_{ni}-\psi_{mj})\right)\right] \right) \quad (18)$$

Since $E[1-\cos(\phi-\psi)] = 1-\frac{I_1^2(\kappa)}{I_0^2(\kappa)}$, then we get

$$E(\mathcal{T}_d(S_1,S_2)) = p\left(\frac{(nm)^3}{n+m}\right)^{\frac{1}{2}}\left(1-\frac{I_1^2(\kappa)}{I_0^2(\kappa)}\right). \quad (19)$$

The variance of the statistic $\mathcal{T}_d(S_1,S_2)$ is calculated as follows:

$$var(\mathcal{T}_d(S_1,S_2)) = var\left(\left(\frac{nm}{n+m}\right)^{\frac{1}{2}} \sum_{i=1}^{n} \sum_{j=1}^{m} D\left(\phi_{1i},\psi_{1j}\right)+\cdots+D\left(\phi_{ni},\psi_{mj}\right) \right)$$

$$= \left(\frac{nm}{n+m}\right) var\left(\sum_{i=1}^{n} \sum_{j=1}^{m} D\left(\phi_{1i},\psi_{1j}\right)+\cdots+D\left(\phi_{ni},\psi_{mj}\right) \right)$$

Since $D(\phi,\psi) = 1-\cos(\phi-\psi)$, then we get

$$var(\mathcal{T}_d) = \left(\frac{nm}{n+m}\right) var\left(\sum_{i=1}^{n} \sum_{i=1}^{m} \left(1-cos(\phi_{1i}-\psi_{1j})\right)+\cdots+\left(1-cos(\phi_{ni}-\psi_{mj})\right) \right). \quad (20)$$

Applying Equation (16) in Equation (20) we get the result

$$var(\mathcal{T}_d) = \left(\frac{p(nm)^2}{n+m}\right)\left[\frac{I_0^4(\kappa)+2I_0^2(\kappa)I_2^2-2I_1^4(\kappa)}{2I_0^4(\kappa)} + \frac{(n+m-2)(I_0^2I_1^2(\kappa)+I_0^2(\kappa)I_1^2(\kappa)I_2(\kappa)-2I_1^4(\kappa))}{2I_0^4(\kappa)}\right] \quad (21)$$

4. SIMULATION STUDY

A simulation analysis is conducted in this section to examine the theoretical and empirical properties of the energy-like statistic that we have theoretically calculated. The simulation is described as follows:

1. When $p = 1$:

- Generate two groups $S_1 = \{\phi_1,\phi_2,\dots,\phi_n\}$ and $S_2 = \{\psi_1,\ \psi_2,\dots,\psi_m\}$ from $vM(\mu_1,\kappa_1)$ and $vM(\mu_2,\kappa_2)$ under the assumptions $\mu_1=\mu_2=0$ and $\kappa_1=\kappa_2=\kappa$.
- Calculate the circular distance $D(\phi_i,\psi_j)$, for each $\phi_i \in S_1$ and $\psi_j \in S_2$, which is given by

$$D(\phi_i,\psi_j) = 1-\cos(\phi_i-\psi_j), i=1,2,\dots,n; j=1,2,\dots,m$$

- Compute the expected value and variance of the statistic $\mathcal{T}_d$ which are given by

$$E(\mathcal{T}_d(S_1,S_2)) = \left(\frac{(mn)^3}{m+n}\right)^{\frac{1}{2}}\left(1-\frac{I_1^2(\kappa)}{I_0^2(\kappa)}\right).$$

and

$$var(\mathcal{T}_d(S_1,S_2)) = \frac{(nm)^2}{n+m}\left[\frac{I_0^4(\kappa)+2I_0^2(\kappa)I_2^2-2I_1^4(\kappa)}{2I_0^4(\kappa)} + \frac{(n+m-2)(I_0^2I_1^2(\kappa)+I_0^2(\kappa)I_1^2(\kappa)I_2(\kappa)-2I_1^4(\kappa))}{2I_0^4(\kappa)}\right]$$

2. When $p = 2$:

- Generate two independent random samples from $BvM(\mu_1,\Lambda)$ and $BvM(\mu_2,\Lambda)$, where $\mu_1=\mu_2=\binom{0}{0}$, $\Lambda = \begin{pmatrix}\kappa & 0\\ 0 & \kappa\end{pmatrix}$.
- Calculate the torus distance $TD(\Phi_i,\psi_j)$, for each $\Phi_i = (\phi_{1i},\ \phi_{2i}) \in S_1$, $\psi_i = (\psi_{1i},\ \psi_{2i}) \in S_2$, $i=1,2,\dots,n; j=1,2,\dots,m$, which is given by

$$\mathrm{TD}(\Phi_i, \psi_j) = \mathrm{D}(\phi_{ik}, \phi_{jk}) + \mathrm{D}(\psi_{ik}, \psi_{jk})$$

- Calculate the expected value and variance of the statistic $\mathcal{T}_d$ which are given by

$$\mathrm{E}(\mathcal{T}_d(S_1, S_2)) = p\left(\frac{(\mathrm{mn})^3}{\mathrm{m}+\mathrm{n}}\right)^{\frac{1}{2}}\left(1 - \frac{\mathrm{I}_1^2(\kappa)}{\mathrm{I}_0^2(\kappa)}\right)$$

and

$$var(\mathcal{T}_d) = \left(\frac{p(nm)^2}{n+m}\right)\left[\frac{I_0^4(\kappa) + 2I_0^2(\kappa)I_2^2 - 2I_1^4(\kappa)}{2I_0^4(\kappa)} + \frac{(n+m-2)(I_0^2 I_1^2(\kappa) + I_0^2(\kappa)I_1^2(\kappa)I_2(\kappa) - 2I_1^4(\kappa))}{2I_0^4(\kappa)}\right]$$

Tables 1-4 summarize the simulation's results. For both one-dimensional and two-dimensional cases, we can observe that the empirical expected values and variances are close to the theoretical expected values and variances.

Table 1: Comparison between the theoretical and empirical expected value and variance of the statistic $\mathcal{T}_d(S_1, S_2)$, where the samples S_1 and S_2 have been generated from $\mathrm{vM}(\mu, \kappa)$ distributions with different sample sizes.

Sample size	Empirical		Theoretical	
	Expected value	Variance	Expected value	Variance
10	178.9507	868.3998	179.05	868.4966
20	1013.152	12538.32	1012.86	12596.65
30	2789.683	61684.46	2791.113	61577.95
40	5725.523	190828.8	5729.6	191151.9
50	10025.38	469692	10009.2	461604.1
60	15782.25	953590.4	15788.92	950166.1
70	23218.64	1724469	23212.37	1751015
80	32399.33	2935259	32411.51	2975276
90	43530.18	4757111	43509.15	4751016
100	56651.91	7210335	56620.59	7223253

Table 2: Comparison between the theoretical and empirical expected value and variance of the statistic $\mathcal{T}_d(S_1, S_2)$, where the samples S_1 and S_2 were generated from $\mathrm{vM}(\mu,\kappa)$ distributions with sizes $n = 60$ and $m = 70$, mean $\mu = 0$ and different variances

κ	Empirical		Theoretical	
	Expected value	Variance	Expected value	Variance
0.5	22476.44	538656.7	22468.88	533386.1
1	19155.55	1288710	19115.76	1289609
1.5	15398.33	1595181	15388.97	1579393
2	12233.36	1458050	12249.35	1430645
2.5	9963.549	1134299	9901.934	1137291
3	8183.725	866104.6	8210.404	866224
3.5	6999.711	660479.2	6983.836	659999.7
4	6053.882	518651.9	6071.523	512256.3
4.5	5353.591	409025.5	5371.767	407059.7
5	4791.876	331578.4	4819.108	330897.9

Table 3: Comparison between the theoretical and empirical expected value and variance of the statistic $\mathcal{T}_d(S_1, S_2)$ with different sample sizes which have been generated from $BvM(\mu, \kappa)$ distributions..

Sample size	Empirical		Theoretical	
	Expected value	Variance	Expected value	Variance
10	357.6182	1734.989	358.1	1736.993
20	2017.963	25130.77	2025.72	25193.29
30	5572.333	123328.6	5582.227	123155.9
40	11458.8	382905.4	11459.2	382303.9
50	20017.56	921921.9	20018.4	923208.2
60	31634.75	1918410	31577.84	1900332
70	46442.57	3519432	46424.74	3502031
80	64694.33	5708044	64823.03	5950551
90	87006.13	9302072	87018.31	9502033
100	113149.6	14529473	113241.2	14446507

Table 4: Comparison between the theoretical and empirical expected value and variance of the statistic $\mathcal{T}_d(S_1, S_2)$, where the samples S_1 and S_2 of sizes $n = 60,\ m = 70$ and different variances have been generated from $BvM(\mu, \kappa)$.

κ	Empirical		Theoretical	
	Expected value	Variance	Expected value	Variance
0.5	44977.84	1062292	44937.76	1066772
1	38201.79	2553813	38231.52	2579217
1.5	30704.28	3193712	30777.94	3158786
2	24512.14	2864423	24498.71	2861290
2.5	19781.92	2244229	19803.87	2274582
3	16440.43	1769330	16420.81	1732448
3.5	13961.4	1322599	13967.67	1319999
4	12130.82	1008783	12143.05	1024513
4.5	10754.05	10743.53	828970.7	814119.5
5	9628.369	657730	9638.216	661795.8

5. CONCLUSION

The theoretical properties of the energy-like statistic based on torus distance and for multidimensional cases are presented in this article. This statistical test could be used with a permutation test as a two-sample test to determine the homogeneity between the distributions of two groups of circular data. Moreover, a simulation study is performed to validate the properties that we have calculated using the R program version 3.6.2.

.

6. REFERENCES

[1] S. Jammalamadaka, and A. Sengupta, "Topics in Circular Statistics, " 1st ed. Singapore: World Scientific Publishing Company, (2001).

[2] K.V. Mardia and P.E. Jupp, "Directional Statistics, " (Vol. 494), John Wiley & Sons, 2009.

[3] A.Pewsey, M.Neuhäuser, and G. D. Ruxton, "Circular statistics in R, " Oxford University Press, (2013).

[4] G. Watson, "Goodness-of-fit tests on a circle, " II. Biometrika, **49**, pp. 57-63, (1962).

[5] P. Mégevand, pierremegevand/watsons_u2 (https://github.com/pierremegevand/watsons_u2), GitHub. Retrieved February 20, 2022 (2022).

[6] C. Agostinelli and M. C. Agostinelli, Package "CircStats, " See https://cran. r-project. org/web/packages/CircStats/CircStats. pdf, (2018).

[7] S. Wheeler and G. S. Watson, (1964). A distribution-free two-sample test on a circle. Biometrika, pp. 256-257.

[8] M. L. Rizzo and G. J. Székely, "Energy distance, " wiley interdisciplinary reviews: Computational statistics, 8(1), pp. 27-38, (2016).

[9] G. J. Székely and M. L. Rizzo, "Energy statistics: A class of statistics based on distances," Journal of statistical planning and inference, 143(8), pp. 1249-1272, (2013).

[10] Benoist, T., Jakšić, V. and Pillet, C.A. , " Energy statistics in open harmonic networks, " Journal of Statistical Physics, 168(5), pp.1016-1030, (2017).

[11] Khanikar, K., Sinha, R. and Bhattacharjee, R. , " Cooperative spectrum sensing using quantized energy statistics in the absence of dedicated reporting channel, " IEEE Transactions on Vehicular Technology, 67(5), pp.4149-4160, (2018).

[12] Nguyen, H.D., Arbel, J., Lü, H. and Forbes, F. , " Approximate Bayesian computation via the energy statistic, " IEEE Access, 8, pp.131683-131698, (2020).

[13] Green, E. and DeLisser, C.D. , " Advancing a culture of effective energy statistics and information management: A study of the CARICOM economic space, " World Journal of Science, Technology and Sustainable Development, (2018).

[14] Ali, A., Abushilah, S. Distribution-free two-sample homogeneity test for circular data based on geodesic distance. International Journal of Nonlinear Analysis and Applications, 13(1), pp. 2703-2711 (2022).

[15] P. Good, "Permutation tests: a practical guide to resampling methods for testing hypotheses, " Springer Science & Business Media, (2013).