In Response to 'Advice on Comparing Two Independent Samples of Circular Data in Biology'

Patrick Mellady

2024-05-01

Motivation

Many examples of directional data can be found in the natural sciences. The recent paper by Landler, Ruxton, and Malkemper [1] focuses specifically on circular data in a biological context. In this response, we will provide a critique of their analysis of the currently available methods for comparing circular data.

Directional Data

To work with directional data, we first define our terms. The (d-1)-sphere is a d-1 dimensional embedded sub-manifold of $\mathbb{R}^d$ defined by $S^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. Directional data is a collection of *random directions*, which are random vectors in S^{d-1}. Specifically, when working with circular data, we are denoting the special case where our random directions come from S^1 (the unit circle in $\mathbb{R}^2$). For the analyses in this paper, we will assume a Langevin-von Mises-Fisher distribution on S^{d-1} which has density

$$f(\mathbf{u}; \boldsymbol{\mu}, \kappa) = \frac{(\kappa/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\kappa)} e^{\kappa \boldsymbol{\mu}^T \mathbf{u}}, \quad \mathbf{u} \in S^{d-1} \tag{1}$$

where

- $\boldsymbol{\mu} \in S^{d-1}$ is the mean direction (defined below)
- $\kappa > 0$
- Γ is the gamma function
- I_ν is the modified Bessel function of the first kind and order ν

Fundamental Quantities

In order to work with directional data, we define some fundamental quantities. Firstly, given a random sample of directions $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$, we define the *resultant vector* to be $\mathbf{v} = \sum_{i=1}^n \mathbf{u}_i$, which naturally implies the sample *mean vector* is $\overline{\mathbf{u}} = \frac{1}{n}\mathbf{v}$.

Although the above are both important quantities, we see a problem. By calculating the resultant or mean vectors, we may violate the unit length condition that keeps the data on the sphere. To return to the realm of directions on the sphere, we need to standardize the above

quantities by calculating and utilizing the *resultant length* $R = \|\mathbf{v}\|$. Equipped with the resultant length, we can define the sample *mean direction* as $\frac{\overline{\mathbf{u}}}{R/n} = \frac{\mathbf{v}}{R}$, when $R \neq 0$.

These constructions give us insight into where we will start our analysis: we will examine the sample mean resultant length of two vectors in S^{d-1}. See that:

$$\begin{aligned}
\|\overline{\mathbf{u}}\|^2 = \left\|\frac{1}{2}(\mathbf{u}_1 + \mathbf{u}_2)\right\|^2 &= \frac{1}{4}(\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + 2\mathbf{u}_1 \cdot \mathbf{u}_2) \\
&= \frac{1}{4}(\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + 2\|\mathbf{u}_1\|\|\mathbf{u}_2\|\cos(\theta)) \\
&= \frac{1}{4}(1 + 1 + 2\cos(\theta)) \\
&= \frac{1}{2}(1 + \cos(\theta)) \\
&\leq 1
\end{aligned}$$

Here it is made clear that, as long as the sample does not share a common direction, the mean resultant length will be less than 1. It also becomes clear that as the random directions become more different, the sample mean resultant length becomes closer to zero. This gives us a starting point of using the sample mean resultant length as a measure of concentration of the data: the closer the sample mean resultant length is to 1, the more concentrated the random directions are. Additionally, the calculations above show that if the sample mean resultant length is zero, $\overline{\mathbf{u}} = 0$.

Although the above holds for a single sample of random directions distributed on the sphere, we desire to compare multiple samples. In which case, we generalize the above in an obvious way: Suppose we have i samples of size j where $i = 1, 2, 3, \cdots, k$ and $j = 1, 2, 3, \cdots, n_i$. Then

- Resultant vector: $\sum_{j=1}^{n_i} \mathbf{u}_{ij}$
- Sample mean vector: $\overline{\mathbf{u}}_i = n_i^{-1} \sum_{j=1}^{n_i} \mathbf{u}_{ij}$
- Resultant length: $R_i = \left\|\sum_{j=1}^{n_i} \mathbf{u}_{ij}\right\|$
- Sample mean resultant length: $\overline{R}_i = n_i^{-1} R_i$
- Sample mean direction vector: $(\overline{R}_i)^{-1}\overline{\mathbf{u}}_i$

Additionally, we can generalize some of the above quantities to hold at the population level, not just for a sample. Supposing that $\mathbf{u} \sim P$ where P is some distribution on the sphere, we have:

- Mean direction: $\boldsymbol{\mu} = \frac{E(\mathbf{u})}{\|E(\mathbf{u})\|}$
- Mean resultant length: $\rho = \|E(\mathbf{u})\|$

Here is where we see a connection between the mean direction and mean resultant length from these definitions and the definition of $\boldsymbol{\mu}$ in (1). Additionally, $\rho = A_d(\kappa) = \frac{I_{d/2}(\kappa)}{I_{d/2-1}(\kappa)}$ is the mean resultant length of the Langevin-von Mises-Fisher distribution.

Testing

Possible goals for comparing two independent samples of directional data include: $^{i)}$ testing to see if the concentrations are equal and $^{ii)}$ testing to see if the mean directions are equal, and $^{iii)}$ testing for any difference in the distributions. It is important to know what kind of test the

researcher is interested in performing so that, in choosing one of the currently available methods, the correct test will be selected. Testing for equality of concentration by using a test for the equality of mean direction does not work. Employing an omnibus test for $^{(iii)}$ may lead to a substantial loss in power when compared with tests focused on $^{(i)}$ or $^{(ii)}$. This work will focus on testing for equality of mean directions between independent samples of directional data.

With the above framework in hand, we begin to develop a test for the difference of mean directions between groups of directional data. We suppose that $\mathbf{u}_{11}, \cdots, \mathbf{u}_{1n_1}, \mathbf{u}_{21}, \cdots, \mathbf{u}_{2n_2}, \cdots, \mathbf{u}_{k1}, \cdots, \mathbf{u}_{kn_k}$ are a collection of k random samples of size n_i for $i = 1, 2, 3, \cdots, k$, possibly from different distributions described in (1) but all with the same concentration. We wish to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k \tag{2}$$

$$H_a : \boldsymbol{\mu}_i \neq \boldsymbol{\mu}_j \text{ for at least one (i,j) pair}$$

The first test we may be interested in trying is a likelihood ratio test, in which case we must find the likelihood of the sample. Letting $n = \sum_{i=1}^{k} n_i$ and $U = (\mathbf{u}_{11}, \mathbf{u}_{12}, \cdots, \mathbf{u}_{k,n_k})$ be a matrix whose columns are the n total random directions, we can describe the likelihood as:

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_k, \kappa : U) = \Big(\frac{(\kappa/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\kappa)}\Big)^n e^{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \kappa \boldsymbol{\mu}_i^T \mathbf{u}_{ij}}$$

Under the null hypothesis, we obtain the following maximum likelihood estimates:

$$\hat{\boldsymbol{\mu}}_0 = (\overline{R})^{-1} \overline{\mathbf{u}}, \quad \hat{\kappa_0} = A_d^{-1}(\overline{R})$$

and under the alternative hypothesis, we obtain the following maximum likelihood estimates:

$$\hat{\boldsymbol{\mu}}_{ai} = (\overline{R}_i)^{-1} \overline{\mathbf{u}}_i, \quad \hat{\kappa_a} = A_d^{-1}(\tilde{R}), \text{ where } \tilde{R} = n^{-1} \sum_{i=1}^{k} R_i$$

Now, using the above, we can describe the ratio of the likelihoods as

$$\Lambda = \frac{L(\hat{\boldsymbol{\mu}}_0, \hat{\kappa}_0; U)}{L(\hat{\boldsymbol{\mu}}_{a1}, \cdots, \hat{\boldsymbol{\mu}}_{ak}, \hat{\kappa}_a; U)} = \frac{\Big(\frac{(\hat{\kappa}_0/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\hat{\kappa}_0)}\Big)^n e^{\hat{\kappa}_0 A_d(\hat{\kappa_0})}}{\Big(\frac{(\hat{\kappa}_a/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\hat{\kappa}_a)}\Big)^n e^{\hat{\kappa}_a A_d(\hat{\kappa_a})}}$$

We may directly use the likelihood ratio test at this point. However, asymptotic considerations of directional data will prove that the soon-to-be derived P statistic, based on the likelihood ratio statistic, consistently outperforms all other tests for equality of mean directions in both asymptotic settings.

Distributional Results

The likelihood ratio above has two asymptotic considerations: the large sample and the high concentration cases. To obtain the large sample case, we fix κ to be some positive number and

assume $\min(n_1, n_2, \cdots, n_k) \to \infty$. Then, under the null hypothesis, the large sample likelihood ratio test statistic is approximately distributed as:

$$G = -2\ln(\Lambda) \sim \chi^2((k-1)(d-1))$$

Alternatively, in the high concentration case, we only concern ourselves with what happens when κ becomes large. In this case, we perform a transformation to the likelihood ratio, analogous to the ratio os sums of squares in classical one-way ANOVA, to define the P statistic. As shown by Presnell and Rumcheva [2], the P statistic is approximately distributed as:

$$P = \frac{n-k}{k-1}(\Lambda^{-2/n(d-1)} - 1) \sim F((k-1)(d-1), (n-k)(d-1))$$

Although this was derived to be in the high concentration case, it is worth mentioning that it has been shown in [2] that P performs reasonably well in the low concentration settings. Now note that, by creating a Maclaurin series of the P statistic, we obtain the following relationship between the P statistic and the likelihood ratio test statistic

$$P = \frac{n-k}{k-1}(e^{\frac{G}{n(d-1)}} - 1) = \frac{G}{(k-1)(d-1)} + O_p(n^{-1})$$

With this approximation, in the large sample setting we have that $P \xrightarrow{d} \frac{G}{(k-1)(d-1)}$, which has a scaled χ^2 distribution. It is important to note that an $F((k-1)(d-1), (n-k)(d-1))$ converges to the same scaled χ^2, so that the F distribution correctly approximates the null distribution of P when the sample size is large. This means that the test based on P has applicability in either the high concentration or large sample settings.

Until recently, tests for (2) have been conducted using either a high concentration test by Watson & Williams based on W, which is approximately distributed:

$$W = \frac{(n-k)(\sum_{i=1}^k R_i - R)}{(k-1)(n - \sum_{i=1}^k R_i)} \sim F((k-1)(d-1), (n-k)(d-1))$$

or a test by Mardia & Jupp based on A, which is approximately distributed:

$$A = \frac{(n-k)(\sum_{i=1}^k n_i R_i^2 - n\overline{R}^2)}{n - \sum_{i=1}^k n_i R_i^2} \sim F((k-1)(d-1), (n-k)(d-1))$$

However, these approximations are only valid in the high concentration setting.

Implementation

The main concern of this current work is to comment on an incorrect application of the above P statistic as seen in [1]. In the code provided in the supplemental material for [1], the authors used the following *approximation* to the P statistic

$$P = W + \frac{n-k}{k-1}\Big(\frac{1-(d-2)^2}{2(d-1)^2}\Big)\frac{(\tilde{R}-\overline{R})(1-\overline{R})}{(1-\tilde{R})} + O_p(\kappa^{-2})$$

Thus, the problems that arise from [1] are due to an incorrect application of the theory derived in [2] rather than any shortcomings in the theory. As stated earlier, derived, and verified via simulation in [2], the test based on P outperforms any of the other tests mentioned here or in [2] in both the high-concentration and large-sample settings. So, the choice to use an approximation of P based on the Watson & Williams statistic in [1] is a dangerous misapplication of the results of [2] that does not accurately reflect the efficacy of using the P statistic.

The remainder of this work will focus on re-performing the simulations and examples done in [1] to verify that, indeed, the P statistic is the best choice for performing tests like those in (2).

Power Analysis

For power analysis of the P test, we will reproduce figure 4 from [1] by using a correct implementation of the P statistic along with the code provided in the supplemental material for [1]. We analyze the power by observing how the difference in angle between the two mean directions affects the rejection rate of the test. We will work with circular data and two random samples of directions on the unit circle.

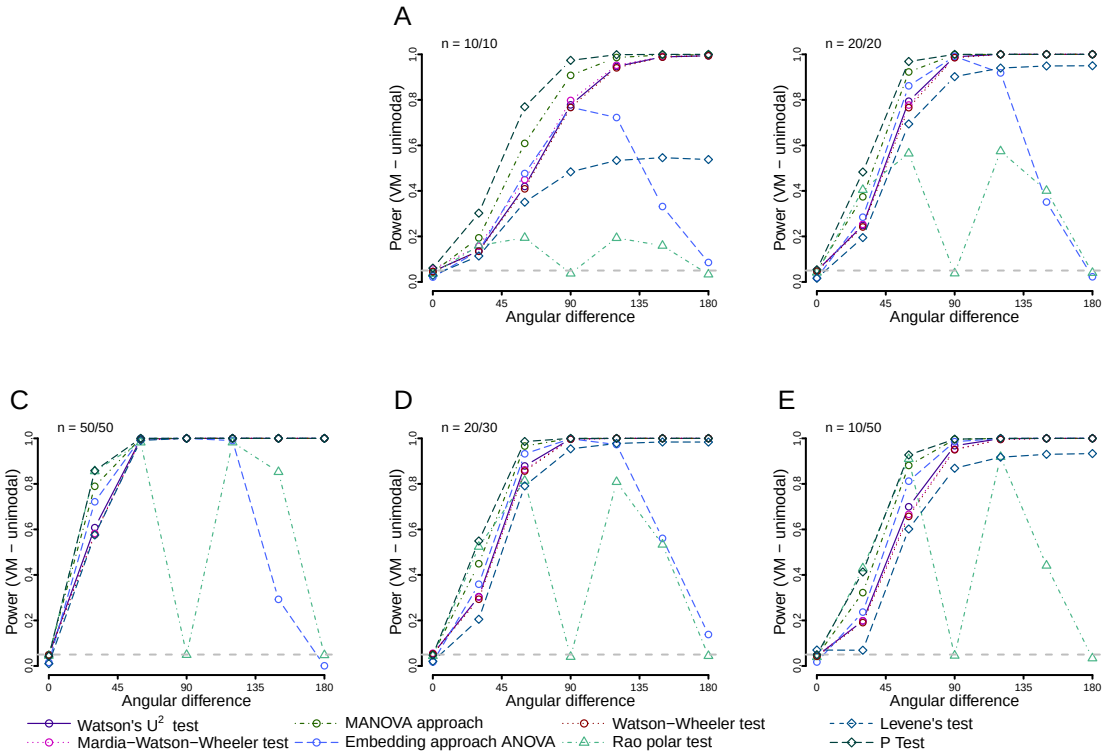


Figure 1: Power Analysis

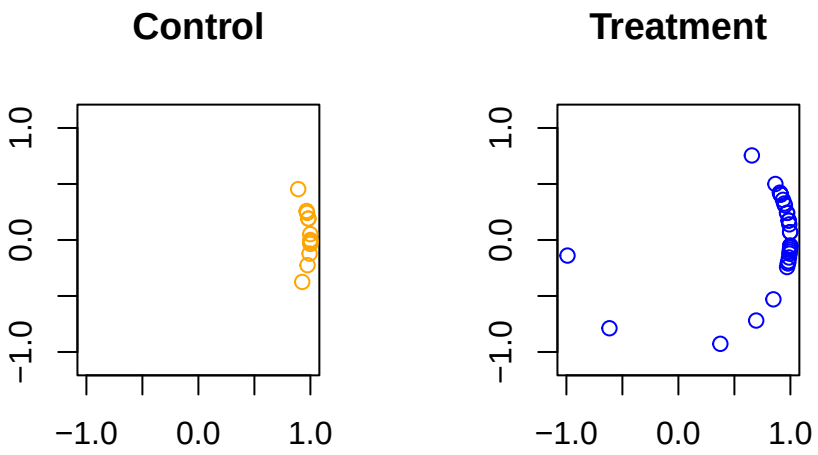
By standardizing one distribution to have mean direction that makes an angle of 0 with the horizontal and we will vary the mean direction of the second distribution to have make angles

$0, 30, 60, 90, 120, 150, 180$ with the horizontal. We perform the tests with sample size combinations $(10, 10), (20, 20), (50, 50), (20, 30), (10, 50)$, and common concentration $\kappa = 2$, just like was done in figure 4 in [1]. We repeat the test for each combination of sample sizes 9999 times and store the rejection rate of the test to obtain the plots.

Real Data Examples

The authors of [1] apply methods for testing equality of mean direction to biological data for bats, and ants. Unfortunately, due to the misapplication of the P statistic, they do not use the P statistic to test for equality of mean direction as they claim it does not perform well. This claim is based on the incorrect implementation of the P statistic, so, in this work, we will test the equality of mean direction in these three cases using the correct P statistic.

We first examine the ant data, where the control group was a certain species of ant trained with both eyes open and the treatment was the same species trained with one eye closed. The researchers were interested in whether or not the ants could transfer visual information from one eye to another. It was suspected that the ants could do this and there should be no difference in mean direction or concentration.

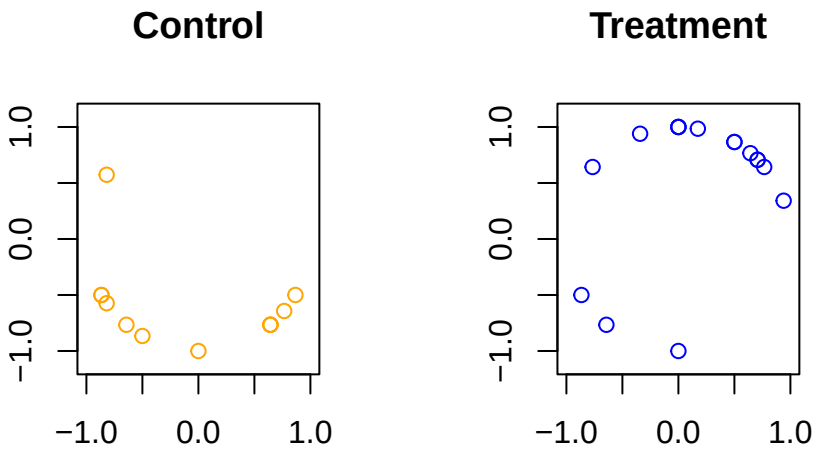



```
## $H0
## [1] "Equal Mean Directions"
##
## $P.stat
## [1] 0.016442
##
## $pval
## [1] 0.8985968
```

Intuition would reason that these two groups have no significantly different mean direction, as the control and treatment groups both have an average orientation in the same direction. The test based on the P statistic detects this and fails to reject the null hypothesis that the mean directions

are equal.

The data on the direction of bats were collected to see if a difference in mean direction existed between bats that were exposed to the sun's position at dusk and those that were not. It was expected that the bats with treatment (no exposure to the sun's direction) would have different mean directions than the control group. The data is plotted below.




```
## $H0
## [1] "Equal Mean Directions"
##
## $P.stat
## [1] 109.212
##
## $pval
## [1] 2.064711e-10
```

It was expected that the bats would have different mean directions, which can be intuited from the plot of the directions. The P test recognizes this difference at the $\alpha = 0.05$ level as we reject the null hypothesis that the mean directions are the same.

Conclusion

There are many tests available for statistical analysis of directional data for testing different flavors of hypotheses. The test based on the P statistic presented in [2] is specifically for testing equality of mean direction while concentrations are held equal and, when properly implemented, it outperforms other tests for this kind of hypothesis. The implementation and corresponding recommendations about the P statistic presented in [1] fail to accurately represent this test. This work was meant to elaborate on the shortcomings of the implementation in [1] and correct any misconceptions it imparted. It is maintained, as stated in [2], that the test based on the P statistic has no reason to not be adopted in preference to the other tests for comparing mean directions.

References

1. Landler, Lukas, et al. “Advice on comparing two independent samples of circular data in Biology.” Scientific Reports, vol. 11, no. 1, 13 Oct. 2021, https://doi.org/10.1038/s41598-021-99299-5.

2. Rumcheva, Pavlina, and Brett Presnell. “An improved test of equality of mean directions for the Langevin-von Mises-Fisher Distribution.” Australian & New Zealand Journal of Statistics, vol. 59, no. 1, Mar. 2017, pp. 119–135, https://doi.org/10.1111/anzs.12183.