Chilean Journal of Statistics
Volume 14, Number 2 (December 2023), Pages 85–98
DOI: 10.32372/chjs.14-02-02

STATISTICAL INFERENCE
RESEARCH ARTICLE

An alternative procedure for testing equality of treatment effects under circular responses

ATANU BISWAS[1], RAHUL BHATTACHARYA[2,*], and TARANGA MUKHERJEE[2]

[1]Applied Statistics Unit, Indian Statistical Institute, Kolkata, India
[2]Department of Statistics, University of Calcutta, Kolkata, India

(Received: 08 July 2022 · Accepted in final form: 27 December 2023)

Abstract

In real life, we often observe circular data (for example, in the fields of geology, meteorology, oceanography and medicine, among others). However, the common practice is to ignore the periodic feature of such observations employing the usual available methods, which often ends in misleading results in terms of inferential measures. The loss of statistical power is a serious issue if the method is applied to clinical trials involving human subjects. In the present article, heuristic procedures for testing equality of treatment effects are developed in the context of clinical trials with the assumption that responses are circular in nature. The developed procedures are studied empirically in terms of statistical power and are compared to relevant existing competitors. The procedures are further applied to a real clinical trial on cataract surgery.

Keywords: Angular data · Clinical trials · Distance metric · Von Mises distribution · Wrapped distributions

Mathematics Subject Classification: Primary 62F03 · Secondary 62F05.

1. INTRODUCTION

The notion of circular data analysis can arise from different real-life situations in geology, meteorology, oceanography, and medicine, among others. Circular data (or angular data) are a special class of observations, which are defined on the circumference of a unit circle; see Mardia and Jupp (2004). The usual continuous probability distributions fail to model circular data because of their bounded domain and periodicity. This necessitates the formulation of a special class of probability models, called circular probability distributions (Rao and Sengupta, 2001). However, most of the clinical studies involving circular responses do not assume any circular model; instead, they stick to the normality assumption. For example, Mihata et al. (2012) conducted a comparison study on the rotator caff muscle under five different conditions, where the total rotational range in motion (in degree) was recorded. They carried out a pairwise comparison of the different conditions and reported p-values of significant tests assuming normality. The rationale behind such an assumption is that in real life, the angular responses lie in the first two quadrants; thus, the aspect of periodicity does not arise. However, the assumption of normality is inappropriate, unless the circular measurements are highly concentrated at a location, producing misleading conclusions.

*Corresponding author. Email: rahul_bhattya@yahoo.com (R. Bhattacharya)

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

In statistical inference, an important aspect is the comparison of two or several populations with respect to location. Since, circular data are periodic, such comparison cannot be carried out based on the methods exclusive to linear data. Another major difficulty of applying linear statistical methods to circular data is that the observations cannot be compared numerically, because of their periodic nature. As an alternative, the observations can be compared with respect to a reference point, called a preferred direction. In the field of medical research, the preferred direction is usually set according to the desired medical condition. For example, in studies related to shoulder movement, it is generally seen that a perfect shoulder allows 90^0 internal rotation (Jain et al., 2013) and hence the preferred direction should be considered as 90^0. Therefore, in an interventional clinical trial with circular outcomes, treatment that has outcomes near the preferred direction can be considered favourable over others. Taking recourse to these facts, Biswas et al. (2015) developed a procedure for testing the equality of treatment effects considering a treatment as favourable if the responses of the treatment have a lower circular distance from the prescribed preferred direction than its competitor. Biswas et al. (2015) applied their methodology to a real-life clinical trial on cataract surgery in which the responses were circular. However, their methodology was limited to the von Mises distribution only, involving two treatments. In fact, most of the inferential studies related to circular responses (as the goodness of fit of Watson and Williams (1956)) are based on the von Mises assumption (von Mises, 1918; Gumbel et al., 1953) due to its applicability.

The present article develops a convenient testing procedure to compare the treatment effects in the context of a clinical trial under the assumption that the responses are circular with a distribution having finite mean direction and concentration. Naturally, the proposed procedure is not limited to von Mises distribution but is applicable to any circular family of distributions (for example, wrapped normal, wrapped Cauchy).

The rest of the article is organized as follows. In Section 2, starting from a short review of circular probability models, the relevant hypothesis testing problem is developed for two populations. A multi-treatment extension of the proposed procedure is described in Section 3. In Section 4, we conduct an extensive simulation study to compare the powers of the proposed and existing procedures. In Section 5, we consider a real clinical trial with circular responses, use the proposed procedure to reach a decision regarding equality of treatment effects, and compare the decision with the actual one. Finally, Section 6 concludes with some related and upcoming issues.

2. A two sample test for equality of treatment effects

2.1 Circular probability models

Circular random variables (RVs) are a special class of continuous RVs that are defined on the circumference of a unit circle. Since, each point on the circumference represents a direction, such a distribution is a way of assigning probabilities to different directions. The range of a circular RV θ, measured in radians, can be taken to be $[0, 2\pi)$ or $[-\pi, \pi)$. The probability density function (PDF) $f(\theta)$ of circular RV θ satisfies the following conditions:

(i) $f(\theta) > 0$, $\forall \theta$;
(ii) $\int_0^{2\pi} f(\theta)\, \mathrm{d}\theta = 1$; and
(iii) $f(\theta) = f(\theta + 2\pi k)$, for any integer k, that is, f is periodic in nature.

Several probability distributions are available in the literature for modelling a circular RV. First, we consider the most popular von Mises or circular normal model, described by the PDF stated as

$$f(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos(\theta - \mu)\right), \quad 0 \le \theta < 2\pi,$$

where $\mu \in [0, 2\pi)$ is the mean direction, $\kappa > 0$ is the concentration parameter and

$$I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \exp(\kappa \cos(\theta)) \cos(p\theta) \mathrm{d}\theta$$

is the modified Bessel function of order $p \geq 0$.

In addition to the von Mises distribution, there are several other distributions to model circular data. The wrapped normal distribution is one of such distributions with PDF expressed as

$$f(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left(- \frac{(\theta - \mu + 2\pi k)^2}{2\sigma^2} \right), \quad 0 \leq \theta < 2\pi,$$

where $\mu \in [0, 2\pi)$ is the mean direction and $\rho = \exp(-\sigma^2/2)$ is the concentration parameter. Another distribution in the sequence is the wrapped Cauchy distribution having the PDF stated as

$$f(\theta) = \frac{1 - \rho^2}{2\pi(1 + \rho^2 - 2\rho\cos(\theta - \mu))}, \quad 0 \leq \theta < 2\pi,$$

where μ is the modal direction and ρ is the concentration parameter. Mean (or modal) direction is analogous to mean (or mode) of the usual RVs. Furthermore, the concentration parameter in a circular probability model acts as opposite to usual variance; as the concentration increases, the accumulation of observations near the mean(or modal) direction increases as well. An account of these models can be found in Fisher (1993), Rao and Sengupta (2001) and Mardia and Jupp (2004). In addition to the standard circular models, a circular probability model related to skew-normal distribution can be found in Hernández-Sánchez and Scarpa (2012).

2.2 Formulating the testing problem

Consider two competing treatments A and B in a phase III clinical trial. We denote the corresponding responses from treatments A and B as Y_A and Y_B, respectively. As the responses are circular, a suitable circular distribution is assumed for Y_g with mean direction μ_g and concentration parameter κ_g, for $g = A, B$. However, numerical comparison of circular responses is not possible and hence such responses are compared with respect to some study-specific reference point, called preferred direction. Consequently, we designate the treatment that has responses nearer to the preferred direction as most promising. Since the responses are circular, the departure from the preferred direction cannot be measured by a straightforward difference, and therefore, we use the notion of circular distances (Mardia and Jupp, 2004). For two points a and b on the circumference of a unit circle, a circular distance is defined as $d(a, b) = \min\{a - b, 2\pi - (a - b)\}$, which is the minimum of the two arc lengths between these points along the circumference. Therefore, treatment A is considered promising over treatment B when the response of a typical A-treated patient is closer to the preferred direction than the response of a typical B-treated patient. However, without loss of generality, we can set the preferred direction at 0^0 through a transformation (Fisher, 1993) without losing any information. Keeping 0^0 as the preferred direction, we find that treatment A is promising over treatment B if $d(Y_A, 0) < d(Y_B, 0)$ or equivalently if either of the events $(0 < Y_A < \pi, Y_A < Y_B < 2\pi - Y_A)$ and $(\pi < Y_A < 2\pi, 2\pi - Y_A < Y_B < Y_A)$ hold. Consequently, the two treatments are considered equally promising if $d(Y_A, 0) = d(Y_B, 0)$ or equivalently if $Y_A = Y_B$ or $Y_A + Y_B = 2\pi$.

Now consider two treatments A and B producing n_A and n_B independent and identically distributed (IID) observations, respectively. Specifically, we denote the (circular) responses corresponding to treatment g by $Y_{g1}, \ldots, Y_{gn_g}$, where we assume that the circular distribution corresponding to treatment g responses has mean direction μ_g, concentration parameter κ_g, for $g = A, B$ and responses from different treatments are independent. Then it follows from the preceding discussion that treatment A is best if $d(\mu_A, 0) < d(\mu_B, 0)$ and the two treatments are equally effective if $d(\mu_A, 0) = d(\mu_B, 0)$. Therefore, testing the equality of treatment effects can be formulated as testing the null hypothesis H_0: $d(\mu_A, 0) = d(\mu_B, 0)$ against the alternative H_1: $d(\mu_A, 0) \neq d(\mu_B, 0)$. Since, $d(\mu_A, 0) = d(\mu_B, 0)$ implies $\mu_A = \mu_B$ or $\mu_A = 2\pi - \mu_B$, the equality of treatment effects holds if $\mu_A = \mu_B$ or $\mu_A = 2\pi - \mu_B$. Note that the distance measure d suffers from the drawback that it is not a conventional distance metric. Consequently, a conventional distance metric $D(a, b) = 1 - \cos(a - b)$ is popularly used to measure the circular distance. If the preferred direction is set at 0^0 and we use the already introduced criterion of better treatment, we find that the inequalities $D(\mu_A, 0) \leq D(\mu_B, 0)$ and $d(\mu_A, 0) \leq d(\mu_B, 0)$ are equivalent and hence the criterion for equally effective treatments remains the same. Therefore, the notion of testing H_0 against H_1 remains valid even if the distance measure is altered.

2.3 Test procedure for equal concentration

For our purpose, we use the transformed responses $d_{gj} = d(Y_{gj}, 0)$, for $j = 1, \ldots, n_g$ and $g = A, B$ as the available set of observations and assume that the response distributions have the same concentration, that is, $\kappa_A = \kappa_B$. Since the observations d_{gj} are linear in nature, an independent sample t-type statistic based on the observations d_{gj} can be used to test H_0. In particular, we suggest the statistic given by

$$Z_c = \frac{\bar{d}_B - \bar{d}_A}{\widehat{v}\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}},$$

where $\widehat{v}^2 = (n_A \widehat{\mathrm{Var}}(d_{A1}) + n_B \widehat{\mathrm{Var}}(d_{B1}))/(n_A + n_B)$ is the pooled estimator of the common variability parameter and $\widehat{\mathrm{Var}}(d_{g1})$ is the sample variance based on d_{gj}, $\bar{d}_g = \sum_{j=1}^{n_g} d_{gj}/n_g$, for $j = 1, \ldots, n_g$ and $g = A, B$. We observe that as the parameter values drift away from the null situation, $|Z_c|$ is expected to be larger, and hence a right-tailed test based on $|Z_c|$ seems appropriate. Since, d_{gj}s are an IID set of linear variables for each g, the central limit theorem coupled with the Slutsky Theorem reveals that under the null hypothesis, Z_c follows approximately a standard normal distribution for large n_A and n_B. Hence, a large sample test of size α rejects H_0 if $|Z_c|$ exceeds the upper percentage point $\alpha/2$ of a standard normal distribution. However, a similar statistic based on the other distance measure D could be defined as

$$Z_c^* = \frac{\bar{D}_B - \bar{D}_A}{\widehat{v}^*\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}},$$

with $\bar{D}_g = \sum_{j=1}^{n_g} D_{gj}/n_g$, for $g = A, B$, $\widehat{v}^{*2} = (n_A \widehat{\mathrm{Var}}(D_{A1}) + n_B \widehat{\mathrm{Var}}(D_{B1}))/(n_A + n_B)$, where $\widehat{\mathrm{Var}}(D_{g1})$ is the usual sample variance based on $D_{g1} = D(Y_{g1}, 0)$. At this point, it is interesting to note that the development of Z_c and Z_c^* is very general and hence are well applicable for a wide class of circular distributions.

2.4 Test procedure for unequal concentration

In real practice, the assumption of equal concentration is difficult to justify apriori and hence we give the development for $\kappa_A \neq \kappa_B$. First of all, we observe from Appendix A.2 on the exact distribution of $d(Y) = d(Y, 0)$ that $\mathrm{E}(d(Y))$ will depend, in general, on the parameters of the distribution. Thus, unlike an equal concentration situation, a naive test based on the estimated mean difference $\bar{d}_B - \bar{d}_A$ cannot be constructed, as $E(\bar{d}_B - \bar{d}_A)$ is different from zero for unequal sample sizes even under the null hypothesis. As a reasonable alternative, we, therefore, develop a Wald type test based on the transformed set of observations assuming further that the circular family under consideration possesses asymptotically normal maximum likelihood (ML) estimators of mean direction. That is, we assume that there exists ML estimator $\hat{\mu}_g$ for $g = A, B$ such that $\widehat{\mu}_g$ follows asymptotically a normal distribution with mean μ_g and asymptotic variance $\sigma_g^2 = 1/I_g$, where I_g is the corresponding Fisher information (may involve parameters). Again, it follows from Appendix A.1 that under such conditions, $d(\widehat{\mu}_g) = d(\widehat{\mu}_g, 0)$ is also asymptotically normal with mean $d(\mu_g, 0)$ and asymptotic variance σ_g^2. Since $d(\widehat{\mu}_g)$ is the ML estimator of $d(\mu_g, 0)$ and, $\widehat{\sigma}_g^2$ is a consistent estimator of σ_g^2, with $g = A, B$, a Wald type statistic for testing H_0 can be expressed as

$$W_c = \left(\frac{d(\widehat{\mu}_B) - d(\widehat{\mu}_A)}{\sqrt{\widehat{\sigma}_B^2/n_B + \widehat{\sigma}_A^2/n_A}} \right)^2 .$$

Thus, a right tailed test based on W_c is appropriate. Further, it follows from Appendix A.1 that W_c follows asymptotically a chi-square distribution with one degree of freedom under the null hypothesis. Therefore, for large sample sizes, a test of size α rejects the null if W_c exceeds the upper percentage point α of a chi-square distribution with one degree of freedom. However, for the other distance measure "D", it follows from the Appendix A.1 that under a similar set of conditions $D(\widehat{\mu}_g) = D(\widehat{\mu}_g, 0)$ is asymptotically normal with mean $D(\mu_g, 0)$ and asymptotic variance $\tau_g^2 = \sin^2(\mu_g)\sigma_g^2$. Then due to the consistency property of ML estimator, $\widehat{\mu}_g$, $D(\widehat{\mu}_g)$ and $\widehat{\tau}_g^2$ are both consistent and hence a Wald type statistic for testing H_0 can be defined as

$$W_c^* = \left(\frac{D(\widehat{\mu}_B) - D(\widehat{\mu}_A)}{\sqrt{\widehat{\tau}_B^2/n_B + \widehat{\tau}_A^2/n_A}} \right)^2 .$$

Since W_c^* follows asymptotically a chi-square distribution with one degree of freedom under the null hypothesis, a concerned large sample test rejects H_0 at size α if W_c^* exceeds upper percentage point α of a chi-square distribution with one degree of freedom. These test procedures are capable of identifying a departure from the equality of true treatment effects for a wide class of circular distributions.

3. A multi-sample test for equality of treatment effects

3.1 Context

In any clinical trial, the problem of multiple treatment comparison frequently arises. However, the complexity involved is increased many folds when the outcome is circular in nature. Therefore, we develop a treatment comparison procedure that involves multiple treatments, when the treatment outcomes are circular in nature. For development, we assume multiple treatments (say $t \geq 3$), each treatment producing circular outcomes based on n_g independent observations from the gth treatment arm.

3.2 Hypotheses

We assume that the distribution of the responses from treatment g is a member of a circular family of distributions with mean direction μ_g and concentration parameter κ_g, for $g = 1, \ldots, t$. However, the definition of treatment effect remains the same as in the two-sample case, and hence testing equality of treatments can be expressed through testing the null hypothesis defined as H_0: $d(\mu_1, 0) = d(\mu_2, 0) = \cdots = d(\mu_t, 0)$ against the general alternative stated as H_1: At least one inequality in H_0. These hypotheses can be equivalently written as H_0: $\boldsymbol{Cd} = \boldsymbol{0}$ against H_1: at least one component of $\boldsymbol{Cd}$ is non zero, where

$$\boldsymbol{C}^{(t-1)\times t} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}, \quad \boldsymbol{d}^{t\times 1} = \begin{pmatrix} d(\mu_1, 0) \\ d(\mu_2, 0) \\ \vdots \\ d(\mu_t, 0) \end{pmatrix}.$$

3.3 Test statistic

The above hypotheses can be thought of as a test of equality of elementary treatment contrasts in the transformed scale. Now, from Appendix A.1, we find that the large sample distribution of $d(\widehat{\mu}_g, 0)$ is normal with mean $d(\mu_g, 0)$ and variance $n_g^{-1}\sigma_g^2$ and that $d(\widehat{\mu}_g, 0)$ are independent, for $g = 1, \ldots, t$. Therefore, it follows that the large sample distribution of $\boldsymbol{C}\widehat{\boldsymbol{d}}$ is $(t-1)$ variate normal with mean vector $\boldsymbol{Cd}$ and dispersion matrix $\Psi = \boldsymbol{C}\mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_t^2)\boldsymbol{C}^\top$, where

$$\widehat{\boldsymbol{d}}^{t\times 1} = \begin{pmatrix} d(\widehat{\mu}_1, 0) \\ \vdots \\ d(\widehat{\mu}_t, 0) \end{pmatrix}.$$

Motivated by the above, we propose a statistic based on the estimated distance metric given by $T_c = (\boldsymbol{C}\widehat{\boldsymbol{d}})^\top(\widehat{\Psi})^{-1}(\boldsymbol{C}\widehat{\boldsymbol{d}})$, where $\widehat{\Psi}$ is a consistent estimator of Ψ. Thus, T_c is expected to be larger whenever the treatment effects drift away from the null configuration, and hence a right-tailed test based on it is appropriate. However, under the null hypothesis, T_c follows asymptotically a chi-square distribution with $(t-1)$ degrees of freedom, and hence a large sample test rejects the null if observed T_c exceeds the upper percentage point α of a chi-square distribution with $(t-1)$ degrees of freedom. However, the corresponding development with the other distance function is straightforward and, therefore we skip details for brevity.

4. Performance evaluation of the proposed test

4.1 Context

To explore the performance (that is, statistical power) of the proposed procedures, we carry out an extensive simulation with programs written in the `R` software using the packages `CircStat` and `Circular`, dedicated for circular data. The average run-time for simulation of size 10,000 is approximately 4 minutes in a 64 bit machine with core i_3 processor. However, for a meaningful evaluation, we need to justify assumptions on the underlying distributions and relevant competitors. Further, the development assumed large samples and consequently we start with the relevant quantile-quantile (QQ) plot for the selected sample sizes, check closeness to normality and envisage the nature of power. Specifically, we consider sample sizes 40 and 80, equally distributed to each treatment arm, and carry out the simulation study with 10,000 repetitions. Furthermore, for a valid comparison, we need to select valid competitors. But competitors in this context are only available for the von Mises distribution. Consequently, we compare the performance of the proposed procedure for von Mises

responses and also include a few other distributions from the circular family of distributions for our purposes.

4.2 Competitors

To describe suitable competitors, we assume that Y_{gi} are independent observations from the von Mises (Mardia and Jupp, 2004) distribution with mean direction μ_g and concentration parameter κ_g, for $i = 1, \ldots, n_g$. As $\cos\mu_A > \cos\mu_B$ is an indicative of the fact that treatment A performs better (that is, has a lower mean direction), Biswas et al. (2015) considered testing H_0: $\cos\mu_A = \cos\mu_B$ with an objective to develop an analogous test for the equality of mean directions. Assuming equal concentration, Biswas et al. (2015) suggested a test statistic given by

$$W^{=} = \frac{\bar{C}_A - \bar{C}_B}{\sqrt{\widehat{\sigma}\,(1/n_A + 1/n_B)}},$$

where $\bar{C}_g = \sum_{i=1}^{n_g} \cos(Y_{gi})/n_g$, for $g = A, B$, $\widehat{\sigma} = (1 + \widehat{\alpha}_2 - 2\widehat{\alpha}_1^2)/2$ with $\widehat{\alpha}_s = (\sum_{g=A,B}\sum_{i=1}^{n_g}\cos(sY_{gi}))/n$, and $s = 1, 2$. Under the alternative hypothesis, the statistic $|W^{=}|$ is expected to be larger, and hence a right-tailed test based on $|W^{=}|$ is suggested. However, under the null hypothesis, for large sample sizes, the statistic $W^{=}$ converges to an RV with standard normal distribution and consequently the square of $W^{=}$ converges to an RV with chi-square distribution with degree of freedom unity. Therefore, a large sample size α test rejects the null hypothesis if the squared $W^{=}$ exceeds the upper percentage point α of a chi-square distribution with one degree of freedom. However, under $\kappa_A \neq \kappa_B$, Biswas et al. (2015) suggested to use the statistic given by

$$W^{\neq} = \left(\frac{T_A^2}{S_A^2} + \frac{T_B^2}{S_B^2} - \left(\frac{T_A}{S_A^2} + \frac{T_B}{S_B^2}\right)^2\right) \Big/ \left(\frac{1}{S_A^2} + \frac{1}{S_B^2}\right),$$

where

$$T_g = \arccos\frac{\bar{C}_g}{\sqrt{\bar{C}_g^2 + \bar{S}_g^2}}, \quad S_g^2 = \frac{1 - A(\widehat{\kappa}_g, 1)}{2N_{gn}A(\widehat{\kappa}_g, 0)},$$

with $\widehat{\kappa}_g$ being the ML estimator of κ_g, for $g = A, B$ and A being a function of modified Bessel functions. Following Biswas et al. (2015), we note that for mean direction parameters not close to zero direction, the asymptotic null distribution of the statistic $W^{\neq}$ is chi-square with one degree of freedom. As larger values of $W^{\neq}$ indicate a deviation from the null hypothesis, a large sample size α test rejects H_0 if $W^{\neq}$ exceeds the upper α percentage point of a chi-square distribution with one degree of freedom.

As another competitor, we consider the procedure suggested by Watson and Williams (1956) to test H_0: $\mu_A = \mu_B$. Under equal concentration $\kappa_A = \kappa_B = \kappa$, the authors suggested the statistic given by

$$V = \frac{(R_A + R_B - R)(n_A + n_B - 2)}{(n_A + n_B - R_A - R_B)},$$

where R_g is the mean resultant length for sample g, with $g = A, B$, and R is the mean resultant length of the combined sample. For moderately large values of κ, V is approximately distributed as F with degrees of freedom $(1, n_A + n_B - 2)$ under the null hypothesis.

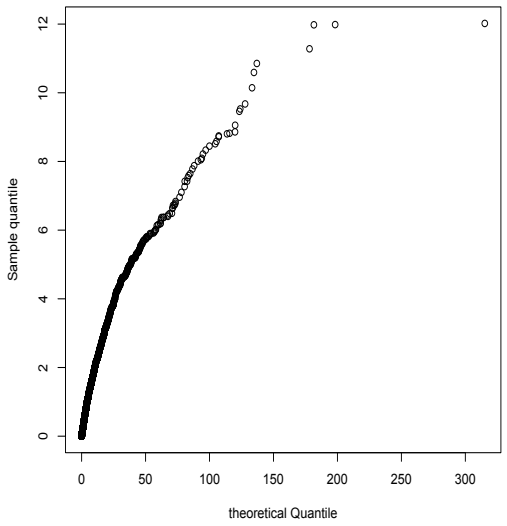
4.3 Simulation studies

First of all, we assume two treatment arms (that is, A and B), response distribution corresponding to treatment g as von Mises with mean direction μ_g and concentration parameter κ_g with n_g observations from treatment g, for $g = A, B$.

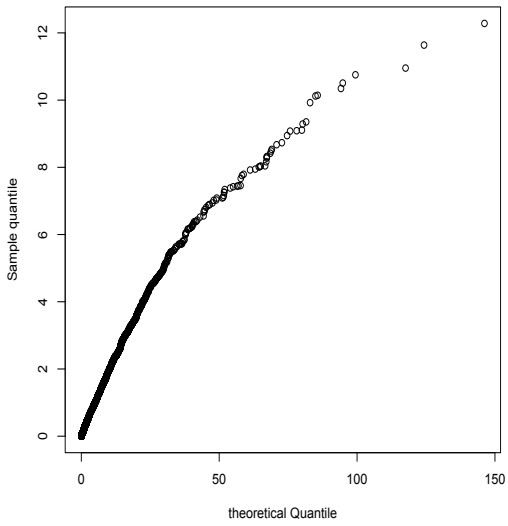
As the proposed tests are based on asymptotic null distributions and the exact null distribution is difficult to obtain even for small samples, we examine the null distributions of Z_c and W_c through simulated data based QQ plots in Figure 1 for moderate choices of (n_A, n_B) and observe the closeness with the desired distributions (that is, asymptotic normality for Z_c test and asymptotic chi square for W_c test). Now for the sake of power computation, we take preferred direction at 0^0 and fix μ_A at 5^0 making treatment A better and vary μ_B upto 120^0 and simulate the power for each of the proposed and competing procedures at 5 percent level of significance. All these are computed assuming equal sample sizes (that is, n) for each treatment arm and different choices of (κ_A, κ_B) and the powers of the corresponding tests are reported in Table 1 and 2. From these power figures, it is readily noticed that both the proposed procedures (that is, Z_c and W_c tests) outperform the $W^=$ test of Biswas et al. (2015) at equal concentration. However, for unequal concentration, apart from minor exceptions, the proposed W_c test is more powerful than the $W^{\neq}$ test of Biswas et al. (2015) when the better treatment has higher concentration. Thus, the proposed procedures are capable of detecting a difference in the treatment effects with considerably higher power under the von Mises responses.

Since the development of the proposed procedure is not based on any specific distributional assumption, it is expected to perform well for any distribution of the circular family. We, therefore, consider wrapped Cauchy and wrapped Normal distributions further to model the distribution of for responses and compute the relevant power figures for the already mentioned Z_c and W_c tests. Specifically, we assume that the responses for treatment g are distributed as wrapped normal with mean direction μ_g and concentration parameter $\rho_g \in (0, 1)$, for $g = A, B$ and compute the relevant powers for the proposed tests at 5 percent level considering different configuration of $(\mu_A, \mu_B, \rho_A, \rho_B)$ and equal sample size n. For the wrapped Cauchy distribution, we continue with the same set of notation as above, but with the exception that μ_g denotes the modal direction. All of these can be found in Tables 3 and 4.

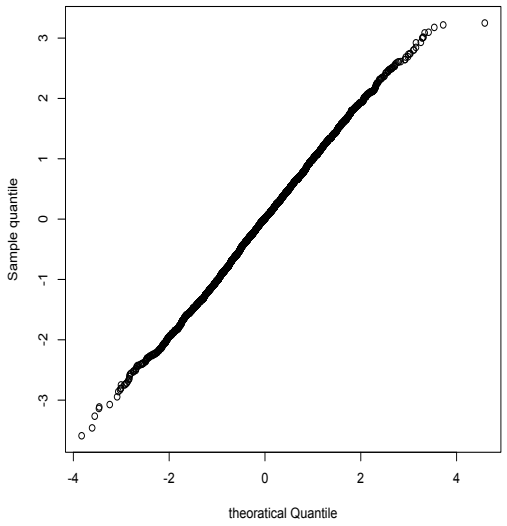
As expected, the powers stated in Tables 3 and 4 increase as μ_B deviates more from μ_A irrespective of the configuration of the concentration parameters, resembling the features for the von Mises responses. In fact, for either response distributions, a little deviation from the null configuration causes a stable increase in power indicating sensitivity. We further observe that power rises more sharply for wrapped Cauchy responses under equal concentration, whereas for unequal concentration, computation with wrapped normal responses produces higher powers compared to those for wrapped Cauchy responses. We further focus towards multi-sample test under the von Mises responses. Specifically, we consider three treatments, where the response from treatment g has a von Mises distribution with mean direction μ_g and concentration κ_g, for $g = 1, 2, 3$. We calculate the powers of the T_c test by setting (μ_1, μ_2, μ_3) in such a way that Treatment 1 is the superior followed by Treatment 3 and Treatment 2. Two sets of concentration parameters are considered ensuring higher and lower concentrations, respectively, for the superior treatment. For the computation, μ_1 is fixed at 5^0, μ_2 and μ_3 are varied according to the assumed ordering. For all the selected set of parameters, the power is computed for equal sample sizes (n) 20 and 40 per treatment arm and are reported in Table 5. It is easy to observe the upward movement of power with increasing difference in treatment effects with the further observation of rapid increase in power when the superior treatment (that is,Treatment 1) has the lower concentration. Thus the T_c test has good ability of detecting a little deviation from the null configuration.



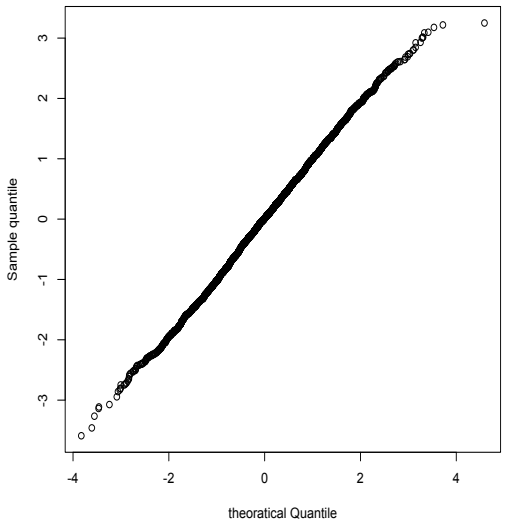
(a) W_C with $n_A = n_B = 20$, $\kappa_A = 2, \kappa_B = 1$.



(b) W_C with $n_A = n_B = 40$, $\kappa_A = 2, \kappa_B = 1$.



(c) Z_C with $n_A = n_B = 20$, $\kappa_A = 1, \kappa_B = 1$.



(d) Z_C with $n_A = n_B = 40$, $\kappa_A = 1, \kappa_B = 1$.

Figure 1. QQ plots of Z_c and W_c for von Mises responses under $\mu_A = \mu_B = 5^0$.

Table 1. Power comparison for von Mises response under equal concentration with $\kappa_A = \kappa_B = 1$.

		Z_c		W_c		$W^=$		V	
μ_A	μ_B	$n = 20$	$n = 40$	$n = 20$	$n = 40$	$n = 20$	$n = 40$	$n = 20$	$n = 40$
5	5	0.0520	0.0500	0.0542	0.0500	0.0520	0.0500	0.0500	0.0500
5	20	0.0730	0.0730	0.1202	0.1768	0.0688	0.0668	0.0556	0.0610
5	45	0.1736	0.2454	0.3638	0.6082	0.1666	0.2398	0.1194	0.1408
5	60	0.3192	0.4862	0.5622	0.8514	0.3124	0.4762	0.1884	0.2470
5	75	0.5132	0.7466	0.7432	0.9660	0.4990	0.7402	0.2766	0.3756
5	90	0.7030	0.9160	0.8866	0.9942	0.6944	0.9130	0.3740	0.5274
5	120	0.9438	0.9978	0.9792	0.9996	0.9410	0.9980	0.6062	0.7956

Table 2. Power comparison for von Mises response under unequal concentration.

	W_c		$W^{\neq}$	
$(\mu_A, \mu_B, \kappa_A, \kappa_B)$	$n = 20$	$n = 40$	$n = 20$	$n = 40$
$(5, 5, 2, 1)$	0.0500	0.0500	0.0500	0.0500
$(5, 30, 2, 1)$	0.2334	0.4600	0.2446	0.4224
$(5, 45, 2, 1)$	0.4658	0.8052	0.4576	0.7248
$(5, 60, 2, 1)$	0.7136	0.9568	0.6412	0.8828
$(5, 75, 2, 1)$	0.8892	0.9970	0.7736	0.9518
$(5, 90, 2, 1)$	0.9626	1.0000	0.8398	0.9750
$(5, 120, 2, 1)$	0.9916	1.0000	0.9238	0.9918
$(5, 5, 1, 2)$	0.0500	0.0500	0.0500	0.0500
$(5, 30, 1, 2)$	0.2814	0.4680	0.1708	0.3312
$(5, 45, 1, 2)$	0.5132	0.7936	0.3744	0.6588
$(5, 60, 1, 2)$	0.7562	0.9628	0.5952	0.8630
$(5, 75, 1, 2)$	0.9082	0.9964	0.7448	0.9452
$(5, 90, 1, 2)$	0.9726	1.0000	0.8274	0.9728
$(5, 120, 1, 2)$	0.9984	1.0000	0.9176	0.9910

Table 3. Power evaluation for wrapped normal and wrapped Cauchy responses under equal concentration with $\rho_A = \rho_B = 0.5$

	Wrapped normal		Wrapped Cauchy	
$(\mu_A, \mu_B, \rho_A, \rho_B)$	$n = 20$	$n = 40$	$n = 20$	$n = 40$
$(5, 5, 0.5, 0.5)$	0.0500	0.0500	0.0500	0.0500
$(5, 30, 0.5, 0.5)$	0.0808	0.1204	0.0974	0.1394
$(5, 45, 0.5, 0.5)$	0.1508	0.2734	0.2106	0.3572
$(5, 60, 0.5, 0.5)$	0.1508	0.5218	0.4220	0.6574
$(5, 75, 0.5, 0.5)$	0.4830	0.7946	0.6610	0.8968
$(5, 90, 0.5, 0.5)$	0.6894	0.9438	0.8564	0.9840
$(5, 120, 0.5, 0.5)$	0.9538	0.9996	0.9906	1.0000

Table 4. Power evaluation for wrapped normal and wrapped Cauchy responses under unequal concentration

	Wrapped normal		Wrapped Cauchy	
$(\mu_A, \mu_B, \rho_A, \rho_B)$	$n = 20$	$n = 40$	$n = 20$	$n = 40$
$(5, 5, 0.8, 0.5)$	0.0500	0.0500	0.0500	0.0500
$(5, 30, 0.8, 0.5)$	0.2564	0.3926	0.1218	0.2322
$(5, 45, 0.8, 0.5)$	0.6564	0.9040	0.5032	0.9192
$(5, 60, 0.8, 0.5)$	0.8630	0.9944	0.8600	0.9952
$(5, 75, 0.8, 0.5)$	0.9726	0.9990	0.9692	0.9998
$(5, 90, 0.8, 0.5)$	0.9960	1.0000	0.9940	1.0000
$(5, 120, 0.8, 0.5)$	0.9994	1.0000	0.9990	1.0000

Table 5. Power comparison of T_c test for $t = 3$.

			$\kappa_1 = 2, \kappa_2 = 2, \kappa_3 = 1$		$\kappa_1 = 1, \kappa_2 = 2, \kappa_3 = 2$	
μ_1	μ_2	μ_3	$n = 20$	$n = 40$	$n = 20$	$n = 40$
5	5	5	0.0500	0.0500	0.0500	0.0500
5	15	10	0.0522	0.0544	0.0539	0.0721
5	25	15	0.0771	0.1104	0.1071	0.2068
5	45	30	0.1384	0.3775	0.2557	0.5617
5	60	45	0.3128	0.8003	0.4886	0.8647
5	75	60	0.6070	0.9836	0.7418	0.9822
5	90	75	0.8478	0.9997	0.9059	0.9988
5	120	100	0.9935	1.0000	0.9936	1.0000

5. Application to a real clinical trial

5.1 Context

To judge the consistency of the decision of the proposed testing procedure in real situations, we consider a real clinical trial involving circular responses. A small incision cataract surgery (SICS) trial was conducted at the Disha Eye Hospital and Research Center, Barrackpore, West Bengal, India, over a period of two years (2008-2010) by Bakshi (2010). We take into account two competing treatments from the study, namely SICS with Snare technique (Rao et al., 1993) and SICS with Irrigating Vectis technique (Masket, 2004) with 19 and 18 observations, respectively. In the original study, the treatments were concluded to be equally performing. Here, we analyze the same data, but in the light of the proposed test, W_c namely.

5.2 Data analysis

For the current analysis, the original study variable is multiplied by 4 in the modulo 2π system to make the preferred direction 0^0. Next, the data are transformed using the circular distance function d as described in Section 2.2. Now we prepare a side-by-side boxplot of the transformed data for each of the treatment groups and provide in Figure 2. However, the boxplot indicates presence of outliers; one for each treatment group. We carry out the test W_c twice: one with the outliers and the other, without the outliers.

The concerned p-value comes out as 0.39 when outliers are retained, whereas removal of outliers produces the updated p-value 0.61. The difference in these p-values indicates the impact of presence of outliers in the analysis. However, for both analyzes, the two treatments (that is, Snare and Irrigating Vectis techniques) do not show significant differences with respect to treatment effects and hence mimic the conclusion obtained in the original trial.

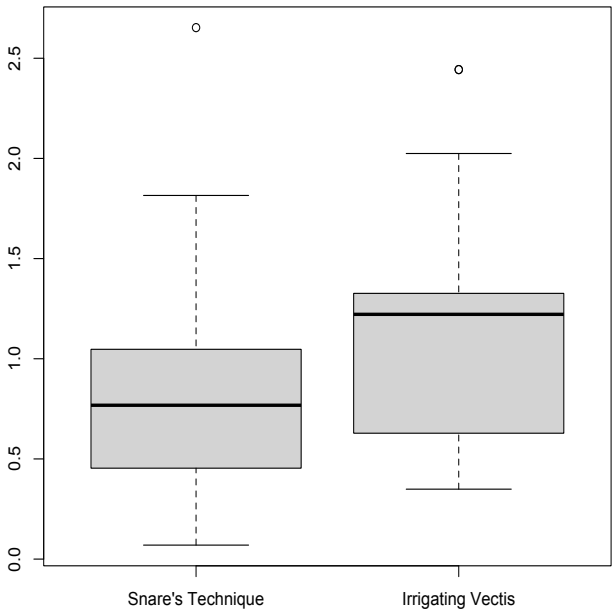


Figure 2. Boxplot for transformed data.

6. Conclusions, limitations, and future research

In the current work, we have suggested a measure of treatment effect in the context of clinical trials with circular responses, developed two treatment and multiple treatment tests for equality of treatment effects and assessed the statistical power of the proposed procedures empirically assuming responses from different circular distributions. In the context of circular data, the proposed procedures are analogous to two sample t test multi-sample contrast based tests. Moreover, the development is based on a relevant asymptotic distribution and hence may produce lower power for small samples. However, in any clinical trial, responses are often influenced by covariates and we intend to develop relevant measures of treatment effect and corresponding testing procedures, when the covariates are either circular or linear in nature.

Appendix A.1

Suppose $X_1, \ldots, X_n$ are IID RVs from a regular circular family of distributions. Assume that the true mean direction is μ and the family of distributions admits the ML estimator of μ. Then, for the distance function d and the ML estimator $\widehat{\mu}$, under certain regularity conditions, we have that $\sqrt{n}(d(\widehat{\mu}) - d(\mu)) \to \mathrm{N}(0, \sigma^2)$ provided $\mu \neq \pi$, where σ^2 is the inverse of the Fisher information contained in μ. However, for $\mu = \pi$, we have $\lim_{n\to\infty} \mathrm{P}\left(\sqrt{n}(d(\widehat{\mu}) - d(\pi)) \leq x\right) = 2\Phi(x/\sigma)$, if $x < 0$, or it is equal to one, if $x \geq 0$.

The previous result can be proven as follows. Under certain regularity conditions the asymptotic distribution of the ML estimator of μ is given by (Mardia and Jupp, 2004)

$$\sqrt{n}(\widehat{\mu} - \mu) \to \mathrm{N}(0, \sigma^2).$$

Now, $d(\widehat{\mu})$ is $\widehat{\mu}$ when $0 < \widehat{\mu} < \pi$ and it is $2\pi - \widehat{\mu}$ when $\pi < \widehat{\mu} < 2\pi$. Thus, d is piecewise continuous and differentiable except at π. For $\mu \neq \pi$, we note that $(d'(\mu))^2 = 1$ and hence by the delta method, the result follows. However, for $\mu = \pi$, assume $x \geq 0$ and consider the representation

$$\begin{aligned} \mathrm{P}\left(\sqrt{n}(d(\widehat{\mu}) - d(\pi)) \leq x\right) &= \mathrm{P}\left(\sqrt{n}(\widehat{\mu} - \pi) \leq x, \widehat{\mu} \leq \pi\right) + \mathrm{P}\left(\sqrt{n}(2\pi - \widehat{\mu} - \pi) \leq x, \widehat{\mu} > \pi\right) \\ &= \mathrm{P}\left(\sqrt{n}(\widehat{\mu} - \pi) \leq 0\right) + \mathrm{P}\left(\sqrt{n}(\widehat{\mu} - \pi) > 0\right) \end{aligned} \tag{1}$$

Since $\sqrt{n}(\widehat{\mu} - \pi)$ converges in distribution to a normal variable with mean zero and variance σ^2, the right hand side of Equation (1) converges to unity as $n \to \infty$. In a similar way, for $x < 0$, we get

$$\mathrm{P}\left(\sqrt{n}(d(\widehat{\mu}) - d(\pi)) \leq x\right) = \mathrm{P}\left(\sqrt{n}(\widehat{\mu} - \pi) \leq x\right) + \mathrm{P}\left(\sqrt{n}(\widehat{\mu} - \pi) \geq -x\right). \tag{2}$$

Now, it follows easily from the asymptotic normality of the distribution of $\widehat{\mu}$ that, for large n, the right hand side of Equation (2) converges to $2\Phi(x/\sigma)$.

Remark: Under the above assumptions, the asymptotic distribution of $D(\widehat{\mu}) = 1 - \cos(\widehat{\mu})$ is normal with mean $D(\mu)$ and variance $\tau^2 = \sin^2(\mu)\sigma^2$, the proof of which follows from the Delta method and the fact that $D(\mu)$ is a continuous differentiable function having $D'(\mu) \neq 0$ except for $\mu = \pi, 2\pi$.

Appendix A.2

The PDF of the distance function $d = d(Y, 0)$ is given by $f_d(y) = f_Y(y) + f_Y(2\pi - y)$, where f is the PDF of the circular RV Y.

The previous result can be proven as follows. Note that

$$\begin{aligned} \mathrm{P}(d(Y) \leq y) &= \mathrm{P}(d(Y) \leq y, 0 < Y < \pi) + \mathrm{P}(d(Y) \leq y, \pi < Y < 2\pi) \\ &= \mathrm{P}(Y \leq y) + \mathrm{P}(2\pi - Y \leq y) = F_Y(y) + 1 - F_Y(2\pi - y). \end{aligned}$$

The result follows from differentiating the above with respect to y.

Acknowledgements

We are grateful to the anonymous referees, associate editor, and editors for their valuable comments that have improved the quality of the originally submitted article.

Author contributions

Conceptualization, A.B., R.B.; data curation, A.B., R.B., T.M.; formal analysis, A.B., R.B., T.M.; investigation, A.B., R.B., T.M.; methodology, A.B., R.B., T.M.; software, T.M.; supervision, A.B., R.B., T.M.; visualization, A.B., R.B. , T.M.; validation, A.B., R.B., T.M.; writing–original draft preparation, A.B., R.B., T.M.; writing–review and editing, A.B., R.B., T.M. All authors have read and agreed to the published version of the article.

Conflicts of interest

The authors declare no conflict of interest.

References

Bakshi, P., 2010. Evaluation of various surgical techniques in brunescent cataracts. Unpublished thesis, Disha Eye Hospital, India.

Biswas, A., Dutta, S., Laha, A.K., and Bakshi, P.K., 2015. Comparison of treatments in a cataract surgery with circular response. Statistical Methods in Medical Research, 25, 2238–2249.

Collett, D. and Lewis, T., 1981. Discriminating between the Von Mises and wrapped normal distributions. Australian Journal of Statistics, 23, 73–79.

Fisher, N.I., 1995. Statistical Analysis of Circular Data. Cambridge University Press, Cambridge, UK.

Gumbel, E.J., Greenwood, J.A., and Durand, D., 1953. The circular normal distribution: Theory and tables. Journal of the American Statistical Association, 48, 131–152.

Hernández-Sánchez, E. and Scarpa, B., 2012. A wrapped flexible generalized skew-normal model for a bimodal circular distribution of wind directions. Chilean Journal of Statistics, 3(2), 129–141.

Jain, N.B., Wilcox III, R.B., Katz, J.N., and Higgins, L.D., 2013. Clinical examination of the rotator cuff. PM& R, 5, 45–56.

Mardia, K.V. and Jupp, P.E., 2004. Directional Statistics. Wiley, Chichester, London, UK.

Masket, S., 2004. The beginning of modern cataract surgery. The evaluation of small incision cataract surgery-A short history of ophthalmologists in progress. Cataract and Refractory Surgeries Today, 77-80.

Mihata, T., McGarry, M.H, Pirolo, J.M, and Lee, T.Q., 2012. Superior capsule reconstruction to restore superior stability in irreparable rotator cuff tears: a biomechanical cadaveric study. The American Journal of Sports Medicine, 40, 2248–2255.

Rao, G.N., Basti, S., Vasavada, A.R., Thomas, R., Braganza, A., Challa, J.K., and George, T., 1993. Extracapsular cataract extraction: Surgical techniques. Indian Journal of Ophthalmology, 41, 195–210.

Rao, J.S. and Sengupta, A., 2001. Topics in Circular Statistics. World Scientific Press, Singapore, 10, p. 4031.

von Mises, R., 1918. Über die "Ganzzahligkeit" der Atomgewichte und verwandete Fragen. Physikalische Zeitschrift, 19, 490.

Watson, G.S. and Williams, E.J., 1956. On the construction of significance tests on the circle and the sphere. Biometrika, 43, 344–352.