A preliminary test estimator in circular statistics

The von Mises distribution plays a pivotal role in circular data analysis. This paper addresses the question as to how much better one can do in estimating its concentration parameter if partial information is available on the mean direction, either (i) through a prior or (ii) through a "pre-test". These two alternate scenarios and the resulting estimators are compared with the standard maximum likelihood estimator, and we explore when one estimator is superior to the other.

**Keywords:** preliminary test estimator, circular statistics, circular normal distribution, concentration parameter, maximum likelihood estimates, bayesian, mean square error, nuisance parameters

**Introduction**

Saleh AME,1 provides an introduction and thorough review on PTEs and Stein-type estimators for various linear models. In statistical inference, the use of prior information on other parameters in a statistical model, usually leads to improved inference on the parameter of interest. Prior information may be (i) known and deterministic which is then incorporated into the model in the form of constraints on the parameter space, leading to a restricted model, or (ii) uncertain and specified in the form of a prior distribution or a verified null hypothesis. In (ii), choosing certain restricted estimators may be justified when the prior information can be quantified i.e. comes with known confidence levels.

In some statistical models, certain parameters are of primary interest while other parameters may be considered as nuisance parameters. One procedure to mitigate the presence of nuisance parameters is to assess what value(s) such nuisance parameter(s) take, by a preliminary test with a null hypothesis restricting the nuisance parameter values. The null hypothesized value(s) of the nuisance parameter are either used or not, depending on whether the observed preliminary test statistic falls in the acceptance or rejection region of the hypothesis. That is, our final estimator for the parameter of interest is thus a linear combination, conditional on whether the preliminary test statistic is in the acceptance or rejection region of the test, and is called a Preliminary Test Estimator (PTE).2-5 and6 were among the first to implement the idea of preliminary test estimation (PTE) in an analysis of variance (ANOVA) framework to analyze the effect of the preliminary test on the estimation of variance. The idea goes back to a suggestion in,1 which considers testing differences between two means after testing for the equality of variances; then using the usual t-test with the pooled estimate for variance, if the variance test shows equality; otherwise, it falls into the category of Behrens-Fisher problem. In these problems it became clear that the performance of the PTE depended heavily on the significance level of the preliminary test. Han c7 were the first to attempt to find an optimum size of significance level for the preliminary test for this two-sample problem.

All Stein-type estimators2,3 involve appropriate test statistics for testing the adequacy of uncertain prior information on the parameter space, which is incorporated into the actual formulation of the estimator. Stein-type estimators adjust the unrestricted estimator by an amount of the difference between unrestricted and restricted estimators scaled by the adjusted test statistics for the uncertain prior information. Usually, the test statistics are the normalized distance between the unrestricted and restricted estimators and follow a noncentral chi-square or an F-distribution. The risk or the MSE of Stein-type estimators depends on the non-centrality parameter, which represents the distance between the full model and restricted model. The PTE may be considered a precursor of the Stein-type estimator. A simple replacement of the indicator function that we will see in the PTE with a multiple of the test statistic, leads to a Stein-type estimator.

**Other estimators**

The CND is the most widely used circular distribution in circular statistics. It plays as central role as the Normal distribution does in usual ‘linear’ statistics. The probability density for a CND denoted by CND(µ, κ), is:

\[
\frac{1}{2\pi I_0(κ)} \exp(κ \cos(α-µ)), 0 \leq α < 2\pi
\]

(1)

The mean direction is also referred to as the preferred direction and the concentration parameter can be thought of as the inverse of variance as it is a measure of concentration around the mean direction. A larger value for implies that observations are more concentrated around the mean direction, while a value of close to 0 implies there may not be a strongly preferred direction. We consider now the maximum likelihood estimates (MLEs) for the parameter in a classical and Bayesian setting.

**Maximum likelihood estimate for concentration parameter**

Given a random sample α_1, ..., α_n from a CND(µ, κ), the MLE for κ when µ in unknown is given by Jammalamadaka SR.6

\[
κ_{MLE} = \text{the solution to: } \frac{1}{n} \sum_{j=1}^{n} \cos(α_j-κ) = \frac{l_{1}(κ)}{l_{0}(κ)}
\]

(2)

When the mean direction µ is known, then the MLE for κ is obtained by substituting µ in place of κ in 2. Since the estimation of concentration parameter is of main interest here, we will denote
A preliminary test estimator in circular statistics

and \( \hat{\kappa} \) the MLEs for \( \kappa \) when sample mean direction is used (if \( \mu \) unknown), and when the mean direction \( \mu \) is known, respectively. In both cases the MLEs carry the usual asymptotic properties. Analogous to the case of a linear Normal distribution, \( \hat{\kappa} \) is superior (has smaller MSE) than \( \hat{\kappa}_{MLE} \).

If the sample comes from a population with population mean \( \mu \), then the MSE \( \text{MSE} (\hat{\kappa}) < \text{MSE} (\hat{\kappa}_{MLE}) \).

\[
\frac{1}{n(1 - \hat{\kappa}/\hat{\kappa}_{MLE}^2)} \leq \frac{1}{n(1 - \hat{\kappa}/\hat{\kappa}_{MLE}^2)}
\]

where we have inequality if and only if \( \mu = \overline{\alpha} \). We denote \( \hat{\kappa} = R^2 / \hat{\mu} \) and \( \hat{\kappa} = V / R \), where \( V = \sum_{i=1}^{n} \cos(\alpha_i - \mu) \), \( R = \sum_{i=1}^{n} \cos(\alpha_i - \overline{\alpha}) \), and \( V \leq R \).

This raises the question whether we can go somewhere in between if we have partial information on \( \mu \).

**MLE for \( \kappa \) when there is a prior on \( \mu \)**

In this semi-Bayesian setting we will place a prior on the nuisance mean direction \( \mu \), a convenient choice being a \( \text{CN} \):

\[
\pi (\mu) = \frac{1}{2\pi I_0 (\mu)} \exp \left( \tau \cos (\mu - \mu_0) \right), \quad 0 \leq \mu < 2\pi
\]

where \( \mu \) and \( \tau \) are the mean direction and concentration parameters for the prior. The value for \( \tau \) measures confidence in the prior mean direction \( \mu_0 \). A larger value of \( \tau \) makes the prior distribution have higher concentration around \( \mu_0 \). A value of \( \tau = 0 \) implies a uniform prior on \( (0, 2\pi) \) for \( \mu \).

In this context, the parameter \( \mu \) has a prior distribution, while the parameter \( \kappa \) is an unknown parameter as in the classical setting. The parameter \( \kappa \) is of interest, while \( \mu \) is the nuisance parameter. We thus blend together classical and Bayesian methods to get an estimate for \( \kappa \).

We begin with the usual likelihood given the data \( (\alpha_1, ..., \alpha_n) \) and independently identically distributed:

\[
L(\mu, \kappa | \alpha_1, ..., \alpha_n) = \frac{1}{2\pi I_0 (\mu)} \exp \left( \kappa \sum_{i=1}^{n} \cos(\alpha_i - \mu) \right), 0 \leq \alpha_i < 2\pi
\]

Given the prior distribution on \( \mu \), we wish to estimate the concentration parameter \( \kappa \). We derive the likelihood function for \( \kappa \) by first averaging our prior knowledge on \( \mu \). The result is the likelihood for \( \kappa \) given by:

\[
\frac{2\pi}{\alpha_0} \int L(\mu, \kappa | \alpha_1, ..., \alpha_n) \pi (\mu) d\mu = L(\kappa | \alpha_1, ..., \alpha_n)
\]

In 6, we begin with joint likelihood for the \( \mu \) and \( \kappa \) which is just the joint density of the data. We then derive marginal distribution for the observations by integrating with respect to \( \mu \). After incorporating our prior knowledge on \( \mu \) and integrating with respect to \( \mu \), we obtain a valid likelihood \( L(\kappa | \alpha_1, ..., \alpha_n) \) for \( \kappa \) which we want to maximize with respect to \( \kappa \).

\[
\frac{2\pi}{\alpha_0} \int L(\mu, \kappa | \alpha_1, ..., \alpha_n) \pi (\mu) d\mu = L(\kappa | \alpha_1, ..., \alpha_n)
\]

where setting \( \Sigma \cos(\alpha_i) = R \cos(\tau) \) and \( \Sigma \sin(\alpha_i) = \sin(\alpha) \) gives

\[
\frac{2\pi}{\alpha_0} \exp \left( (kR \cos(\tau) + \tau \cos(\mu_0)) \cos(\mu) + K \sin(\tau) \sin(\mu) + \tau \sin(\mu_0) \sin(\mu) \right) d\mu
\]

Putting \( \kappa R \cos(\tau) + \tau \cos(\mu_0) = \gamma \cos(\alpha^*) \) and \( \kappa R \sin(\tau) + \tau \sin(\mu_0) = \gamma \sin(\alpha^*) \), and by the definition of the Bessel function \( I_0(\cdot) \), the resulting integral in 8 is our likelihood for \( \kappa \) which is given by:

\[
L(\kappa | \alpha_1, ..., \alpha_n) = \frac{2\pi I_0 \left( \sqrt{k^2 R^2 + 2kR \cos(\pi/2) - \gamma^2} \right)}{(2\pi)^{1/2} I_0(\gamma)}
\]

The likelihood is a ratio of Bessel functions. Given the likelihood, prior distribution on \( \mu \), and data we can find the MLE for \( \kappa \). There is not a simple analytical solution for the MLE, so numerical methods are required for the maximization of 9 with respect to \( \kappa \) leading to the semi-Bayesian MLE \( \hat{\kappa}_{Bay} \).

One interesting comparison would be of the frequentist MLE for \( \kappa \) as in 2 with the semi-Bayesian MLE obtained from 9, using a circular uniform prior distribution on \( \mu \) in the latter, i.e. setting \( \tau = 0 \). In some cases, placing uniform priors result in Bayes estimates that are similar to classical MLEs. Using a circular uniform prior distribution on \( \mu \) in 9, we derive the Fisher Information to find the variance of our semi-Bayesian MLE. From 9, with a circular uniform prior distribution on \( \mu \), we obtain the Fisher Information

\[
\mathbb{E} \left( \left( \frac{\partial^2}{\partial \kappa^2} \log L(\kappa | \alpha_1, ..., \alpha_n) \right) \right) = \frac{2\pi}{\alpha_0} \int \frac{\partial^2}{\partial \kappa^2} \log L(\mu, \kappa | \alpha_1, ..., \alpha_n) \pi (\mu) d\mu
\]

The solution for MLE \( \kappa \) in this case is found by:

\[
\hat{\kappa}_{Bay} = \text{the solution to } \frac{R}{\alpha_0 I_0(\kappa R)} = \frac{I_0(\kappa R)}{I_0(\kappa R)}
\]

Taking another derivative of 11 we have obtained the Hessian where

\[
\hat{\kappa} = \frac{\partial^2}{\partial \kappa^2} \log L(\mu, \kappa | \alpha_1, ..., \alpha_n)
\]

Citation: Nava MM, Jammalamadaka SR. A preliminary test estimator in circular statistics. Biom Biostat Int J. 2017;6(3):342-348. DOI: 10.15406/bbij.2017.06.00168
A preliminary test estimator in circular statistics

A preliminary test estimator (PTE) is a method of estimation that introduces sample-based prior information via a hypothesis test on the nuisance parameter to aid in estimating the parameter of interest. If we fail to reject the null, then we use an estimator evaluated using the null hypothesis value. If we reject the null hypothesis, we use an estimator evaluated using the conditional distribution of the parameter given the sample resultant vector.

The Fisher Information (I) is given by,

\[ I = \frac{I_2(\hat{\kappa}_{\text{Bay}})}{I_0(\kappa)} \left( R^2 - n \right) + n A^2(\hat{\kappa}_{\text{Bay}}) - R^2 \hat{A}^2(\kappa) \]

(13-15)

Then the Fisher Information (I) is given by,

\[ I = \frac{I_2(\kappa)}{I_0(\kappa)} \left( R^2 - n \right) + n A^2(\kappa) - R^2 \hat{A}^2(\kappa) \]

(16)

where substituting the semi-Bayes MLE, the asymptotic variance (V) of \( \hat{\kappa}_{\text{Bay}} \) is given by,

\[ V = \left( \frac{I_2(\kappa)}{I_0(\kappa)} \right)^{-1} \left( R^2 - n \right) + n A^2(\kappa) - R^2 \hat{A}^2(\kappa) \]

(17)

Therefore the asymptotic variance of the MLE can be found using (13-15) of \( \hat{\kappa}_{\text{Bay}} \) :

\[ \hat{k}_{\text{MLE}} \sim Z_{(\alpha/2)} \left( \frac{1}{n(1 - R/k_{\text{MLE}} - R)} \right) \]

\[ \hat{k}_{\text{Bay}} \sim Z_{(\alpha/2)} \left( \frac{I_2(\hat{\kappa}_{\text{Bay}})}{I_0(\hat{\kappa}_{\text{Bay}})} \right)^{-1} \left( R^2 - n \right) + n A^2(\hat{\kappa}_{\text{Bay}}) - R^2 \hat{A}^2(\hat{\kappa}_{\text{Bay}}) \]

Figure 1 displays histograms for \( \hat{\kappa}_{\text{MLE}} \) and \( \hat{\kappa}_{\text{Bay}} \) based on 1000 simulations from the CND with \( \kappa = 1, 3 \), using sample size of \( n = 30 \). In each \( \kappa \) setting the histograms of estimated values are nearly identical for \( \hat{\kappa}_{\text{MLE}} \) and \( \hat{\kappa}_{\text{Bay}} \).

**Preliminary test estimators**

A preliminary test estimator (PTE) is a method of estimation that introduces sample-based prior information via a hypothesis test on the nuisance parameter to aid in estimating the parameter of interest. If we fail to reject the null, then we use an estimator evaluated using the null hypothesis value. If we reject the null hypothesis, we use an estimator based directly on the sample, the usual MLE. The parameter value in the null hypothesis represents our prior knowledge. The idea is when the true parameter value is in or near the null hypothesis value, the PTE will provide a better estimator in terms of mean squared error (MSE), or any other risk function.

We observe data from a CND with unknown mean direction and concentration parameter. We are interested in estimating the concentration parameter, with the mean direction being a nuisance parameter. Our preliminary test has null hypothesis of mean direction equal to a pre-specified direction, versus a two-sided alternative.

Figure 1 Histograms of \( \hat{\kappa}_{\text{MLE}} \) (MLE) and \( \hat{\kappa}_{\text{Bay}} \) (Bayes) with circular uniform prior for 1000 simulations from CND(\( \mu, \kappa \)).

Our PTE for the concentration performs better than the usual MLE and Bayesian estimates for the parameter. The result is similar to the linear case where we have a normal distribution with unknown mean and variance. This methodology can be used to improve the estimation accuracy in many existing applications since the CND is one of the most commonly used distributions in circular statistics.

**Test for assumed mean direction**

Suppose we have observations \( \alpha_1, ..., \alpha_n \) from a CND with both mean direction and concentration parameter unknown. We want to test:

\[ H_0: \mu = \mu_0 \ vs H_1: \mu \neq \mu_0 \]

(18)

In the linear case with data from a Normal distribution, this is parallel to the standard Student’s t-test. In the Likelihood Ratio Test (LRT) is based on the test statistic:

\[ V = \sum_{i=1}^{n} \cos(\alpha_i - \mu_0) = R \cos(\bar{z}_0 - \mu_0) \]

(19)

where we reject the null hypothesis for small values of the test statistic. Note the distribution for \( V \) and \( V^0 \) depend on the nuisance parameter \( \kappa \). However, the exact conditional test for the mean direction of the CN can be obtained by using the conditional distribution of \( R | V \), which is independent of \( \kappa \). \( V \) is the length of the projection of sample resultant vector, \( R \), towards the null hypothesized mean direction, \( \mu_0 = (\cos(\mu_0), \sin(\mu_0)) \). In the conditional test we reject null if \( V \) is too small for a given \( R \), or equivalently, we reject the null if \( R \) is too large for a given \( V \).

To illustrate the geometry of the test, suppose we have polar vector given by the null hypothesis, \( (\cos(\mu_0), \sin(\mu_0)) \). Next, we have
n observations and we calculate the length of projection, c, of the sample resultant vector on the polar vector. Conditioning on the value of c, we find the probability of observing our sample resultant vector, \( R \), and larger values when the null direction is true, conditional on the observed value of \( V_0 = V \).

The space consists of sample resultant vectors that have projection length, c, on the polar vector. Suppose \( R_1 \) and \( R_2 \) are two resultant vector with equal projection length and \( R_1 > R_2 \). Then the direction of \( R_1 \) is further away from \( \mu_o \), than \( R_2 \)’s direction.

For significance level \( \gamma \), we find the rejection region via the exact conditional distribution of \( R | V_0 \). That is, \( r_o \) is the solution to the equation that satisfies:

\[
P(R > r_o | V_0 = v) = \gamma
\]  

As shown in, this critical point \( r_o \) is the solution to:

\[
\int f(r|v) dr = \int_{r_o}^{\infty} \frac{\psi_{\mu_o}(r)}{\psi_{\mu_o}(r_o)} dr = \alpha
\]  

Equations for \( \psi_{\mu_o}(r) \) and \( f_{\mu_o}(v) \) can be found in. There is no analytical solution for \( r_o \) in this case, and provides a table of rejection regions for various values of \( v \). To simplify our hypothesis test we use results in, where approximate confidence intervals for the mean direction are provided. Our test statistic derived from the approximate LRT is broken into two cases:

For \( \bar{R} \leq 0.9 \), reject \( H_0 \) if:

\[
R_2 > V_0^2 + \frac{1}{4n}(2n^2 - V_0^2) Z_{\gamma} \]  

Where \( \bar{R} = \frac{R}{n} \) and \( Z_{\gamma} \) is the upper quantile of the standard Normal distribution.

For \( \bar{R} \geq 0.9 \), reject \( H_0 \) if:

\[
n \log \left( \frac{n^2 - V_0^2}{n^2 - R^2} \right) > \chi^2_{1, \gamma} \]  

These approximations hold well for even small sample sizes when the concentration parameter is high.

The PTE for the concentration parameter

Now we introduce our PTE for estimating the concentration parameter, where the mean direction is a nuisance parameter. Given observations, \( \alpha_1, ..., \alpha_n \), with unknown mean direction and concentration parameter we test our null hypothesized mean direction via the aforementioned hypothesis test. Our PTE is given by:

(i) For \( \bar{R} \leq 0.9 \),

\[
\hat{\kappa}_{\text{PTE}} = \hat{\kappa}_{\text{MLE}} I(Z > Z_{\gamma}) + \hat{\kappa}_{v_0} I(Z < Z_{\gamma})
\]  

where \( Z \) is found by using 22 and solving for \( Z_{\gamma} \).

(ii) For \( \bar{R} \geq 0.9 \),

\[
\hat{\kappa}_{\text{PTE}} = \hat{\kappa}_{\text{MLE}} I(Z > Z_{\gamma}) + \hat{\kappa}_{v_0} I(Z < Z_{\gamma})
\]  

where \( Z \) is found by using 23.

we break the estimator into the two cases according to our hypothesis test. The PTE in either case selects only one of the two estimators according to the result of the hypothesis test. The performance of the PTE depends on the level of the test and the proximity of the true mean direction to the null hypothesized value. We measure performance in terms of mean squared error (MSE) of our estimator over different significance levels \( \gamma \), and different true differences between the mean directions \( \delta = \mu - \mu_o \).

In Figure 2, we observe the simulation-based MSE of the PTE and MLE for the concentration parameter. We perform 1000 simulations of \( \{\alpha_1, ..., \alpha_{20}\} \sim CN(\delta, \kappa) \), for \( j = 1, ..., 50 \). Here \( \delta_1, ..., \delta_{50} \) represent the 50 equally spaced points between 0 and \( \pi \). For each \( \delta_j \), we record the MSE. Each line represents MSE of an estimator over values of \( \delta = \mu - \mu_o \), where \( \delta \) represents the true difference between the population mean direction and the null hypothesized mean direction.

MSE Comparisons: MLE vs. Various PTE

Figure 2 Simulation-based MSE of MLE and PTE for different significance levels \( \gamma \).

For significance levels \( \gamma = 0.10, 0.15, 0.25 \), the PTE performs at least as good as the MLE, and performs better when the true mean direction is closer to the null hypothesized value. For larger significance levels the test requires less evidence to reject the null hypothesis, and when we reject the null the PTE is equivalent to the MLE, \( \hat{\kappa}_{\text{MLE}} \). In Figure 2, we observe that as the significance level increases the PTE is more likely to use \( \hat{\kappa}_{\text{MLE}} \) for smaller values of \( \delta \). To show the vast improvement in our PTE, we examine the mean-square relative efficiency (MRE) of the 2 estimators PTE and MLE, defined by

\[
\varepsilon(\hat{\kappa}_{\text{PTE}}, \hat{\kappa}_{\text{MLE}}) = \frac{\text{MSE}(\hat{\kappa}_{\text{MLE}})}{\text{MSE}(\hat{\kappa}_{\text{PTE}})}
\]  

Values larger than unity imply that the PTE performs better than the MLE. In Figure 3, we have the MRE of the MLE and PTE with \( \gamma = 0.01 \) across all values of \( \delta \). The relative efficiency is greater than 1 for all \( \delta \) less than approximately 0.65 radians. In this example, the PTE can reduce the MSE by 20% when the true difference in
mean directions is small. For $65 < \delta < 1.5$, the MRE is less than 1 implying the MLE has the smaller MLE. This is due to our preliminary test failing to reject the null hypothesis. For $\delta > 1.5$, the preliminary test will almost always reject the null hypothesis value and the PTE will be the same as the MLE resulting in the MRE being equal to one.

**Figure 3** Mean-squared error relative efficiency of MLE and PTE with $\gamma = 0.01$.

In Equation 3, the MRE is maximum for $\delta = 0$, and when the PTE will almost always reject the null for large enough $\delta$ the MRE is equal to one. For $0 < \delta < \pi$ the PTE may reject or fail to reject the null hypothesis depending on the sample observed. In the case it fails to reject $V_0 = \sum_{i=1}^{n} \cos(\alpha_i - \mu_0)$ is no longer minimized at $\mu_0 \mu_0$ since $\mu$ is the population mean (Recall $\mu = \mu - \mu_0$). Therefore $V < R$ if $\mu_0$ is closer to $\mu$ than $\alpha_0$ and $V > R$ if $\alpha_0$ is closer to $\mu$ than $\mu_0$. If the latter case appears more often than the former case for some intermediate values of $\delta$, then the MRE will be less than 1.

In Figure 2, we compare the MLE and PTE with $\gamma = 0.01$ from Figure 2. The PTE with $\gamma = 0.01$ has the best results for smaller values of $\delta$, but could perform worse than the MLE for intermediate values of $\delta$. PTEs with $\gamma = 0.1, 0.15, 0.25$ perform at least as good as the MLE. Now, we illustrate other possibilities that can occur and the performance of the PTE.

In Figure 4, we simulate from four different realities and examine the performance of our PTE for the same significance levels as used in Figure 2. Note that the lines have same labels as in Figure 2. For each plot we have simulated-based MSEs for each line. We perform 1000 simulations of $(\alpha_1, \ldots, \alpha_n) \sim \text{CN}(\delta_0, \kappa)$, for $j = 1, \ldots, 50$. Here $\delta_1, \ldots, \delta_5$ represent the 50 equally spaced points between 0 and $\pi$. For each $\delta_j$, we record the MSE which creates our MSE curve over $\delta$ for each scenario.

First examine that in all scenarios, the PTE with significance level $\gamma = 0.01$ performs the best when the true difference in mean direction is null or small. In the top-left plot we have $n = 50$ simulated observations from $\text{CN}(\delta_0, \kappa = 0.5)$; top-right plot we have $n = 10$ simulated observations from $\text{CN}(\delta_0, \kappa = 0.5)$; bottom-left plot we have $n = 40$ simulated observations from $\text{CN}(\delta_0, \kappa = 3)$; bottom-right plot we have $n = 10$ simulated observations from $\text{CN}(\delta_0, \kappa = 2.5)$.

In applications, the values of $\delta$ and $\kappa$ are unknown. So how do we select the optimal significance level given $n$ observations from $\text{CN}(\mu, \kappa)$? Following the work of, we create tables to find a PTE with minimum and maximum MREs.

Tables were constructed through simulations. Given a sample size $n$ and value for $\kappa$, we generate values from a $\text{CN}(\delta, \kappa)$ distribution to estimate the MRE over a grid of $\gamma$ and $\delta$ values, where $0 \leq \delta \leq \pi$. For each $\gamma$, we compute the maximum MRE,

$E_{\text{min}}$, minimum MRE, $E_{\text{max}}$, over all $\delta$, and record the $\delta$ where $E_{\text{min}}$ is located, $\Delta_{\text{min}}$. For almost all cases the location of the maximum MRE is located at $\delta = 0$ and the function $\text{MRE}(\delta)$ is monotone decreasing from $\delta = 0$ to $\delta = E_{\text{max}}$. For values $\delta > E_{\text{max}}$, the function MRE$(\delta)$ increases back to unity since the PTE will reject the null hypothesized values for larger $\delta$. We then repeated this procedure for different parameter values for $\kappa$.

The mean resultant vector is the normalized length of $R$ since $0 \leq R \leq 1$ and is a measure of concentration for a sample of observations. A value close to 1 implies high concentration and a value close to 0 implies little to no concentration around any single direction. This estimate does not depend on the knowledge of $\kappa$ or of the mean $\mu$ of the distribution. For the CND, there is a one-to-one correspondence between statistic $R$ and the concentration parameter $\kappa$. Given a sample size $n$ and $\kappa$, we observe the average $\overline{R}$ over...
our simulations and use the average as an indication of strength of concentration. In practice, we advise the user to find the sample observed $\tilde{R}$ of the $n$ observations, and then use the column of the table with the nearest $\tilde{R}$ value.

In Table 1, we provide a list of potential PTEs for $n = 20$. The rows list various significance levels $\gamma$ for the PTE ranging from 1% to 50%. The columns list the different observed values for $\tilde{R}$. Suppose we have a sample size of 20 observations and observe $\tilde{R}$ close to $0.779$. Following the procedure in,1 we then decide the minimum MRE preferred is $E_{\text{min}} = 0.471$. Then using the Table 1, the optimal PTE corresponds to using $\alpha = 0.20$. Upon request, we provide tables for various sample sizes, where the tables require only knowledge of sample size, $\tilde{R}$, and the predetermined $E_{\text{min}}$.

Table 1: n = 20: Mallmen and minimum guaranteed efficiencies for the PTE

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$E_{\text{min}}$</th>
<th>$\Delta E_{\text{min}}$</th>
<th>$\Delta E_{\text{max}}$</th>
<th>$\Delta E_{\text{max}}$</th>
<th>$\min(\Delta E_{\text{min}}, \Delta E_{\text{max}})$</th>
<th>$\min(\Delta E_{\text{min}}, \Delta E_{\text{max}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>1.016</td>
<td>1.027</td>
<td>1.676</td>
<td>1.38</td>
<td>1.244</td>
<td>1.075</td>
</tr>
<tr>
<td>0.02</td>
<td>1.008</td>
<td>1.015</td>
<td>0.995</td>
<td>0.688</td>
<td>0.695</td>
<td>0.869</td>
</tr>
<tr>
<td>0.05</td>
<td>1.055</td>
<td>1.103</td>
<td>1.555</td>
<td>1.231</td>
<td>1.105</td>
<td>1.049</td>
</tr>
<tr>
<td>0.1</td>
<td>1.116</td>
<td>1.102</td>
<td>0.991</td>
<td>0.743</td>
<td>0.924</td>
<td>0.968</td>
</tr>
<tr>
<td>0.15</td>
<td>1.103</td>
<td>1.195</td>
<td>1.139</td>
<td>1.007</td>
<td>1.002</td>
<td>1.02</td>
</tr>
<tr>
<td>0.2</td>
<td>1.084</td>
<td>1.144</td>
<td>1.085</td>
<td>1.005</td>
<td>1.001</td>
<td>1.015</td>
</tr>
<tr>
<td>0.25</td>
<td>1.059</td>
<td>1.040</td>
<td>0.965</td>
<td>0.96</td>
<td>0.991</td>
<td>0.997</td>
</tr>
<tr>
<td>0.3</td>
<td>1.049</td>
<td>1.08</td>
<td>1.015</td>
<td>1.006</td>
<td>1.001</td>
<td>1.013</td>
</tr>
<tr>
<td>0.35</td>
<td>1.046</td>
<td>1.032</td>
<td>0.963</td>
<td>0.971</td>
<td>0.995</td>
<td>0.998</td>
</tr>
<tr>
<td>0.4</td>
<td>1.036</td>
<td>1.059</td>
<td>1.002</td>
<td>1.004</td>
<td>1.001</td>
<td>1.01</td>
</tr>
<tr>
<td>0.45</td>
<td>1.019</td>
<td>1.031</td>
<td>0.99</td>
<td>1.001</td>
<td>1.007</td>
<td>1.008</td>
</tr>
<tr>
<td>0.5</td>
<td>1.017</td>
<td>1.011</td>
<td>0.969</td>
<td>0.987</td>
<td>0.998</td>
<td>1.001</td>
</tr>
<tr>
<td>0.6</td>
<td>1.015</td>
<td>1.023</td>
<td>0.993</td>
<td>1.001</td>
<td>1.005</td>
<td>1.003</td>
</tr>
<tr>
<td>0.7</td>
<td>1.013</td>
<td>1.009</td>
<td>0.971</td>
<td>0.999</td>
<td>1.001</td>
<td>1.000</td>
</tr>
<tr>
<td>0.8</td>
<td>1.012</td>
<td>0.962</td>
<td>0.049</td>
<td>0.898</td>
<td>0.898</td>
<td>0.898</td>
</tr>
</tbody>
</table>

Citation: Nava MM, Jammalamadaka SR. A preliminary test estimator in circular statistics. Biom Biostat Int J. 2017;6(3):342-348. DOI: 10.15406/bbij.2017.06.00168
A preliminary test estimator in circular statistics

is much larger for large values of . In this case for large values of the MSE of . The MSE of PTE and Bayes estimators over . In the previously mentioned Bayesian setting 4. If we su the . We 0. Another interesting result is in comparison for the and with . For , but the . Also, for all sample sizes, if less than the PTE table value, then use otherwise use .

Comparison of the PTE and bayes estimators

Both the PTE and Bayes estimators in 12, use prior information on the mean direction to aid in estimation of the concentration parameter. A smaller significance level for the PTE requires stronger evidence to reject the null hypothesized value . A smaller significance level may be chosen to coincide with a stronger belief in the mean direction . In the previously mentioned Bayesian setting of this chapter, a larger value for the concentration parameter focuses our prior distribution around the mean direction . A larger value in parameter represents a stronger belief in prior mean direction .

In Figure 5 we make a comparison of the MSE of our PTE with significance level of 1% with the Bayes estimator with CN prior centered around the null hypothesis value and with . We plot the MSE curve of each estimator over values of .

In each plot, the solid line is MSE curve for the PTE and the dashed line is the MSE curve for the Bayes estimator. For , performs better overall for all sample sizes. For , performs uniformly better than . For estimators have similar performances for small values of , but the MSE for is much larger for large values of .

If , we have different results when comparing the MSES. In all sample sizes of , the MSE of is best for small . Also, for all sample sizes, has the smaller MSE for the larger values of . In this case for large values of , would be the preferred estimator since the performance is better overall.

In reality we do not know the value of , so need a data driven way to select versus . If we suspect a high concentration then we suggest to use , and for a weak concentration then use .

Conclusion

In our work we have shown a superior estimator for the concentration parameter when the mean direction is unknown of a CND. In all cases the PTE has better performance in terms of MSE if is null or small. Another interesting result is in comparison for the PTE we developed an MLE for the concentration parameter based from the prior distribution of mean direction.

Acknowledgements

None.

Conflicts of interest

None.

References


Figure 5 MSE of PTE and Bayes estimators over δ: (—) and (—).