Statistica Sinica **29** (2019), 67-80
doi:https://doi.org/10.5705/ss.202016.0067

INVARIANCE PROPERTIES AND STATISTICAL INFERENCE FOR CIRCULAR DATA

Gianluca Mastrantonio[1], Giovanna Jona Lasinio[2],
Antonello Maruotti[3] and Gianfranco Calise[2]

[1]*Polytechnic of Turin,* [2]*University of Rome "Sapienza"*
and [3]*University of Southampton*

Abstract: Statistical inference on the circle may strongly depend on the chosen reference system. Here, we introduce necessary and sufficient conditions to avoid inferential problems and misinterpretation of parameter estimates for any circular distribution. The construction of invariant distributions, with respect to the reference system, is discussed by introducing specific properties. Numerical examples on simulations and data are presented to corroborate and illustrate theoretical results.

Key words and phrases: Circular data, initial direction, invariance, orientation.

1. Introduction

Circular data (for a review see e.g. Lee (2010)) arise naturally in many scientific fields where observations are recorded as directions or angles. Such data are encountered in environmental science (Bulla et al. (2012); Di Marzio, Panzera and Taylor (2013); Wang and Gelfand (2014); Lagona et al. (2015); Lagona, Picone and Maruotti (2015); Mastrantonio, Maruotti and Jona Lasinio (2015); Mastrantonio, Jona Lasinio and Gelfand (2016); Mastrantonio, Gelfand and Jona Lasinio (2016)), animal movements (Eckert et al. (2008); Langrock et al. (2012, 2014); McLellan et al. (2015); Maruotti et al. (2016); Maruotti (2016)), social science (Gill and Hangartner (2010)) and musicology (Lee and Oh (2008)). Standard techniques cannot be used to analyze circular data, mainly due to the circular geometry of the sample space (for details, see Section 2). Many *ad-hoc* methods and statistical techniques have been developed to analyze and understand circular data (Mardia (1972); Fisher (1996); Mardia and Jupp (1999); Jammalamadaka and SenGupta (2001); Pewsey, Neuhäuser and Ruxton (2013)), leading to important probability distribution theory and inferential results.

Probability distributions for circular data often assume a general structure using the unit circle as support and having a closed-form density. However, circular data have some specific features that should be taken into account in any analysis. Indeed, circular data have no designed zero (i.e. initial direction) or end, and the designation of the natural orientation is arbitrary. Despite having tractable forms, the use of well-known circular distributions can lead to misleading inference if the issues of the initial direction and orientation are overlooked. This article studies the impact on statistical inference of overlooking changes in the reference system.

We define two properties for circular distributions: *invariance under changes of initial direction* (ICID) and *invariance under changes of system orientation* (ICO). We demonstrate that only a distribution holding these properties allows inference independent of the reference system. We give necessary and sufficient conditions for a distribution to be ICID and ICO, and we investigate the nature of existing circular distributions in order to check if the invariance properties hold. We also show that by introducing two additional parameters, accounting for changes in the reference system, an invariant circular distribution can be built from one that is not-invariant.

This article is organized as follows. Section 2 introduces the notation used throughout the article and basic definitions on circular variables. It discusses the ICID and ICO properties and provides examples demonstrating the importance of those properties. Section 3 deals with the construction of invariant distributions. Focusing on widely-used circular distributions, we present the main statistical properties of invariant distributions and we stress the inferential issues of overlooking the initial direction and orientation in empirical analysis. Section 4 considers some numerical examples pointing out the issues in parameters interpretation and model fitting if the mentioned data features are ignored. Section 5 concludes with a summary of the main results and some concluding remarks. Extensive examples of the use of our proposal are reported in the supplementary material; Appendix A illustrates analytical applications, and Appendix B collects some further numerical examples.

2. Invariance in Circular Distributions

Let $\{\mathbb{S}, \mathcal{A}, P\}$ be a probability space, where the sample space $\mathbb{S} = \{(x, y) : x^2 + y^2 = 1\}$ is the unit circle, $\mathcal{A}$ is the $\sigma-$algebra on $\mathbb{S}$ and $P : \mathbb{S} \to [0, 1]$ is the normalized Lebesgue measure on the measurable space $\{\mathbb{S}, \mathcal{A}\}$.

Let $\mathbb{D} = [a, b)$ with $b - a = 2\pi$ and consider the measurable function $\Theta : \mathbb{S} \to \mathbb{D}$, with $\Theta^{-1}(d) = (x, y) = (\cos d, \sin d)$, let $\mathcal{D} = \sigma(\mathbb{D})$ be the $\sigma-$algebra of $\mathbb{D}$ induced by Θ, $A_{\Theta,D} \equiv \{(x, y) : \Theta(x, y) \in D\}$ and $\mathbb{P}_\Theta(D) = P(\Theta^{-1}(D)) =$

$P\left(A_{\Theta,D}\right), \forall D \in \mathcal{D}$. Then the measurable space induced by Θ, $(\mathbb{D}, \mathcal{D}, \mathbb{P}_{\Theta})$, is a probability space.

It follows that Θ is a circular random variable and $\mathbb{P}_{\Theta}$ is its *probability distribution.* Accordingly, for all $d \in \mathbb{D}$, $\Theta^{-1}(d) = \Theta^{-1}(d \mod (2\pi))$. $\mathbb{D}$ can be either continuous or discrete and, in the latter case, it is composed of $l < \infty$ distinct points equally spaced between 0 and 2π with $\mathbb{D} \equiv \{2\pi j/l\}_{j=0}^{l-1}$.

If $\mathbb{D}$ is a continuous domain, Θ is a continuous circular variable and $\mathbb{P}_{\Theta}$ is the Lebesgue measure; if $\mathbb{D}$ is discrete, Θ is a discrete circular variable and $\mathbb{P}_{\Theta}$ is its counting measure. In either case, let f_{Θ} be the probability density function (pdf) of Θ, with $f_{\Theta} = d\mathbb{P}_{\Theta}/dP_{\Theta} : \mathbb{D} \to \mathbb{R}^{+}$.

In the representation of circular variables, a key role is played by the initial direction and the orientation of the domain (clockwise or anti-clockwise). Both are uniquely determined by the choice of the orthogonal reference system on the plane. Any statistical tool for circular variables should be invariant with respect to different choices of the reference system to avoid conflicting or misleading conclusions, see Section 3.3.

With $\boldsymbol{\psi}$ a vector of parameters, ICID and ICO distributions are defined as follow:

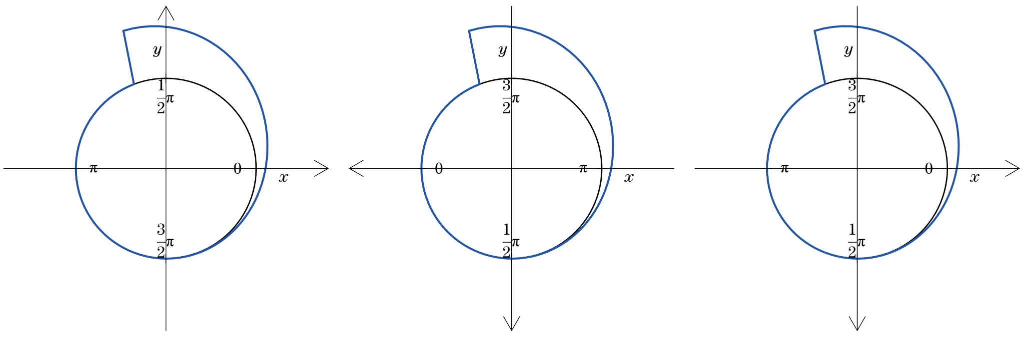
Definition 1. *A probability density function f_{Θ} is invariant under changes of initial direction (ICID) if, for all $\xi \in \mathbb{D}$ and $\boldsymbol{\psi} \in \boldsymbol{\Psi}$, there exists $\boldsymbol{\psi}^{*} \in \boldsymbol{\Psi}$ such that $f_{\Theta}(\theta|\boldsymbol{\psi}) = f_{\Theta}(\theta - \xi|\boldsymbol{\psi}^{*})$ for all $\theta \in \mathbb{D}$.*

Definition 2. *A probability density function f_{Θ} is invariant under changes of the reference system orientation (ICO) if, for all $\boldsymbol{\psi} \in \boldsymbol{\Psi}$, there exists $\boldsymbol{\psi}^{*} \in \boldsymbol{\Psi}$ such that $f_{\Theta}(\theta|\boldsymbol{\psi}) = f_{\Theta}(-\theta|\boldsymbol{\psi}^{*})$ for all $\theta \in \mathbb{D}$.*

The following examples demonstrate that inference should not depend on the reference system. In Figure 1 (a), a wrapped skew normal (WSN) density (Pewsey (2000)) is plotted. The origin (initial direction) is chosen according to a geographical template and set to East. The orientation is anti-clockwise. By changing the initial direction (Figure 1 (b)) or the system orientation (Figure 1 (c)), we can obtain WSN pdfs with shapes exactly as the one in Figure 1 (a): there exists $\boldsymbol{\psi}^{*}$ such that $f_{\Theta}(\theta|\boldsymbol{\psi}) = f_{\Theta}(\theta - \pi|\boldsymbol{\psi}^{*})$ and $f_{\Theta}(\theta|\boldsymbol{\psi}) = f_{\Theta}(-\theta|\boldsymbol{\psi}^{**})$. This example is not sufficient to prove that the WSN verifies the ICID and ICO because the properties must hold for all possible reference systems; the formal demonstration is given in Section 3.3 using Theorem 1, to be introduced in the next section.

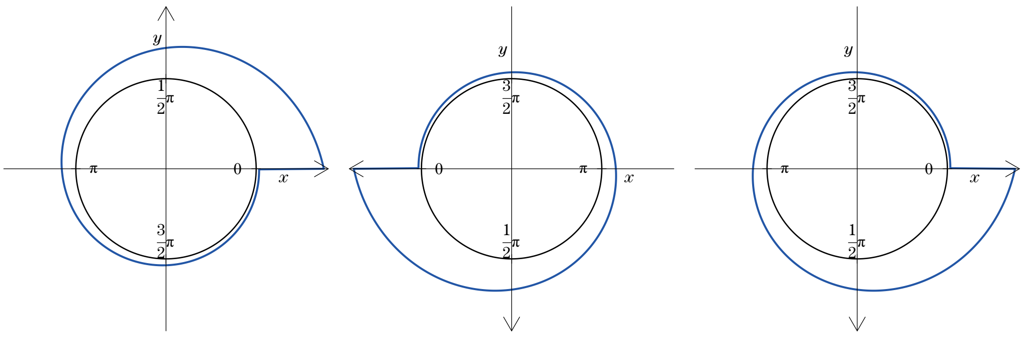
A further example is provided in Figure 1 (d): a wrapped exponential (WE)

Wrapped skew normal



(a) (b) (c)

Wrapped exponential



(d) (e) (f)

Figure 1. Probability density functions of a WSN (a-c) and a WE (d-f) under different initial directions and orientations. The arrows indicate the axis orientation.

distribution (Jammalamadaka and Kozubowski (2004)). We proceed as before, changing the initial direction (Figure 1 (e)) and the system orientation (Figure 1 (f)). The intuition tells us that a necessary (but not sufficient) condition for a circular distribution to be ICID and ICO is that the circular variance must be constant across reference systems (technical details are given in Section 3.2). We find that this holds only for the densities in Figure 1 (e) and (f) that, on the other hand, concentrate their probability mass in different portion of the circle with respect to Figure 1 (d). Hence the wrapped exponential is not ICID nor ICO.

3. Building Invariant Circular Distributions

In this section we introduce necessary and sufficient conditions for a distribution to verify ICID and ICO. We then derive the ICID and ICO counterparts of existing distributions.

3.1. Invariance: necessary and sufficient conditions

Theorem 1. *Let* $f_\Theta(\cdot|\boldsymbol{\psi})$ *be the pdf of the circular variable* $\Theta \in \mathbb{D}$, *with* $\boldsymbol{\psi} \in \boldsymbol{\Psi}$. *Let* $\Theta^* = \delta(\Theta + \xi)$, *with* $\delta = \{-1, 1\}$ *and* $\xi \in \mathbb{D}$, *and* $f_{\Theta^*}(\cdot|\boldsymbol{\psi}^*)$, $\boldsymbol{\psi}^* \in \Psi^*$. *Then* f_Θ *is ICID and ICO iff* $f_{\Theta^*}(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\theta^*|\boldsymbol{\psi}^*)$ *almost everywhere with* $\boldsymbol{\Psi}^* \equiv \boldsymbol{\Psi}$.

Proof. From the rule of variable transformation we have that

$$f_{\Theta^*}(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\delta\theta^* - \xi|\boldsymbol{\psi}), \tag{3.1}$$

where $\boldsymbol{\psi}^*$ is a function of $(\boldsymbol{\psi}, \delta, \xi)$.

For sufficiency, with (3.1) true for all $\theta^* \in \mathbb{D}$, $\xi \in \mathbb{D}$, $\delta \in \{-1, 1\}$, as long as f_{Θ^*} and f_Θ belong to the same parametric family, $\boldsymbol{\Psi}^* \equiv \boldsymbol{\Psi}$. Then we can write

$$f_\Theta(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\delta\theta^* - \xi|\boldsymbol{\psi}). \tag{3.2}$$

Set $\delta = 1$ in (3.2) to satisfy Definition 1 and set $\delta = -1$ and $\xi = 0$ to satisfy Definition 2.

For necessity, according to the ICID and ICO properties, we have $f_\Theta(\delta\theta^* - \xi|\boldsymbol{\psi}) = f_\Theta(\theta^*|\boldsymbol{\psi}^{**})$, where $\boldsymbol{\psi}^{**} \in \boldsymbol{\Psi}$. Since $f_{\Theta^*}(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\delta\theta^* - \xi|\boldsymbol{\psi})$ (see (3.1)) we have $f_{\Theta^*}(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\theta^*|\boldsymbol{\psi}^{**})$ which implies $\boldsymbol{\psi}^* \equiv \boldsymbol{\psi}^{**}$, ending the proof.

It is always possible to transform non-invariant pdf so to obtain its ICID and ICO version.

Proposition 1. *If* $f_\Theta(\cdot|\boldsymbol{\psi})$ *is not ICID and ICO, the density* $f_{\Theta^*}(\cdot|\boldsymbol{\psi}^*)$, *with* $\Theta^* = \delta(\Theta + \xi)$, $\delta \in \{-1, 1\}$, $\xi \in \mathbb{D}$, *and* $\boldsymbol{\psi}^* = (\boldsymbol{\psi}, \delta, \xi) \in \boldsymbol{\Psi}^*$, *is ICID and ICO.*

Proof. Let $\Theta^{**} = \delta^*(\Theta^* + \xi^*)$. Following Theorem 1, if $f_{\Theta^{**}}$ belongs to the same parametric family as f_{Θ^*}, then f_{Θ^*} is ICID and ICO. Because $f_{\Theta^{**}}(\theta^{**}|\boldsymbol{\psi}^{**}) = f_{\Theta^*}(\delta^*\theta^{**} - \xi^*|\boldsymbol{\psi}^*)$ and $f_{\Theta^*}(\theta^*|\boldsymbol{\psi}^*) = f_\Theta(\delta\theta^* - \xi|\boldsymbol{\psi})$, we have $f_{\Theta^{**}}(\theta^{**}|\boldsymbol{\psi}^{**}) = f_{\Theta^*}(\delta^*\theta^{**} - \xi^*|\boldsymbol{\psi}^*) = f_\Theta(\delta(\delta^*\theta^{**} - \xi^*) - \xi|\boldsymbol{\psi})$. Now if $\delta^{**} = \delta\delta^*$ and $\xi^{**} = (\delta\xi^* + \xi)$, we can write $f_{\Theta^{**}}(\theta^{**}|\boldsymbol{\psi}^{**}) = f_\Theta(\delta^{**}\theta^{**} - \xi^{**}|\boldsymbol{\psi})$. Bearing in mind that $\delta^{**} \in \{-1, 1\}$ and $\xi^{**} \in \mathbb{D}$, Θ^{**} is obtained starting from Θ by transforming Θ^* and the vector of parameters $(\boldsymbol{\psi}, \delta^{**}, \xi^{**})$ belonging to the same space $\boldsymbol{\Psi}^*$ of $\boldsymbol{\psi}^* = (\boldsymbol{\psi}, \delta, \xi)$. Then $f_{\Theta^{**}}$ and f_{Θ^*} belong to the same parametric family. Thus one can get the *invariant* version of any circular density.

3.2. Invariance: statistical properties

A number of circular distributions have been studied in the literature, and their characteristic functions, as well as the trigonometric moments, circular means, and concentrations have been defined. Here we show how to obtain these quantities when ICID and ICO distributions are obtained by using Proposition 1.

The trigonometric moments, $\alpha_p = E\cos p\Theta$ and $\beta_p = E\sin p\Theta$, of a circular density are related to the characteristic function $\varphi_\Theta(p)$ of Θ. As $\varphi_\Theta(p) = E(\exp(ip\Theta)) = \alpha_p + i\beta_p$. Let $c_p = |\varphi_\Theta(p)| = \sqrt{\alpha_p^2+\beta_p^2}$ and $\mu_p = \text{atan}^*(\beta_p/\alpha_p)$ (atan* is the modified *inverse tangent function*, see for example Jammalamadaka and SenGupta (2001)). It is well known that $\varphi_\Theta(p) = c_p\exp(i\mu_p) = c_p\cos\mu_p + ic_p\sin\mu_p$, and then $\alpha_p = c_p\cos\mu_p$ and $\beta_p = c_p\sin\mu_p$. When $p = 1$, the quantities μ_1 and c_1 are called the *circular mean* and the *circular concentration*, respectively. Now let $\Theta^* = \delta(\Theta+\xi)$ and suppose that the density of Θ is not-invariant. Following Proposition 1, the density of Θ^* is ICID and ICO. Θ^* is a linear transformation of Θ and its characteristic function is then

$$\varphi_{\Theta^*}(p) = e^{ip\delta\xi}\varphi_\Theta(\delta p) = c_{\delta p}e^{i(p\delta\xi+\mu_{\delta p})}. \tag{3.3}$$

$\varphi_\Theta(-p)$ the complex conjugate of $\varphi_\Theta(p)$ and, since $|\varphi_\Theta(-p)| = |\varphi_\Theta(p)|$, then $c_p = c_{-p}$, $\alpha_{-p} = \alpha_p$, $\beta_{-p} = -\beta_p$ and it follows that $\mu_{-p} = -\mu_p$. Then (3.3) can be written as $c_p\exp(i(p\delta\xi+\delta\mu_p)) = c_p\cos(p\delta\xi+\delta\mu_p) + ic_p\sin(p\delta\xi+\delta\mu_p)$. If α_p^*, β_p^*, μ_p^* and c_p^*, are the trigonometric moments, the circular mean, and concentration of the random variable Θ^*, we have

$$\alpha_p^* = c_p\cos(p\delta\xi+\delta\mu_p) = c_p\cos(p\xi+\mu_p), \tag{3.4}$$

$$\beta_p^* = c_p\sin(p\delta\xi+\delta\mu_p) = \delta c_p\sin(p\xi+\mu_p), \tag{3.5}$$

with $\mu_p^* = p\delta\xi+\delta\mu_p$ and $c_p^* = c_p$. For all invariant distributions, (3.4) and (3.5) can be used to compute the trigonometric moments when the reference system is changed.

3.3. Examples of inferential problems

Consider a WE(λ) circular variable Θ, first we check whether this distribution is ICID and ICO. The WE (Jammalamadaka and Kozubowski (2004)) has density

$$f_\Theta(\theta|\lambda) = \frac{\lambda e^{-\lambda\theta}}{1-e^{-2\pi\lambda}},\ \lambda > 0. \tag{3.6}$$

We find the density of the random variable Θ^*,

$$f_{\Theta^*}(\theta^*|\lambda,\delta,\xi) = \frac{\lambda e^{-\lambda(\delta\theta^*-\xi)}}{1-e^{-2\pi\lambda}}. \tag{3.7}$$

To see that the WE is not ICID and ICO, it is sufficient to prove that for a given λ, δ, and ξ, there is no λ^* such that (3.7) can be written as a WE density; this would necessitate

$$\frac{\lambda e^{-\lambda(\delta\theta^*-\xi)}}{1-e^{-2\pi\lambda}} = \frac{\lambda^* e^{-\lambda^*(\theta^* \bmod (2\pi))}}{1-e^{-2\pi\lambda^*}}. \tag{3.8}$$

In (3.8) the modulus on the right side is required to ensure that the density is well defined. To show that the WE is not ICID and ICO, it is sufficient to show that $-\lambda(\delta\theta^* - \xi) \neq -\lambda^*(\theta^* \bmod (2\pi))$ for a set of values. For example, at $\Theta^* = 2\pi$, with $\delta = 1$, we obtain $-\lambda(2\pi - \xi) = -\lambda^* 0$; thus, the WE is not ICID and ICO. The density of Θ^* in (3.7) is the invariant version of the wrapped exponential (IWE) and it depends on parameters λ, δ, ξ. The domain of Θ^* depends on δ and ξ and, as this may lead to issues in model fitting, we prefer to write (3.7) as

$$f_{\Theta^*}(\theta^*|\lambda,\delta,\xi) = \frac{\lambda e^{-\lambda[(\delta\theta^*-\xi) \bmod (2\pi)]}}{1-e^{-2\pi\lambda}},\ \theta^* \in \mathbb{D}.$$

The trigonometric moments of the WE (see Jammalamadaka and Kozubowski (2004)) are $\lambda/\sqrt{\lambda^2+p^2}\cos\{\text{atan}^*(p/\lambda)\}$ and $\lambda/\sqrt{\lambda^2+p^2}\sin\{\text{atan}^*(p/\lambda)\}$. From them, using the results of Section 3.2, we can derive those of the IWE: $\alpha_p = (\lambda/\sqrt{\lambda^2+p^2})\cos(p\xi + \text{atan}^* p/\lambda)$ and $\beta_p = \delta(\lambda/\sqrt{\lambda^2+p^2})\sin(p\xi + \text{atan}^* p/\lambda)$. The circular mean is then $\mu_1 = \delta\xi + \delta\text{atan}^*(1/\lambda)$ and the circular concentration is $c_1 = \lambda/\sqrt{\lambda^2+1^2}$.

Now consider n independent observations $\theta_1, \dots \theta_n$, WE distributed, in a reference system with zero direction set to North and clockwise orientation. Consider finding the maximum likelihood estimator $\hat{\lambda}$ (MLE) of λ. The WE density, (3.6), has the same functional form of the truncated exponential defined over the domain $[0, 2\pi)$. Then, likelihood functions based on the WE or the truncated exponential lead to the same MLE. For a truncated exponential likelihood over a given interval $[a, b)$, the MLE of λ exists only if the arithmetic mean of the observations is smaller than $(b+a)/2$ (Deemer and Votaw (1955)); in our setting $\hat{\lambda}$ is defined only if $\sum \theta_i/n < \pi$. Assume that we observed the same circular variable but we recorded its values using an anticlockwise orientation. The data have the same nature as before, but we have changed the reference system. In it the value of the circular variable is $2\pi - \theta_i$, and the MLE of λ exists only if $2\pi - \sum \theta_i/n < \pi \Rightarrow \sum \theta_i/n > \pi$. The two conditions for the existence of $\hat{\lambda}$ in the two reference systems, $\sum \theta_i/n < \pi$ and $\sum \theta_i/n > \pi$, cannot both hold. It follows that the inference with the WE depends on the choice of the reference system orientation.

If we record the values of the circular variable in another reference system with the zero direction set to be the angle $d \in [0, 2\pi)$ of the first system, the observed values are $\theta_i - d$. If $d \geq \max_{i,i=1,\dots,n} \theta_i$, then $\theta_i - d = 2\pi + \theta_i - d$. Here the condition for the existence of $\hat{\lambda}$ is $2\pi + \sum \theta_i/n - d < \pi \Rightarrow \sum \theta_i/n < d - \pi$. If $d = \pi$, the MLE does not exist.

The MLEs of λ, δ, and ξ, under an IWE likelihood, always exist. In the original reference system $\hat{\lambda}$ exists if $\sum [(\delta\theta_i - \xi) \bmod (2\pi)]/n < \pi$, that is at least when $\delta = 1$ and $\xi = 0$ if $\sum \theta_i/n < \pi$ and $\delta = -1$ and $\xi = \epsilon$ otherwise where $\epsilon \in (0, \min_{i,i=1,\dots,n} \theta_i)$. In the second reference system (anticlockwise orientation) the condition is $\sum \{(\delta 2\pi - \delta\theta_i - \xi) \bmod (2\pi)\}/n < \pi$; this hold at least when $\delta = -1$ and $\xi = 0$ if $\sum \theta_i/n < \pi$, and $\delta = 1$ and $\xi = \epsilon$ otherwise. In the third reference system (changed zero direction) the existence condition is $\sum \{(\delta\theta_i - \delta d - \xi) \bmod (2\pi)\}/n < \pi$ which holds when $\delta = 1$ and $\xi = 2\pi - d$ if $\sum \theta_i/n < \pi$, and $\delta = -1$ and $\xi = d + \epsilon$ otherwise. We are not able to find a closed form for the MLE of δ and ξ, but the log-likelihood is proportional to $-\sum [(\delta\theta_i^* - \xi)]$ that is a finite function of (δ, ξ) and this is a sufficient condition for the existence of their MLEs.

As a further example, consider a discrete circular variable Θ that can assume l equally spaced points on the circle, such variable is said to follow a wrapped Poisson (WP) distribution with parameter $\lambda > 0$ (Mardia and Jupp (1999)), $\Theta \sim WP_l(\lambda)$, if it has pdf $\sum_{k=0}^{\infty} \lambda^{\theta(l/2\pi)+kl} e^{-\lambda}/(\theta l/2\pi + kl)!$. The WP is not ICID and ICO (see the Supplementary material for details) but, using Proposition 1, we can obtain its invariant version (IWP). It has density

$$f_{\Theta^*}(\theta^*|\lambda, \delta, \xi) = \sum_{k=0}^{\infty} \frac{\lambda^{[(\delta\theta^* - \xi) \bmod (2\pi)]l/2\pi + kl} e^{-\lambda}}{[\{(\delta\theta^* - \xi) \bmod (2\pi)\}\, l/2\pi + kl]!},\ \theta^* \in \mathbb{D}. \tag{3.9}$$

Now let $\Theta \sim WP_l(\lambda)$ with n samples θ_i, in a reference system with zero direction North and clockwise orientation. Let $\hat{\lambda}$ be the MLE of λ and $\hat{\mu}_1$ and $\hat{c}_1$ the associated circular mean and concentration, respectively, $\mu_1 = \lambda \sin(2\pi/l)$ and $c_1 = \exp[-\lambda\{1 - \cos(2\pi/l)\}]$. If we change orientation, in the new reference system the circular observations are $2\pi - \theta_i$. Since the transformation $2\pi - \theta_i$ is linear and the data are the same as in the first reference system, the circular concentration should remain the same, $\hat{c}_1^* = \hat{c}_1$, while the circular mean should be $\hat{\mu}_1^* = 2\pi - \hat{\mu}_1$. Since $\hat{c}_1 = \exp[-\hat{\lambda}\{1 - \cos(2\pi/l)\}]$, then $\hat{c}_1^* = \hat{c}_1$ only if $\hat{\lambda}^* = \hat{\lambda}$ and it follows that $\hat{\mu}_1^* = 2\pi - \hat{\mu}_1$ never holds, since $\hat{\lambda}^* \sin(2\pi/l) \neq 2\pi - \hat{\lambda}\sin(2\pi/l)$ if $\hat{\lambda}^* = \hat{\lambda}$.

If we take $d \in [0, 2\pi)$ as the new zero direction, the circular observations are $\theta_i - d$. We expect $\hat{c}_1^* = \hat{c}_1$ and $\hat{\mu}_1^* = \hat{\mu}_1 - d$. Since the two circular concentrations should be the same, then $\hat{\lambda}^* = \hat{\lambda}$ and, again, the equivalence $\hat{\lambda}^* \sin(2\pi/l) = \hat{\lambda}\sin(2\pi/l) - d$ does not hold if $d > 0$. On the other hand the circular mean of any invariant density can be written as $\delta\xi + \delta\mu_1$, where we can think of μ_1 as a "baseline" circular mean, and parameters δ and ξ can account for changes in the reference system in a coherent way without affecting the circular concentration, see Section 3.2.

4. Numerical Examples

Two examples are considered in this section. The first is based on simulated continuous data and the other is a data problem where discrete circular variables were observed. Our intent lies in highlighting the consequences of the lack of ICO and ICID properties on model inference. We use three reference systems: the first fixes North as the zero direction and chooses a clockwise orientation (RS1), in the second reference system (RS2) we move the zero direction to East, while we obtain the third reference system (RS3) by changing the orientation of RS1. In each reference system, we find the MLEs of the invariant distribution parameters, for the IWE modeling continuous data, the IWP for describing discrete data, and the MLE of the corresponding non-invariant distributions parameters, together with their circular means and concentrations. We indicate with $\hat{\lambda}_i$, $\hat{\mu}_{1,i}$ and $\hat{c}_{1,i}$ the MLEs of the non-invariant distribution parameter, circular mean, and concentration in the i^{th} reference system, respectively, while $\hat{\lambda}^*_i$, $\hat{\delta}^*_i$, $\hat{\xi}^*_i$, $\hat{\mu}^*_{1,i}$ and $\hat{c}^*_{1,i}$ are the corresponding parameters in the i^{th} reference system. The MLEs are found using numerical optimization procedures. An exact evaluation of the IWP density is not possible since it involves the evaluation of an infinite sum (see (3.9)) and a truncation strategy, often used with wrapped distributions (Coles and Casson (1998); Jona Lasinio, Gelfand and Jona Lasinio (2012)), is adopted. We approximate (3.9) with

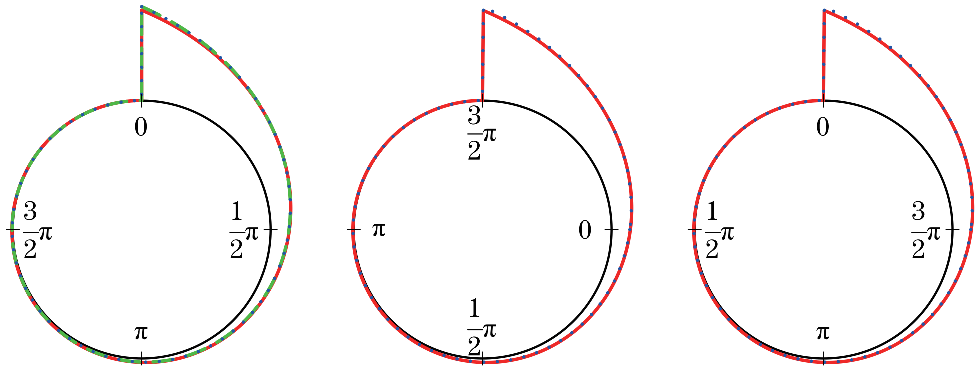
$$f_{\Theta^*}(\theta^*|\lambda,\delta,\xi) = \sum_{k=0}^{k_{\max}} \frac{\lambda^{[(\delta\theta^*-\xi) \bmod (2\pi)]l/2\pi+kl} e^{-\lambda}}{[\{(\delta\theta^*-\xi) \bmod (2\pi)\}\, l/2\pi + kl]!},\ \theta^* \in \mathbb{D},$$

and we choose $k_{\max}$ so that the total probability mass captured by the approximation is > 0.99999.

Artificial data - Wrapped exponential We simulated 500 observations from a $WE(1)$ in the RS1. The MLEs of the WE and IWE parameters are shown in Table 1, while Figure 2 illustrates the density used to simulate the data and the WE and IWE densities obtained with the MLEs in the three reference systems.

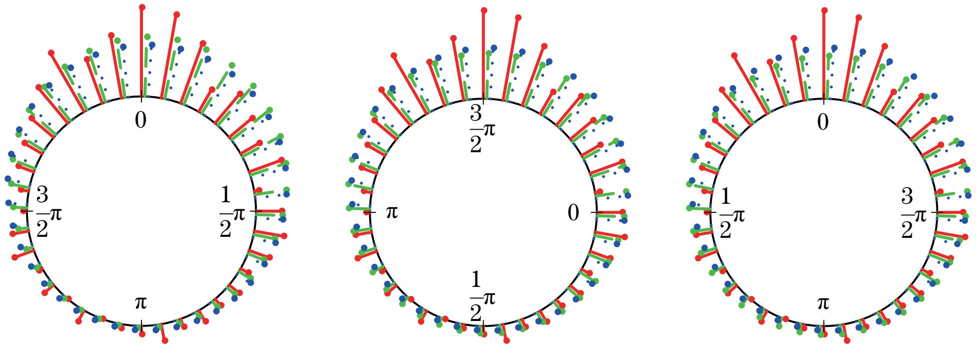
Table 1. Simulated example - MLE of the WE and IWE parameters, circular mean and concentration.

	WE			IWE				
	$\hat{\lambda}_i$	$\hat{\mu}_{1,i}$	$\hat{c}_{1,i}$	$\hat{\lambda}_i^*$	$\hat{\delta}_i^*$	$\hat{\xi}_i^*$	$\hat{\mu}_{1,i}^*$	$\hat{c}_{1,i}^*$
RS1	1.04	0.7658	0.7208	1.04	1	0	0.7658	0.7208
RS2	·	·	·	1.04	1	4.7124	5.4782	0.7208
RS3	·	·	·	1.04	-1	0	5.5174	0.7208



(a) RS1 (b) RS2 (c) RS3

Figure 2. Simulated example - Density used to simulate the data (solid line), wrapped exponential (dashed line) and invariant wrapped exponential (dotted line) density computed using the MLE of the parameters in the three reference systems (the WE MLE does not exist in RS2 and RS3).



(a) RS1 (b) RS2 (c) RS3

Figure 3. Wind example - Density estimate of observed data (solid line), wrapped Poisson (dashed line) and invariant wrapped Poisson (dotted line) density computed using the MLE of the parameters in the three reference systems.

Table 2. Wind example - MLE of the WP and IWP parameters, circular mean and circular concentration.

	WP			IWP				
	$\hat{\lambda}_i$	$\hat{\mu}_{1,i}$	$\hat{c}_{1,i}$	$\hat{\lambda}^*_i$	$\hat{\delta}^*_i$	$\hat{\xi}^*_i$	$\hat{\mu}^*_{1,i}$	$\hat{c}^*_{1,i}$
RS1	37.3434	0.2014	0.567	50.1254	1	4.0143	0.1520	0.4657
RS2	64.1283	4.8526	0.3775	50.1254	1	2.4435	4.8644	0.4657
RS3	71.3427	12.3885	0.3383	50.1254	-1	4.0143	6.1312	0.4657

Data - Wind direction The wind direction data were recorded in January 2000 at the monitoring station of Capo Palinuro (South Italy). The monitoring station at Capo Palinuro (WMO code 16310) is one of the coastal stations managed by the Meteorological Service of the Italian Air Force. The station is located on the rocky cape of Capo Palinuro, in the town of Centola in the province of Salerno, South Italy. Wind directions are monitored and routinely collected by several environmental agencies. Analyzed data come from reports prepared at the station and provided by the National Center of Aeronautical Meteorology and Climatology (C.N.M.C.A.), special office of the Meteorological Service of the Italian Air Force. The database includes date and time of registration, direction of the wind in degrees, with eight daily measurements (every three hours) in the month of January 2000, and contains 240 observations. The measuring instrument, anemometer, is placed away from obstacles and at the height of 10 meters above ground. A relevant issue with this measurement instrument is that it measures wind directions on a discrete scale dividing the circle into ten-degrees intervals ($l = 36$). The MLEs are reported in Table 2 while the barplot of the observed data and the WP and IWP densities obtained with the MLEs in the three reference systems, are shown in Figure 3.

General Comments The arithmetic mean $\sum \theta_i/n$ in the first example is 0.9523 in the RS1, 4.4710 in the RS2, and 5.3308 in the RS3. Keeping in mind the results of Section 3.3, it is not surprising that the MLE of the WE parameter can be estimated only in the RS1.

In all the examples, see Table 1 and 2, we can appreciate how the MLEs of the invariant densities parameters are coherent in moving among the three reference systems: $\hat{\lambda}^*_1 = \hat{\lambda}^*_2 = \hat{\lambda}^*_3$. We have $(\hat{\delta}^*_2, \hat{\xi}^*_2) = \left(\hat{\delta}^*_1, \hat{\xi}^*_1 - \pi/2\right)$, since in the RS2 we change the zero direction to $\pi/2$, and in the RS3 we have $(\hat{\delta}^*_3, \hat{\xi}^*_3) = \left(-\hat{\delta}^*_1, \hat{\xi}^*_1\right)$ as we changed the orientation. The circular concentration remains the same in the three reference systems, $\hat{c}^*_{1,1} = \hat{c}^*_{1,2} = \hat{c}^*_{1,3}$, while the circular mean changes according to the reference system, $\hat{\mu}^*_{1,2} = \hat{\mu}^*_{1,1} - \pi/2$ and $\hat{\mu}^*_{1,3} = -\hat{\mu}^*_{1,2}$. For the non-invariant densities, even the circular concentration changes with the reference system as well as the MLE of the parameter λ and the circular mean. The shapes of the invariant densities remain the same in the three reference systems, see Figures 2 and 3.

5. Summary and Concluding Remarks

In this paper we introduced the invariance under changes of initial direction and the invariance under changes of the reference system orientation properties that any circular distribution should hold to avoid misleading inferential results. We introduced the necessary and sufficient conditions for a circular distribution to have the ICID and ICO properties. By considering simulated and data examples, we illustrated how misleading the use of not-invariant distributions can be. Not all the circular distributions proposed in the literature satisfy ICID and ICO and we developed the method to modify them appropriately. Trigonometric moments are then easily obtained from those of the non-invariant version of circular distributions (see Supplementary Material for more details and examples).

Further use of the invariant distributions in complex models can be found in Mastrantonio and Calise (2016), where this is extended, in a Bayesian framework, to address classification issues with discrete circular and linear variables; in that paper a hidden Markov model for discrete valued time series with linear and circular components is introduced. We believe that our proposal opens new possibilities for the practical use of circular information measured on any scale.

Supplementary Materials

The online supplementary material contains more analytic applications and two numerical examples. In particular we verify if the wrapped skew normal, discrete circular uniform, wrapped Poisson, wrapped Weibull, wrapped geometric, wrapped skew Laplace on integers and wrapped binomial distributions are ICI and ICO. We also report on artificial wrapped Poisson data, and on the Drosophila change of direction data from Suster (2000).

Acknowledgment

This work is partially developed under the PRIN2015 supported-project "Environmental processes and human activities: capturing their interactions via statistical methods (EPHASTAT)" funded by MIUR (Italian Ministry of Education,

University and Scientific Research).

References

Bulla, J., Lagona, F., Maruotti, A. and Picone, M. (2012). A multivariate hidden Markov model for the identification of sea regimes from incomplete skewed and circular time series. *Journal of Agricultural, Biological, and Environmental Statistics* **17**, 544–567.

Coles, S. and Casson, E. (1998). Extreme value modelling of hurricane wind speeds. *Structural Safety* **20**, 283–296.

Di Marzio, M., Panzera, A. and Taylor, C. C. (2013). Non-parametric regression for circular responses. *Scandinavian Journal of Statistics* **40**, 238–255.

Eckert, S. A., Moore, J. E., Dunn, D. C., van Buiten, R. S., Eckert, K. L. and Halpin, P. N. (2008). Modeling loggerhead turtle movement in the mediterranean: importance of body size and oceanography. *Ecological Applications* **18**, 290–308.

Fisher, N. I. (1996). *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press.

Gill, J. and Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis* **18**, 316–336.

Jammalamadaka, S. R. and Kozubowski, T. J. (2004). New families of wrapped distributions for modeling skew circular data. *Communications in Statistics - Theory and Methods* **33**, 2059–2074.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. Singapore: World Scientific.

Jona Lasinio, G., Gelfand, A. and Jona Lasinio, M. (2012). Spatial analysis of wave direction data using wrapped Gaussian processes. *Annals of Applied Statistics* **6**, 1478–1498.

Lagona, F., Picone, M. and Maruotti, A. (2015). A hidden Markov model for the analysis of cylindrical time series. *Environmetrics* **26**, 534–544.

Lagona, F., Picone, M., Maruotti, A. and Cosoli, S. (2015). A hidden Markov approach to the analysis of space–time environmental data with linear and circular components. *Stochastic Environmental Research and Risk Assessment* **29**, 397–409.

Langrock, R., Hopcraft, G., Blackwell, P., Goodall, V., King, R., Niu, M., Patterson, T., Pedersen, M., Skarin, A. and Schick, R. (2014). Modelling group dynamic animal movement. *Methods in Ecology and Evolution* **5**, 190–199.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. and Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology* **93**, 2336–2342.

Lee, A. (2010). Circular data. *Wiley Interdisciplinary Reviews: Computational Statistics* **2**, 477–486.

Lee, J.-R. and Oh, H.-S. (2008). Circular statistics in musicology. *Communications of the Korean Statistical Society* **15**, 273–282.

Mardia, K. V. (1972). *Statistics of Directional Data*. London: Academic Press.

Mardia, K. V. and Jupp, P. E. (1999). *Directional Statistics*. Chichester: John Wiley and Sons.

Maruotti, A. (2016). Analyzing longitudinal circular data by projected normal models: a semi-parametric approach based on finite mixture models. *Environmental and Ecological Statis-*

tics **23**, 257–277.

Maruotti, A., Punzo, A., Mastrantonio, G. and Lagona, F. (2016). A time-dependent extension of the projected normal regression model for longitudinal circular data based on a hidden Markov heterogeneity structure. *Stochastic Environmental Research and Risk Assessment* **30**, 1725–1740.

Mastrantonio, G. and Calise, G. (2016). Hidden Markov model for discrete circular–linear wind data time series. *Journal of Statistical Computation and Simulation* **86**, 2611–2624.

Mastrantonio, G., Gelfand, A. E. and Jona Lasinio, G. (2016). The wrapped skew Gaussian process for analyzing spatio-temporal data. *Stochastic Environmental Research and Risk Assessment* **20**, 2231–2242.

Mastrantonio, G., Jona Lasinio, G. and Gelfand, A. E. (2016). Spatio-temporal circular models with non-separable covariance structure. *TEST* **25**, 331–350.

Mastrantonio, G., Maruotti, A. and Jona Lasinio, G. (2015). Bayesian hidden Markov modelling using circular-linear general projected normal distribution. *Environmetrics* **26**, 145–158.

McLellan, C. R., Worton, B. J., Deasy, W. and Birch, A. N. E. (2015). Modelling larval movement data from individual bioassays. *Biometrical Journal* **57**, 485–501.

Pewsey, A. (2000). The wrapped skew-normal distribution on the circle. *Communications in Statistics - Theory and Methods* **29**, 2459–2472.

Pewsey, A., Neuhäuser, M. and Ruxton, G. D. (2013). *Circular Statistics in R.* Croydon: Oxford University Press.

Suster, M. (2000). *Neural control of larval locomotion in Drosophila melanogaster*. Ph.D. thesis, University of Cambridge.

Deemer, W. L. and Votaw, D. F. (1955). Estimation of parameters of truncated or censored exponential distributions. *The Annals of Mathematical Statistics* **26**, 498–504.

Wang, F. and Gelfand, A. E. (2014). Modeling space and space-time directional data using projected Gaussian processes. *Journal of the American Statistical Association* **109**, 1565–1580.

Department of Mathematical Science, Polytechnic of Turin, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy.

E-mail: gianluca.mastrantonio@polito.it

Department of Statistical Sciences, University of Rome “Sapienza”, Piazzale Aldo Moro, 5, 00185 Roma, Italy.

E-mail: giovanna.jonalasinio@uniroma1.it

Centre for Innovation and Leadership in Health Sciences, University of Southampton, Building 67 University Road, Southampton, UK.
Department of Economic, Political Sciences and Modern Languages, LUMSA, Via Pompeo Magno 22, 00192 Roma, Italy.

E-mail: a.maruotti@lumsa.it

Department of Earth Science, University of Rome “Sapienza”, Piazzale Aldo Moro, 5, 00185 Roma, italy.

E-mail: gianfranco.calise@uniroma1.it

(Received February 2016; accepted May 2017)