Wiley StatsRef:
Statistics Reference Online

Directional Statistics: Introduction

By *Riccardo Gatto*[a] *and S. Rao Jammalamadaka*[b]

Keywords: *Directional distributions, goodness-of-fit tests, nonparametric inference, robust estimators, rotational invariance, spacings, spacing-frequencies, two-sample tests, von Mises–Fisher distribution*

Abstract: In many of the natural and physical sciences, measurements are directions, either in two or three dimensions. The analysis of directional data relies on specific statistical models and procedures, which differ from the usual models and methodologies of Cartesian data. This article briefly introduces statistical models and inference for this type of data. The basic von Mises–Fisher distribution is introduced and nonparametric methods such as goodness-of-fit tests are presented. Further references are given for exploring related topics.

1 Introduction

This article explores the novel area of directional statistics. In many diverse scientific fields, measurements are directions. For instance, a biologist may be measuring the direction of flight of a bird or the orientation of an animal, whereas a geologist may be interested in the direction of the Earth's magnetic pole. Such directions may be in two dimensions as in the first two examples or in three dimensions like the last one. A set of such observations on directions is referred to as *directional data* (*see* **Directional Data Analysis**). Langevin[1] and Lévy[2] are two pioneering contributions in this area. One of the first statistical analyses of directional data is Fisher[3].

A two-dimensional direction is a point in $\mathbb{R}^2$ without magnitude, for example, a unit vector. It can also be represented as a point on the circumference of the unit circle centered at the origin, denoted S^1, or as an angle measured with respect to some suitably chosen "zero direction," that is, starting point, and "sense of rotation," that is, whether clockwise or counterclockwise is the positive direction. Because of this circular representation, observations on two-dimensional directions are also called *circular data*. Generally, directions in $p \geq 2$ dimensions can be represented by $p-1$ angles (akin to the representation of points on the Earth's surface by their longitude and latitude when $p = 3$), as unit vectors in $\mathbb{R}^p$ or as points on the unit hypersphere $S^{p-1} = \{\mathbf{x} \in \mathbb{R}^p | \langle \mathbf{x}, \mathbf{x} \rangle = 1\}$. Consequently, directional data in p dimensions are also referred to as *spherical data*.

Directional data have many unique and novel features both in terms of modeling and in their statistical treatment. For instance, the numerical representation of a two-dimensional direction as an angle or a unit

[a]University of Bern, Bern, Switzerland
[b]University of California, Santa Barbara, CA, USA

Update based on original article by S. Rao Jammalamadaka, Wiley StatsRef: Statistics Reference Online,

DOI: 10.1002/9781118445112.stat00201.pub2

vector is not necessarily unique, as the angular value depends on the choice of the zero direction and the sense of rotation. It is therefore important to determine statistical procedures (e.g., data summaries and inference), which are "invariant" with respect to any arbitrary choice for the zero-direction and sense of rotation. Thus, there is no natural ordering or ranking of observations, as whether one direction is "larger" than the other depends on whether clockwise or counterclockwise is treated as being the positive direction as well as where the origin is. This makes rank-based methods essentially inapplicable. Finally, as the "beginning" coincides with the "end" that is, the data is periodic, methods for dealing with directional data should take special care on how to measure the distance between any two points.

Such distinctive features make directional analysis substantially different from the standard "linear" statistical analysis of univariate or multivariate data. The need for the described invariance of statistical methods and measures makes many of the usual linear techniques and measures misleading, if not entirely meaningless. Commonly used summary measures on the real line, such as the sample mean and variance, turn out to be inappropriate as do moments and cumulants. Analytical tools such as the moment generating function (*see* **Moment Generating Function**) and other generating functions are also essentially useless. Many notions such as correlation and regression as well as their statistical measures need to be reinvented for directional data. Similarly, ideas of statistical inference such as unbiasedness, loss functions, variance bounds need to be redefined with caution.

All in all, this area of directional data provides an inquisitive reader with many open research problems and is a fertile area for developing new statistical methods and inferential tools. There is also an opportunity to develop new and novel applications to problems arising in the natural, physical, medical, and social sciences.

One can also note that in studying circadian or other rhythms, the circle may be used to represent one cycle, and the interest may lie in the timing of an event within this cycle, say, for instance, when the body temperature or the blood pressure peaks within the day. Thus, certain aspects in the study of biological rhythms provide an important example from biology that can be put into this circular data framework. Because biological rhythms control characteristics such as sleep-wake cycles, hormonal pulsatility, body temperature, mental alertness, and reproductive cycles, there has been a renewed interest among medical professionals in topics such as chronobiology, chronotherapy, chronomedicine, and the study of the biological clock.

This succinct presentation of directional statistics makes readers aware of the limitations of usual linear statistical methodology. A software for circular data analysis based on *R*, called *CircStats*, is freely available. Books on circular data include those by Batschelet[4], Fisher[3], Mardia and Jupp[5], and Jammalamadaka and SenGupta[6]. Readers interested in directions in three dimensions can consult the books by Watson[7] and Fisher *et al.*[8].

2 Models for Directional Data

Probability distributions for directions in $\mathbb{R}^p$ are presented in this section. A particular attention is given to the case $p = 2$ in Section 2.1 and the case $p > 2$ is overviewed in Section 2.2.

2.1 Circular Distributions

A direction in the plane can be represented in the following equivalent ways: as an angle θ (in radians for example), as a point of S^1, as a unit vector $(\cos\theta, \sin\theta)$ of $\mathbb{R}^2$, or as a complex number of unit modulus, $\mathrm{e}^{\mathrm{i}\theta}$. We follow the first description and define a circular density as any nonnegative and 2π-periodic real function integrating to one over any interval of length 2π.

DOI: 10.1002/9781118445112.stat00201.pub2

One can distinguish four main types of circular distribution: (i) wrapped distributions, (ii) offset distributions, (iii) distributions derived from stereographic projections and (iv) distributions resulting from theoretical characterizations.

(i) Concerning wrapped distributions, let Z be a $\mathbb{R}$-valued random variable with density g, then $X = Z \mod 2\pi$ corresponds to "wrapping" Z around S^1. Thus, X has density

$$f(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi) \tag{1}$$

$\forall\theta \in \mathbb{R}$. From periodicity, we can represent any circular density f in terms of its discrete Fourier transform

$$f(\theta) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \varphi_j \exp\ \{-\mathrm{i}j\theta\}$$

$\forall\theta \in \mathbb{R}$, the equality being in the $\mathscr{L}^2$-sense, where φ_j is the j-th Fourier coefficient, $\forall j \in \mathbb{Z}$. If f is the wrapped density (1), then the Fourier coefficients are the characteristic function (*see* **Characteristic Functions**) of the linear random variable Z at integer values, that is, $\varphi_j = \mathrm{E}[\exp\{\mathrm{i}jZ\}], \forall j \in \mathbb{Z}$. Gatto and Jammalamadaka[9] provide methods of inference for wrapped α-stable distributions, which derive from the important α-stable distributions on $\mathbb{R}$. Two well-known members within this class are the wrapped normal and Cauchy distributions.

(ii) A circular distribution can also be obtained from the directional component of a distribution over $\mathbb{R}^2$. Assume that g is a probability density over $\mathbb{R}^2$, then the marginal density given by

$$f(\theta) = \int_0^{\infty} g(r\cos\theta,\ r\sin\theta) r\ \mathrm{d}r$$

$\forall\theta \in \mathbb{R}$, is called an *offset* or *projected* density, as it corresponds to the radial projection from $\mathbb{R}^2\backslash\{0\}$ onto S^1.

(iii) A third main way of determining circular distributions is to apply a stereographic projection to a random variable on the real line. This is done by taking X as the point of intersection between the unit circle centered around $(0, 1)$ and the straight line connecting the top of the circle with the random point Z on the real line. Note that $\pm\infty$ on the real line are mapped to $X = \pi$ on this circle.

(iv) Circular distributions can also be deduced from important theoretical properties or characterizations. The well-known von Mises (vM) distribution, also called the *circular normal distribution*, possesses many important properties. It has density

$$f(\theta \mid \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\ \{\kappa\cos(\theta - \mu)\} \tag{2}$$

$\forall\theta \in \mathbb{R}, \mu \in [0, 2\pi), \kappa \geq 0$ and where $I_r(z) = (2\pi)^{-1} \int_0^{2\pi} \cos r\theta\ \exp\{z\cos\theta\} d\theta, \forall z \in \mathbb{C}$ (which is the modified Bessel function of integer order r, see for example, Abramowitz and Stegun[10], p. 376).

The generalized von Mises (GvM) distribution has density

$$f(\theta \mid \mu_1, \mu_2, \kappa_1, \kappa_2) = \frac{1}{2\pi G_0(\delta, \kappa_1, \kappa_2)} \exp\{\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos 2(\theta - \mu_2)\} \tag{3}$$

$\forall\theta \in \mathbb{R}, \mu_1 \in [0, 2\pi), \mu_2 \in [0, \pi), \delta = (\mu_1 - \mu_2) \bmod \pi, \kappa_1, \kappa_2 \geq 0$ and where the normalizing constant is given by $G_0(\delta, \kappa_1, \kappa_2) = (2\pi)^{-1} \int_0^{2\pi} \exp\ \{\kappa_1\cos\theta + \kappa_2\cos 2(\theta + \delta)\}\ \mathrm{d}\theta$, see Gatto and Jammalamadaka[11], who provide the following characterizations.

- The GvM distribution can be re-expressed in the canonical form of the exponential class.
- The GvM distribution maximizes the entropy among all distributions with fixed first and second trigonometric moments. Shannon's entropy (*see* **Entropy**)

$$-\int_0^{2\pi} \log\ f(\theta) f(\theta)\ \mathrm{d}\theta$$

DOI: 10.1002/9781118445112.stat00201.pub2

is an appropriate measure of the uncertainty carried by the circular distribution with density f. This characterization is relevant in Bayesian statistics (*see* **Bayesian Methods**; **Bayesian Model Selection**), where partial prior information is available and the most noninformative prior distribution that satisfies the known partial prior information is desired. An optimal prior is the solution of the constrained maximization of the entropy.

- The GvM distribution is the conditional offset distribution of the bivariate normal (*see* **Bivariate Normal Distribution**). A conditional offset distribution is the conditional distribution of the directional component given a fixed length from the origin.

The vM distribution (2) is clearly a special case of the GvM distribution (3) and therefore it shares these characterizations. The GvM density (3) provides a very flexible model that allows for bimodality and for various asymmetric shapes. In fact, the GvM density can have more than two cosines in the exponent, see Gatto and Jammalamadaka[11]. Various other types of models can be found in the literature. However, their use for data analysis should be judged also by the availability of relevant sampling theory for inference.

2.2 Directional Distributions

Directions in $\mathbb{R}^p$ can be represented by the unit vector $\mathbf{x} \in S^{p-1}$, which can be re-expressed by angular coordinates by the transform $\mathbf{x} = (x_1, \ldots, x_p) = \mathbf{g}(\theta_1, \ldots, \theta_{p-1})$ which, for $p \geq 3$, is given by

$$\begin{aligned} x_1 &= \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{p-2}\cos\theta_{p-1} \\ x_2 &= \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{p-2}\sin\theta_{p-1} \\ \vdots \quad & \vdots \\ x_{p-1} &= r\sin\theta_1\sin\theta_2 \\ x_p &= \cos\theta_1 \end{aligned}$$

where $0 \leq \theta_i \leq \pi$, for $i = 1, \ldots, p-2$, and $0 \leq \theta_{p-1} < 2\pi$. For $p = 3, \theta_1$ is the co-latitude and θ_2 is the longitude.

A popular model for directional data is the von Mises–Fisher (vMF), which is due to Langevin[1] and which was rediscovered by von Mises[12]. The density toward any p-dimensional direction $\mathbf{x}$ is given by

$$f(\mathbf{x}|\boldsymbol{\mu}, \kappa) = C_p(\kappa)\exp\{\kappa\langle \mathbf{x}, \boldsymbol{\mu}\rangle\} \tag{4}$$

where $\boldsymbol{\mu} \in S^{p-1}$ denotes the mean direction. It can be checked that the normalizing constant $C_p(\kappa)$ is given by

$$C_p(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)}$$

If the density is given with respect to the uniform distribution on S^{p-1}, then the normalizing constant becomes $C_p(\kappa)a_p$, where $a_p = 2\pi^{p/2}/\Gamma(p/2)$ is the surface area of S^{p-1}, Γ denoting the gamma function. The vM distribution corresponds to the case $p = 2$ and the sampling properties of the case $p = 3$ were discussed by Fisher[13].

The sampling distribution of the resultant vector $\sum_{i=1}^{n} \mathbf{x}_i$, which is a complete and sufficient statistic, can be easily obtained and so inference on $\boldsymbol{\mu}$ and κ can be relatively easily handled. The vMF distribution possesses essentially the same theoretical properties of the vM distribution, including an analogous information theoretic characterization. We can also mention the rotational invariance: if $\mathbf{A} \in \mathbb{R}^{p\times p}$ is an orthogonal matrix, then $f(\mathbf{A}\mathbf{x}|\mathbf{A}\boldsymbol{\mu}, \kappa) = f(\mathbf{x}|\boldsymbol{\mu}, \kappa), \forall \mathbf{x}, \boldsymbol{\mu} \in S^{p-1}$ and $\kappa \geq 0$. Finally, an interesting probabilistic characterization is that the first exit point from S^{p-1} of the drifted Wiener process on $\mathbb{R}^p$ starting at the origin is vMF distributed, see Gatto[14] and references therein.

DOI: 10.1002/9781118445112.stat00201.pub2

3 Nonparametric Tests for Circular Data

The vM distribution is often considered as central as the normal distribution for linear data, even though there is not an analogous asymptotic rationale for invoking it. In this context, nonparametric or distribution-free inference (*see* **Distribution-Free Methods**) becomes very important with directional data. Section 3.1 presents goodness-of-fit tests and Section 3.2 presents two-sample tests of equality of distributions.

3.1 One-Sample Goodness-Of-Fit Tests

Let $X_1, \ldots, X_n$ be independent circular random variables with common circular probability distribution P_X. The goodness-of-fit problem involves testing that the sample arises from a specified directional distribution Q, that is,

$$\mathrm{H}_0 : P_X = Q$$

One can distinguish three main categories of goodness-of-fit tests: (i) those based on grouping data, (ii) those based on the empirical distribution function (e.d.f.) and (iii) those based on the gaps between successive points, called *spacings*. As mentioned in the introduction, there is generally no prescribed null direction or sense of rotation. These attributes are arbitrarily fixed and invariant inference with respect to these arbitrary choices is desired. Let us denote by $\mathcal{G}$ the transformation group $[0, 2\pi)^n \to [0, 2\pi)^n$ consisting of all changes of origin and of the two changes of sense of rotation. The test statistic $T : [0, 2\pi)^n \to \mathbb{R}$ is $\mathcal{G}$-invariant, if $T(X_1, \ldots, X_n) = T(\boldsymbol{g}(X_1, \ldots, X_n)), \forall \boldsymbol{g} \in \mathcal{G}$.

(i) Tests based on grouping data such as the chi-square and the likelihood ratio tests, depend on circular partitions of $[0, 2\pi)$ into cells, which in turn depend on the choice of the null direction and of the sense of rotation. Consequently, they are not directly applicable to circular data. Chi-square tests that are $\mathcal{G}$-invariant are considered in Ajne[15] and Rao[16]. Rao[16] considers the average of χ^2 over all possible choices of the zero direction, to obtain an invariant version of it.

(ii) For the same invariance consideration, tests based on the e.d.f., such as the Kolmogorov–Smirnov or the Cramér–von Mises tests (*see* **Cramér-von Mises Statistic**), are not directly applicable to circular data. Let us fix an arbitrary origin and sense of rotation. Then, we can define the e.d.f. of this circular sample as in the linear case, that is, by $\hat{F}_n(\theta) = n^{-1} \sum_{i=1}^{n} \mathrm{I}\{X_i \leq \theta\}, \forall \theta \in \mathbb{R}$, where $\mathrm{I}\{A\}$ denotes the indicator of statement A. This e.d.f. clearly depends on the choice of this origin as well as whether clockwise or counterclockwise is taken as the positive direction. The Kolmogorov–Smirnov and Cramér–von Mises test statistics were modified so as to make them $\mathcal{G}$-invariant by Rao[16], Kuiper[17], and Watson[18].

The one-sided Kolmogorov–Smirnov test statistics are given by

$$D_n^+ = \sqrt{n} \sup_{\theta \in [0, 2\pi)} \{\hat{F}_n(\theta) - F(\theta)\}$$

and

$$D_n^- = \sqrt{n} \sup_{\theta \in [0, 2\pi)} \{F(\theta) - \hat{F}_n(\theta)\}$$

where F denotes the circular distribution function of P_X. The two-sided Kolmogorov–Smirnov statistic is then given by

$$\sqrt{n} \sup_{\theta \in [0, 2\pi)} |\hat{F}_n(\theta) - F(\theta)| = \max\{D_n^+, D_n^-\}$$

By noticing that D_n^+ gains (or loses) just as much as D_n^- loses (or gains) due to a rotation, Kuiper[17] suggested the statistic

$$D_n^+ + D_n^-$$

DOI: 10.1002/9781118445112.stat00201.pub2

A classical alternative e.d.f.-based goodness-of-fit test on the real line is provided by the Cramér–von Mises test, given by

$$n \int_{-\infty}^{\infty} (\hat{F}_n - F)^2 \, \mathrm{d}F$$

Watson[18] provided the $\mathcal{G}$-invariant modification defined by

$$\int_0^{2\pi} \left\{ (\hat{F}_n - F) - \int_0^{2\pi} (\hat{F}_n - F) \, \mathrm{d}F \right\}^2 \mathrm{d}F$$

We can note that the Cramér–von Mises statistic can be thought of as the second moment of $\hat{F}_n - F$ and Watson's statistic is similar to the expression for the variance, so that a variation of $\hat{F}_n - F$ due to a change of origin has no influence on it.

(iii) The third category of tests is the one based on spacings. Ordering the sample with respect to a given zero direction and sense of rotation yields $0 \leq X_{(1)} \leq \dots \leq X_{(n)} < 2\pi$. The spacings are the n random variables defined by

$$D_i = X_{(i)} - X_{(i-1)}, \quad \text{for } \; i = 1, \dots, n, \quad \text{where } \; X_{(0)} = X_{(n)} - 2\pi$$

Thus, D_1 is the gap between the first and last ordered values that straddle the origin. The spacings form a maximal $\mathcal{G}$-invariant, so that every $\mathcal{G}$-invariant statistic can be expressed in terms of spacings. The one-sample goodness-of-fit problem can be converted into one of testing uniformity, just as one does on the real line, as long as the proposed test statistic is $\mathcal{G}$-invariant. We are thus lead to test uniformity, under which the spacings become n exchangeable random variables with mean n^{-1}.

Rao[19] defines the circular range by $2\pi - \max_{i=1,\dots,n} D_i$, which corresponds to the shortest arc-length that contains the sample. Small values of it indicate clustering and lead to rejection of uniformity.

Generally, one considers spacing statistics of the form

$$\sum_{i=1}^{n} h(nD_i)$$

for a real-valued function h satisfying some regularity conditions. Typical examples are $h(x) = x^r$, for $r > 0$ and $r \neq 1$, where the case $r = 2$ is called *Greenwood Statistic*, $h(x) = \log\ x$ and Rao's test, $h(x) = |x - 1|/2$ (*see* **Circular Data, Rao's Spacing Test for**).

3.2 Two-Sample Tests

Two circular samples can be compared using two-sample versions of Kuiper's and Watson's tests, among others, or using tests based on the "spacing-frequencies" defined below. Let $X_1, \dots, X_m$ be independent with common circular probability distribution P_X and let $Y_1, \dots, Y_n$ be independent with common circular probability distribution P_Y. The two samples are independent. The null hypothesis is that the two samples arise from the same model, that is

$$\mathrm{H}_0 \; : P_X = P_Y$$

Let $X_{(1)} \leq \dots \leq X_{(m)}$ denote the circularly ordered values $X_1, \dots, X_m$, for a given origin and sense of rotation. The spacing-frequencies are given by

$$S_j = \sum_{i=1}^{n} \mathrm{I}\{Y_i \in [X_{(j)}, X_{(j+1)})\} \;\; \text{for} \;\; j = 1, \dots, m-1, \quad \text{and } S_m = n - \sum_{j=1}^{m-1} S_j$$

They count the number of Y-values that lie in-between successive gaps made by $X_{(1)}, \dots, X_{(m)}$.

DOI: 10.1002/9781118445112.stat00201.pub2

Let us now denote by $\mathcal{G}$ the transformation group $[0, 2\pi)^{m+n} \to [0, 2\pi)^{m+n}$ consisting of all changes of origin and senses of rotation, now for the two samples. The two-sample test statistic $T : [0, 2\pi)^{m+n} \to \mathbb{R}$ is $\mathcal{G}$-invariant, if $T(X_1, \dots, X_m, Y_1, \dots, Y_n) = T(\boldsymbol{g}(X_1, \dots, X_m, Y_1, \dots, Y_n)), \forall \boldsymbol{g} \in \mathcal{G}$. Obviously, $(S_1, \dots, S_m)$ is not $\mathcal{G}$-invariant: if we change, for example, the zero direction, then the new spacing-frequencies are obtained by a permutation of the original ones. If however T is a symmetric function of $S_1, \dots, S_m$, then T is $\mathcal{G}$-invariant. Thus, the popular nonparametric Wilcoxon test statistic, which takes the nonsymmetric form $\sum_{k=1}^{m} kS_k$, is inadequate for circular data.

Holst and Rao[20] consider the class of test statistics

$$\sum_{j=1}^{m} h(S_j)$$

where $h : \{0, \dots, n\} \to \mathbb{R}$ satisfies certain regularity conditions. Two important examples are the Dixon test, with $h(x) = x^2$, and the Wald–Wolfowitz run test, with $h(x) = \mathrm{I}\{x > 0\}$.

Note that reverting the roles of the X- and Y-samples in the construction of the spacing-frequencies yields a set of dual spacing-frequencies. However, because these dual values can be obtained as a one-to-one function from the original spacing-frequencies and conversely, the tests can be based on either set of spacing-frequencies.

4 Other Topics

There are several other important topics of directional statistics and we briefly mention two of them.

There are several important higher dimensional extensions of directional models. One can consider two or more directional variables simultaneously, or directional variables in conjunction with linear variables. For instance, models for bivariate circular data have the torus $\mathrm{S}^1 \times \mathrm{S}^1$ as their sample space, whereas one circular variable and one linear variable take their values on the surface of a cylinder $\mathrm{S}^1 \times \mathbb{R}$ (*see* **Cylindrical Data**). Questions of modeling, correlation, and regression (*see* **Spherical Regression**) in such a context are discussed in Fisher[3], Chap. 6, Mardia and Jupp[5], Chap. 11, and Jammalamadaka and SenGupta[6], Chap. 8.

Next, robust statistical procedures (*see* **Robust Estimation**) are also important with directional data. The influence function has a central role in the analysis of robustness. It is the ratio of the infinitesimal bias of a statistic resulting from an infinitesimal contamination of the model under consideration, over the amount of infinitesimal contamination (*see* **Influence Functions**). Because S^{p-1} is compact, the mean and the median directions appear quite robust, see for example, Wehrli and Shine[21]. However, one can object that a small value of the influence function should not neglected when it is substantially larger than the standard deviation of the estimator. With this in mind, Ko and Guttorp[22] suggested the use of standardized influence functions and showed that they can be very large when computed for the mean direction under various directional distribution.

The spherical median of $\mathbf{X}_1, \dots, \mathbf{X}_n$, S^{p-1}-valued, can be defined as the value $\boldsymbol{t} \in \mathrm{S}^{p-1}$ minimizing the sum of arc-lengths

$$\sum_{i=1}^{n} \arccos\langle \mathbf{X}_i, \boldsymbol{t} \rangle$$

Generalizing on this form, the M-estimator of the mean direction (*see* **M Estimators of Location**) is defined by $\boldsymbol{t} \in \mathrm{S}^{p-1}$ minimizing

$$\sum_{i=1}^{n} \rho(\langle \mathbf{X}_i, \boldsymbol{t} \rangle)$$

for some appropriate function ρ, see for example, Lenth[23] and Ko and Chang[24].

DOI: 10.1002/9781118445112.stat00201.pub2

Estimators of the concentration parameter κ of the vMF distribution (4) can be found, for example, in Ducharme and Milasevic[25] and Ronchetti[26]. A more detailed review on robustness can be found in Mardia and Jupp[5], Sect. 12.4.

Acknowledgment

The authors thank the Editor W. W. Piegorsch for various corrections.

Related Articles

Bayesian Model Selection; **Directional Data Analysis**; **Entropy**; **Circular Data, Rao's Spacing Test for**; **Influence Functions**; **Robust Estimation**; **Distribution-Free Methods**; **Bayesian Methods**; **Bivariate Normal Distribution**; **Moment Generating Function**; **M Estimators of Location**; **Cylindrical Data**; **Characteristic Functions**; **Spherical Regression**.

References

[1] Langevin, P. (1905) Magnétisme et théorie des éléctrons. *Ann. Chim. Phys.*, **5**, 71–127.
[2] Lévy, P. (1939) L'addition des variables aléatoires définies sur une circonférence. *Bull. Soc. Math. Fr.*, **67**, 1–41.
[3] Fisher, N.I. (1993) *Statistical Analysis of Circular Data*, Cambridge University Press, Cambridge.
[4] Batschelet, M. (1981) *Circular Statistics in Biology*, Academic Press, London.
[5] Mardia, K.V. and Jupp, P.E. (2000) *Directional Statistics*, John Wiley & Sons, Inc., New York.
[6] Jammalamadaka, S.R. and SenGupta, A. (2001) *Topics in Circular Statistics*, World Scientific, Singapore.
[7] Watson, G.S. (1983) *Statistics on Spheres*, John Wiley & Sons, Inc., New York.
[8] Fisher, N.I., Lewis, T., and Embleton, B.J.J. (1987) *Statistical Analysis of Spherical Data*, Cambridge University Press, Cambridge.
[9] Gatto, R. and Jammalamadaka, S.R. (2003) Inference for wrapped symmetric α-stable circular models. *Sankhya*, **65**, 333–355.
[10] Abramowitz, M. and Stegun, I.E. (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, 9-th Printing, New York; originally published by the National Bureau of Standards, USA, 10-th Printing.
[11] Gatto, R. and Jammalamadaka, S.R. (2007) The generalized von Mises distribution. *Stat. Methodol.*, **4**, 341–353.
[12] von Mises, R. (1918) Über die "Ganzzahligkeit" der Atomgewichhte und verwandte Fragen. *Phys. Z.*, **19**, 490–500.
[13] Fisher, R.A. (1953) Dispersion on a sphere. *Proc. R. Soc. London, Ser. A*, **217**, 295–305.
[14] Gatto, R. (2013) The von Mises-Fisher distribution of the first exit point from the hypersphere of the drifted Brownian motion and the density of the first exit time. *Stat. Probab. Lett.*, **83**, 1669–1676.
[15] Ajne, B. (1968) A simple test for uniformity of a circular distribution. *Biometrika*, **55**, 343–354.
[16] Rao, J.S. (1972) Some variants of chi-square for testing uniformity on the circle. *Z. Wahrscheinlichkeitstheor. Verwandte Geb.*, **22**, 33–44.
[17] Kuiper, N.H. (1960) Tests concerning random points on a circle. *Ned. Akad. Wet. Proc.*, **63**, 38–47.
[18] Watson, G.S. (1961) Goodness-of-fit tests on the circle. *Biometrika*, **48**, 109–114.
[19] Rao, J.S. (1969) *Some contributions to the analysis of circular data*. PhD thesis, Indian Statistical Institute, Calcutta, India.
[20] Holst, L. and Rao, J.S. (1980) Asymptotic theory for families of two-sample nonparametric statistics. *Sankhya, Ser. A*, **42**, 19–52.
[21] Wehrli, T. and Shine, E.P. (1981) Influence curves for directional data. *Biometrika*, **68**, 634–635.
[22] Ko, D. and Guttorp, P. (1988) Robustness of estimators for directional data. *Ann. Stat.*, **16**, 600–618.
[23] Lenth, R.V. (1981) Robust measures of location for directional data. *Technometrics*, **23**, 77–81.
[24] Ko, D. and Chang, T. (1993). Robust M-estimators on spheres. *J. Multivariate Anal.*, **45**, 104–136.
[25] Ducharme, G.R. and Milasevic, P. (1990) Estimating the concentration of the Langevin distribution. *Can. J. Stat.*, **18**, 163–169.
[26] Ronchetti, E. (1992) Optimal robust estimators for the concentration parameter of a Fisher-von Mises distribution, in *The Art of Statistical Science: A Tribute to G. S. Watson* (ed. K.V. Mardia), John Wiley & Sons, Inc., Chichester, pp. 65–74.

DOI: 10.1002/9781118445112.stat00201.pub2