International Statistical Review (1989), **57**, 3, pp. 261–294. Printed in Great Britain

A Unified View of the Theory of Directional Statistics, 1975–1988

P.E. Jupp[1] and K.V. Mardia[2]

[1]Department of Mathematical Sciences, University of St. Andrews, St. Andrews, Fife KY16 9SS, U.K. [2]Department of Statistics, School of Mathematics, University of Leeds, Leeds, LS2 9JT, U.K.

Summary

Numerous articles on the theory and practice of directional statistics have appeared since Mardia's (1975a) survey. This paper aims to present a coherent view of the theory of the topic by relating these developments to the following key ideas: exponential families, transformation models, 'tangent-normal' decompositions, transformations to multivariate problems and the central limit theorem. Further unification is attained by identifying three basic approaches to directional statistics, in which the basic sample space, the sphere, is regarded respectively as a subset of Euclidean space, an object in its own right and as something approximated by a tangent plane. Parametric, nonparametric and informal methods are considered. The discussion is mainly of observations on the circle or sphere but a section on non-spherical sample spaces is included.

Key words: Asymptotic theory; Circular data; Density estimation; Embedding; Exploratory data analysis; Exponential models; Nonparametric inference; Outliers; Parametric inference; Regression; Robust methods; Rotations; Spherical data; Transformation models; Wrapping.

1 Introduction—Background and Principles

The topic of directional statistics is concerned with observations which are not the familiar counts, real numbers or (unrestricted) vectors but instead are typically either directions in 2- or 3-dimensional space or rotations of such a space. A direction can be regarded as a unit vector and represented by a column vector $\mathbf{x}$ satisfying $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$. A rotation can be represented by a square matrix $\mathbf{X}$ such that $\mathbf{X}^{\mathrm{T}}\mathbf{X} = \mathbf{I}$ and $\det(\mathbf{X}) = 1$. Thus the typical sample spaces of directional statistics are the unit circle S^1, the unit sphere S^2, and the group so(3) of (proper) rotations of R^3. Such sample spaces occur naturally in a wide variety of scientific problems, notably in the earth sciences, astronomy and biology. The classical spherical sample spaces are the earth and the celestial sphere while the compass dial provides a commonly-occurring circular sample space. An appropriate general mathematical setting for directional statistics considers sample spaces which are general Riemannian manifolds. However, manifolds other than circles, spheres and rotation groups rarely occur in applications. As a reasonable compromise between undue restriction to the circle and the usual sphere on the one hand and excessive generality on the other, we shall here consider the sample space to be S^{p-1}, the unit $(p-1)$-dimensional sphere. Thus S^{p-1} is the set of vectors $\mathbf{x}$ in R^p satisfying $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$.

The aim of this paper is to survey the developments in directional statistics since Mardia (1975a), emphasising that the special nature of the sample spaces gives rise to some distinctive features and showing that this topic provides a fascinating meeting point of elegant mathematics, general theory and practical applications. Many of these developments can be found in the two recent key books on the topic; Watson (1983a) introduced some of the more mathematical ideas whereas the book of Fisher, Lewis & Embleton (1987) brings together current ideas on practical analysis of spherical data. A very brief overview is given in Mardia (1988b) but it does not discuss the unifying themes considered here.

This paper has been arranged in three parts. The first (§§ 2–8) is concerned with parametric methods. The second part (§§ 9–11) deals with nonparametric methods. Various other topics (non-standard applications, informal methods and non-spherical sample spaces) are treated in the final part (§§ 12–15).

From the theoretical viewpoint, the recent developments in directional statistics are based mainly on a few key ideas and three basic approaches which we now describe.

1.1 *Key Ideas*

(i) *Exponential families.* Families of densities

$$f(\mathbf{x};\boldsymbol{\beta}) = \exp\{\boldsymbol{\theta}(\boldsymbol{\beta})^{\mathrm{T}}\mathbf{t}(\mathbf{x}) - c(\boldsymbol{\beta})\}$$

in which the canonical statistic $\mathbf{t}$ is a suitably chosen vector-valued function on the sample space play an important role in modelling directional data.

(ii) *Transformation structure.* The group $\mathrm{so}(p)$ of rotations of R^p consists of the $p \times p$ matrices $\mathbf{U}$ satisfying $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I}_p$ and $\det \mathbf{U} = 1$. As rotations transform unit vectors into unit vectors, $\mathrm{so}(p)$ acts (transitively) on S^{p-1} by $(\mathbf{U}, \mathbf{x}) \mapsto \mathbf{U}\mathbf{x}$. Thus it is natural to consider transformation models (Barndorff-Nielsen et al., 1982) for directional data.

(iii) '*Tangent-normal*' *decomposition.* Given a unit vector $\boldsymbol{\mu}$ in R^p, the subgroup $\mathrm{so}(p-1)$ of rotations about $\boldsymbol{\mu}$ preserves vectors which are multiples of $\boldsymbol{\mu}$ and sends the subspace orthogonal to $\boldsymbol{\mu}$ into itself. Any unit vector $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = t\boldsymbol{\mu} + (1 - t^2)^{\frac{1}{2}}\boldsymbol{\xi}, \tag{1.1}$$

where $\boldsymbol{\xi}$ is a unit tangent at $\boldsymbol{\mu}$ to the unit sphere, that is $\|\boldsymbol{\xi}\| = 1$, $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\mu} = 0$; see Fig. 1. This decomposition is particularly useful for studying concentrated distributions or distributions with symmetry about some axis $\boldsymbol{\mu}$, that is, distributions with probability density function $f(.)$ satisfying $f(\mathbf{U}\mathbf{x}) = f(\mathbf{x})$ if $\mathbf{U}\boldsymbol{\mu} = \boldsymbol{\mu}$.

(iv) *Transformation to a multivariate problem.* The inclusion of the sphere S^{p-1} in its surrounding Euclidean space R^p turns every directional problem into a (singular) multivariate problem. This inclusion can be generalised as follows. Any suitable sequence $\{a_k\}$ of real numbers gives rise to a transformation $\mathbf{t}$ of S^{p-1} into $L^2(S^{p-1})$, the Hilbert space of square-integrable functions on S^{p-1}. Infinite-dimensional versions of ordinary

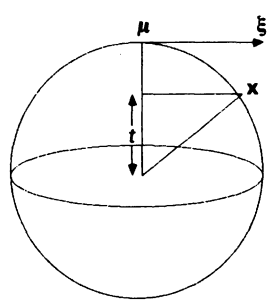

Figure 1. *Tangent-normal decomposition* $\mathbf{x} = t\boldsymbol{\mu} + (1 - t^2)^{\frac{1}{2}}\boldsymbol{\xi}$ *of unit vector* $\mathbf{x}$.

(non-normal) multivariate methods can then be applied. However, the slogan 'directional is singular multivariate' is useful but limited. Note that the singular nature of the sample space leads to moment relationships such as

$$\operatorname{tr} \boldsymbol{\Sigma} + \|\boldsymbol{\mu}\|^2 = 1$$

between the mean $\boldsymbol{\mu}$ and the variance matrix $\boldsymbol{\Sigma}$ of a random vector on the unit sphere.

(v) *Central limit theorem.* As usual, the central limit theorem yields asymptotic ($n \to \infty$) normality of $\bar{\mathbf{x}}$ and

$$n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}},$$

which can be exploited for large-sample inference.

1.2 Basic Approaches

There are three basic approaches to directional statistics, which may be termed the *embedding, wrapping* and *intrinsic* approaches.

(a) In the embedding approach, the sphere is regarded as a subset of p-dimensional Euclidean space.
(b) In the wrapping approach, tangent vectors $\boldsymbol{\xi}$ to the sphere at $\boldsymbol{\mu}$ are wrapped onto the sphere by the mapping

$$\boldsymbol{\xi} \mapsto t\boldsymbol{\mu} + (1 - t^2)^{\frac{1}{2}} \boldsymbol{\xi}, \tag{1.2}$$

where $t = \sin(\|\boldsymbol{\xi}\|)$ and $\boldsymbol{\xi}^{\mathrm{T}} \boldsymbol{\mu} = 0$. In the circular case ($p = 2$), if $\boldsymbol{\mu}$ is regarded as a complex number of unit modulus then (1.2) takes the form $\theta \mapsto e^{i\theta} \boldsymbol{\mu}$ and so is essentially reduction mod 2π. The wrapping map can be used, for example, to transfer distributions on a tangent plane to the sphere into distributions on the sphere itself. (More generally, if the sample space is a Riemannian manifold, wrapping can be carried out using the exponential map.)
(c) In the intrinsic approach the sphere is regarded just as a manifold in its own right without reference to any embedding.

In many ways the intrinsic approach is the most natural. However, the embedding approach, in the general form discussed in §§ 2.1 and 3.1, has been particularly useful for constructing models and tests. The wrapping approach has been used mainly for studying concentrated distributions on the sphere. For a further comparison of the three approaches see, in particular, the discussion of the types of model to which they give rise at the end of § 2.1.

2 Parametric Models

2.1 Univariate Models

Introduction. The simplest model for spherical data is the uniform distribution on the sphere. This is the unique distribution which is invariant under rotations. However, uniformity often provides a poor fit to data and more complicated models are required. Because, in general, there is no preferred coordinate system on the sphere, it is natural to consider families of distributions which are closed under rotations, i.e. if $\mathbf{x}$ has a distribution in this family then so does $\mathbf{Ux}$ for all rotations $\mathbf{U}$. Models generated in this way by transformations of the sample space are known as transformation models. All the families of distributions discussed in this section are transformation models.

Perhaps the most widely-used family of distributions employed for modelling spherical data is that of the von Mises–Fisher (or Langevin) distributions $F(\boldsymbol{\mu}, \kappa)$. This has probability density functions

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = a(\kappa)^{-1} \exp\{\kappa \boldsymbol{\mu}^{\mathrm{T}} \mathbf{x}\} \tag{2.1}$$

with respect to the uniform distribution, where $\kappa \geqslant 0$ and $\|\boldsymbol{\mu}\| = 1$. In equation (2.1) the canonical parameter $\kappa\boldsymbol{\mu}$ has been decomposed into its 'polar coordinates', the concentration parameter κ and the mean direction $\boldsymbol{\mu}$. That this family plays a central role in modelling directional data is due to (i) the particularly pleasant inferential properties of exponential families, and (ii) the symmetry of (2.1) under rotations about the mean direction $\boldsymbol{\mu}$. Each of these properties has given rise to further models for directional data.

Exponential families. The first class of generalisations of (2.1) is to other exponential families. Most generally, Beran (1979) has proposed replacing $\mathbf{x}$ in the exponent in (2.1) by higher polynomials $\mathbf{t}(\mathbf{x})$ in $\mathbf{x}$. In particular, the use of general quadratics in $\mathbf{x}$ yields the Fisher–Bingham family (Mardia, 1975a)

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa, \mathbf{A}) = a(\kappa, \mathbf{A})^{-1} \exp\{\kappa \boldsymbol{\mu}^{\mathrm{T}} \mathbf{x} + \mathbf{x}^{\mathrm{T}} \mathbf{A} \mathbf{x}\}, \tag{2.2}$$

where $\mathbf{A}$ is a symmetric $p \times p$ matrix and the constraint $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$ allows us to assume without loss of generality that $\operatorname{tr}(\mathbf{A}) = 0$. Distributions in this family can also be obtained by conditioning p-variate normal distributions on $\|\mathbf{x}\| = 1$.

Because the family (2.2) has $p(p+1)/2 + p - 1$ identifiable parameters it is rather large for general use and so interest has centred on suitable subfamilies. Perhaps the subfamily with the most potential use is the $\{(p-1)p/2 + p - 1\}$-parameter family introduced by Kent (1982) which is obtained by imposing the restriction $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$ in addition to $\operatorname{tr}(\mathbf{A}) = 0$. This family can be regarded as a spherical analogue of the general $(p-1)$-variate normal family because (under certain restrictions on the parameters) its distributions have modes at $\boldsymbol{\mu}$ and density contours which are approximately elliptical. Relaxing the restriction $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$ to the condition that $\boldsymbol{\mu}$ be an eigenvector of $\mathbf{A}$, yields the 1-parameter extension of the Kent family suggested by Rivest (1984b). A smaller useful family is the $2p$-parameter Fisher(–Dimroth)–Watson family introduced by Wood (1988) which is obtained from (2.2) by the restriction that the matrix $\mathbf{A}$ has rank 1 (instead of $\operatorname{tr}(\mathbf{A}) = 0$) so that the densities have the form

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\mu}_0, \kappa, \kappa_0) = a(\kappa_0, \kappa, \boldsymbol{\mu}_0^{\mathrm{T}}\boldsymbol{\mu})^{-1} \exp\{\kappa_0 \boldsymbol{\mu}_0^{\mathrm{T}} \mathbf{x} + \kappa(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x})^2\}. \tag{2.3}$$

Further families of distributions with interesting geometrical properties appropriate for modelling phenomena from various fields can be obtained by suitable restriction of the parameters of the Fisher–Bingham family. In particular, certain problems in the earth sciences require rotationally-symmetric spherical distributions which have a 'modal ridge' along a small circle rather than a mode at a single point. Suitable models include the 'small circle' family of Bingham & Mardia (1978) which has probability density functions

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa, \nu) = a(\kappa)^{-1} \exp\{\kappa(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x} - \nu)^2\}. \tag{2.4}$$

It is a $(p+1)$-parameter curved exponential subfamily of the Fisher–Watson family (2.3). Another family of 'small circle' distributions on the sphere was introduced by Mardia & Gadsden (1977). As this family has log-densities proportional to $\alpha\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x} + \beta(1 - (\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x})^2)^{\frac{1}{2}}$, it is not an exponential family.

In order to model bimodal data on S^{p-1}, Wood (1982) modified the Fisher family by 'doubling the longitude'. The density functions are

$$f(\mathbf{x}; \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \kappa, \alpha) = a(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \kappa, \alpha)^{-1} \times \exp\{\kappa \cos\alpha[\mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}_1] + \kappa \sin\alpha[2(\mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}_2)^2 - 1]/(1 - (\mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}_1)^2)^{\frac{1}{2}}\}, \quad (2.5)$$

where $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ are orthogonal unit vectors, $\kappa \geqslant 0$ and α is real. These distributions have mean $\boldsymbol{\mu}_1$ and two modes of equal strength at $\cos\alpha\boldsymbol{\mu}_1 \pm \sin\alpha\boldsymbol{\mu}_2$. Note that (2.5) is not an exponential family.

Models with symmetry. The other important property of the von Mises–Fisher distributions is that they are rotationally symmetric about the modal direction. Saw (1978, 1984) and Watson (1983a, pp. 92, 136–186) have abstracted this property by considering general distributions with such symmetry. (See also M.S. Bingham & Mardia, 1975). These distributions have probability density functions

$$f(\mathbf{x}) = g(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x}), \quad (2.6)$$

where $g(.)$ is a known function. Then, in the 'tangent-normal' decomposition

$$\mathbf{x} = t\boldsymbol{\mu} + (1 - t^2)^{\frac{1}{2}}\boldsymbol{\xi} \quad (2.7)$$

(where $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\xi} = 0$), symmetry implies that the unit tangent $\boldsymbol{\xi}$ at $\boldsymbol{\mu}$ to S^{p-1} is uniformly distributed on S^{p-2} and is independent of t. Among the families of distributions of the form (2.6) considered by Saw is a p-parameter family with the property that if $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ has a distribution in this family then $\|\mathbf{x}_1\|^{-1}\mathbf{x}_1$ and $\|\mathbf{x}_2\|^{-1}\mathbf{x}_2$ are independent with distributions in the corresponding families on S^{q-1} and S^{p-q-1}. Watson (1983a, Ch. 4) showed that many families of the form (2.6) have the same asymptotic (large sample-size or high concentration) behaviour as the von Mises–Fisher family. Distributions with symmetry of the form (2.6) can arise through (i) symmetrisation performed by the observation process, e.g. the rotation of the earth as in the motivating problem for the 'rotating spherical cap' distribution of Mardia & Edwards (1982), or (ii) the inability of the observation procedure to distinguish between $\mathbf{x}$ and $\mathbf{U}\mathbf{x}$, where $\mathbf{U}$ is a rotation about $\boldsymbol{\mu}$, so that only the colatitudes (or equivalently $\mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}$) are observed, as in Clark's (1983) motivation for the 'marginal Fisher' distribution introduced by Mardia & Edwards (1982). Watson (1983a, pp 94–95; 1983c) has generalised the axial symmetry associated with (2.6) to rotational symmetry about some subspace V of R^p. The corresponding distributions have probability density functions of the form

$$f(\mathbf{x}) = g(\mathbf{x}_V),$$

where $\mathbf{x}_V$ is the orthogonal projection of $\mathbf{x}$ onto V.

Some types of observation are *axial* in that $\mathbf{x}$ and $-\mathbf{x}$ cannot be distinguished, so that the observations are of *axes* rather than of directions. Thus, for modelling axial data, it is appropriate to use models with antipodal symmetry, i.e. densities $f(.)$ satisfying $f(-\mathbf{x}) = f(\mathbf{x})$. More generally, if $\mathbf{x}$ and $\mathbf{U}\mathbf{x}$ are indistinguishable for all rotations $\mathbf{U}$ in some group G, then it is appropriate to consider only models which are symmetric under G. Rivest (1984b) has shown how such symmetry leads to a simple form of the Fisher information matrix. A model for axial data can be obtained from (2.2) by putting $\kappa = 0$. This is the $\{(p+1)p/2 - 1\}$-parameter Bingham (1974) family, with densities

$$f(\mathbf{x}; \mathbf{A}) = a(\mathbf{A})^{-1} \exp\{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\}. \quad (2.8)$$

A simple p-parameter subfamily of the Bingham family is the (Dimroth–Scheidegger–) Watson family with densities

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = a(\kappa, \boldsymbol{\mu})^{-1} \exp\{\kappa(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x})^2\}, \quad (2.9)$$

obtained by adding the restriction that $\mathbf{A}$ has rank 1 (and removing the condition that $\operatorname{tr}(\mathbf{A})=0$). A one-parameter extension of the Bingham family was introduced by Kelker & Langenberg (1982) in order to model axes concentrated asymmetrically near a small circle. It has density functions

$$f(\mathbf{x}; \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \kappa_1, \kappa_2, \gamma) = a(\kappa_0, \kappa_1, \gamma)^{-1} \exp\{-\kappa_1 \sin^2(\theta-\gamma) + \kappa_2 \cos^2(\theta-\gamma)\sin^2\phi\},$$

where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ are orthogonal unit vectors, κ_1, κ_2 are positive, $\gamma \in [0, 2\pi]$ and (θ, ϕ) is defined by $\cos\theta = \boldsymbol{\mu}_1^{\mathrm{T}}\mathbf{x}$ and $\sin\theta\cos\phi = \boldsymbol{\mu}_2^{\mathrm{T}}\mathbf{x}$. This family is not an exponential family but for given γ it forms a transformation model.

The relationships between the various exponential families discussed in this section are shown in Table 1.

Although exponential families have many pleasant inferential properties, the need to evaluate the norming constant (or at least the first derivative of its logarithm) can be a practical difficulty. For directional distributions these norming constants are often given in terms of special functions such as Bessel functions. Some simplifications have been effected by de Waal (1979) and Wood (1988). Asymptotic expansions for the norming constant in various types of concentrated Bingham distribution have been developed by Kent (1987). In the general context of exponential families on spheres, Beran (1979) has developed a regression-based estimator (using nonparametric density estimation) which bypasses the need for calculation of the norming constant.

Insight into the models. The three approaches to directional statistics mentioned in § 1 can each be used in the construction of models. The embedding approach regards the sphere as a subset of p-dimensional Euclidean space. Taking $\mathbf{t}(\mathbf{x}) = \mathbf{x}$ (for $\|\mathbf{x}\| = 1$) as the canonical statistic of an exponential family leads to the von Mises–Fisher family (2.1). A generalisation of this approach replaces the inclusion of S^{p-1} in R^p by another vector-valued function $\mathbf{t}$ on S^{p-1} (compare key idea (iv) of § 1.1). Taking $\mathbf{t}(\mathbf{x}) = \mathbf{x}\mathbf{x}^{\mathrm{T}}$ as the

Table 1

Some exponential families of distributions on S^{p-1}, showing name, equation number in text, and dimension. Arrows denote inclusions.

$$\begin{tikzcd}
 & \begin{array}{c}\text{Fisher–Bingham (2.2)}\\ p(p+1)/2+p-1\end{array} & \\
 & \begin{array}{c}\text{Kent (2.2) } \mathbf{A}\boldsymbol{\mu}=\mathbf{0}\\ p(p-1)/2+p-1\end{array} \arrow[u] & \\
 & \begin{array}{c}\text{Fisher–Watson (2.3)}\\ 2p\end{array} \arrow[u] & \begin{array}{c}\text{Bingham (2.8)}\\ p(p+1)/2-1\end{array} \arrow[ul] \\
\begin{array}{c}\text{von Mises–Fisher (2.1)}\\ p\end{array} \arrow[ur] & \begin{array}{c}\text{Bingham–Mardia (2.4)}\\ p+1\end{array} \arrow[u] & \\
 & \begin{array}{c}\text{Watson (2.9)}\\ p\end{array} \arrow[u] \arrow[uur] &
\end{tikzcd}$$

canonical statistic of an exponential family yields the Bingham family (2.8). The embedding approach leads also to the use of 'offset normal' (or 'angular Gaussian') distributions. These are obtained by projecting p-variate normal distributions onto S^{p-1} using the radial map $\mathbf{x} \mapsto \|\mathbf{x}\|^{-1}\mathbf{x}$, that is, the length of $\mathbf{x}$ is ignored. In particular, use of normal distributions with mean zero yields the angular central Gaussian distributions considered by Tyler (1987), which possess antipodal symmetry and so can be used for modelling axial data.

In the wrapping approach, tangent vectors to the sphere at a given point $\boldsymbol{\mu}$ are wrapped onto the sphere by the mapping (1.2) and so distributions on the tangent plane give rise to distributions on the sphere. In particular, giving $\boldsymbol{\xi}$ a $(p-1)$-variate normal distribution and letting the parameter $\boldsymbol{\mu}$ vary over S^{p-1} yields the family of wrapped normal distributions.

In the intrinsic approach, intrinsic properties of the sphere (without reference to any embedding) can be used for the construction of models. For example, the distribution at a given time of particles starting at a fixed point and then undergoing Brownian motion on the sphere gives rise to the Brownian motion (or 'spherical normal') distributions. In fact, the von Mises–Fisher, Brownian motion and 'offset normal' distributions are surprisingly close. See Roberts & Ursell (1960), Kendall (1974a,b), Kent (1978a) and Watson (1982a). Further evidence of this closeness was given in the circular case by Collett & Lewis (1981), who showed that large sample sizes are required in order to distinguish between the von Mises and wrapped normal (Brownian motion) distributions.

2.2 Bivariate Models and Regression Models

The natural analogue of (2.1) for modelling bivariate spherical data is the $p(p+2)$-parameter family of densities

$$f(\mathbf{x}, \mathbf{y}; \kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{A}) = a(\kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{A}) \exp\{\kappa_1\boldsymbol{\mu}_1^{\mathrm{T}}\mathbf{x} + \kappa_2\boldsymbol{\mu}_2^{\mathrm{T}}\mathbf{y} + \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y}\}, \tag{2.10}$$

introduced by Mardia (1975a,c). Although the conditional distributions of $\mathbf{x}$ and $\mathbf{y}$ belong to the Fisher family (2.1), the marginal distributions are more complicated. For the special case in which the dependency matrix $\mathbf{A}$ is a multiple of a rotation matrix, Jupp & Mardia (1980) found an expansion in terms of products of Bessel functions for the norming constant $a(\kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$. Rivest (1988) considers the $\{(p^2+p+6)/2\}$-parameter subfamily of (2.10) for which the densities have rotational symmetry in that they are invariant under $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{U}\mathbf{x}, \mathbf{R}\mathbf{U}\mathbf{R}^{-1}\mathbf{y})$, for rotations $\mathbf{U}$ about $\boldsymbol{\mu}_1$, where $\mathbf{R}$ is a rotation such that $\mathbf{R}\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. Thus the matrix $\mathbf{A}$ has the form $\mathbf{A} = \mathbf{H}_1^{\mathrm{T}} \operatorname{diag}(\alpha, \beta\mathbf{Q})\mathbf{H}_2$, where $\mathbf{Q} \in O(p-1)$, α and β are real numbers, and $\mathbf{H}_i$ rotates $\boldsymbol{\mu}_i$ to $(1, 0, \ldots, 0)^{\mathrm{T}}$ for $i = 1,2$.

More general bivariate models are obtained by using densities of the form

$$f(\mathbf{x}, \mathbf{y}; \kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{A}) = a(\kappa_1, \kappa_2, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{A}) \exp\{\kappa_1\boldsymbol{\mu}_1^{\mathrm{T}}\mathbf{t}(\mathbf{x}) + \kappa_2\boldsymbol{\mu}_2^{\mathrm{T}}\mathbf{u}(\mathbf{y}) + \mathbf{t}(\mathbf{x})^{\mathrm{T}}\mathbf{A}\mathbf{u}(\mathbf{y})\}, \tag{2.11}$$

where $\mathbf{t}(.)$ and $\mathbf{u}(.)$ are suitable vector-valued functions. For example, in the case where $\mathbf{x}$ is scalar and $\mathbf{y}$ is circular (so that the sample space is a cylinder), Mardia & Sutton (1978) used a subfamily of (2.11) with $\mathbf{t}(x) = (x, x^2)$, $\mathbf{u}(\mathbf{y}) = \mathbf{y}$ and convenient restrictions on $\mathbf{A}$. Other subfamilies of (2.11) were obtained by Johnson & Wehrly (1978) from maximum entropy arguments.

Johnson & Wehrly (1978) and Wehrly & Johnson (1980) produced a way of constructing bivariate circular models which uses the cumulative distribution function and so seems peculiar to the circular case, since there appears to be no coordinate-free definition of cumulative distribution function on general spheres.

A wide variety of regression models are obtained as generalised linear models with error distributions in an appropriate exponential family so that the density of $\mathbf{y}$ given $\mathbf{x}$ is proportional to

$$\exp\{[\boldsymbol{\phi} + \mathbf{A}^{\mathrm{T}}\mathbf{t}(\mathbf{x})]^{\mathrm{T}}\mathbf{u}(\mathbf{y})\}. \tag{2.12}$$

An example of (2.12) in which $\mathbf{x}$ takes values in S^2 and $\mathbf{y}$ is axial in S^1 is the model used by Kendall & Young (1984) to examine a fascinating problem in radio-astronomy. Other examples of (2.12) are given by Johnson & Wehrly (1978).

Nonlinear regression occurs naturally in directional statistics. In some contexts, such as finding the source $\mathbf{a}$ of a signal or the centre of an explosion, directions $\mathbf{y}_i$ are observed at positions $\mathbf{x}_i$. A convenient model for regression of a spherical $\mathbf{y}$ on a vector $\mathbf{x}$ takes $\mathbf{y} \mid \mathbf{x}$ to have a Fisher distribution with mean direction pointing towards an unknown 'centre', $\mathbf{a}$, so that

$$\mathbf{y} \mid \mathbf{x} \sim F(\|\mathbf{x}-\mathbf{a}\|^{-1}(\mathbf{x}-\mathbf{a}), \kappa_0 + \kappa\,\|\mathbf{x}-\mathbf{a}\|^c). \tag{2.13}$$

For various particular cases of (2.13) with $\kappa_0 = 0$ see Lenth (1981b) (who considers also robust estimation in this context), Jupp, Spurr, Nichols & Hirst (1987) and Jupp & Spurr (1989). Another important non-linear model is that used by Mackenzie (1957) and Chang (1986) for regression of a spherical variate $\mathbf{y}$ on a spherical predictor $\mathbf{x}$. This takes $\mathbf{y} \mid \mathbf{x}$ to have a circularly symmetric distribution with mean direction $\mathbf{Ax}$, so that

$$f(\mathbf{y}; \mathbf{A} \mid \mathbf{x}) = g(\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}), \tag{2.14}$$

where $\mathbf{A}$ is an unknown rotation. Chang (1989) investigated the associated errors in variables model in which $(\mathbf{x}_i, \mathbf{y}_i)$, $i = 1, \ldots, n$, are independent with densities

$$f(\mathbf{x}_i, \mathbf{y}_i; \boldsymbol{\xi}_i, \mathbf{A}) = g(\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\xi}_i)g(\mathbf{y}_i^{\mathrm{T}}\mathbf{A}\boldsymbol{\xi}_i)$$

for some known function g(.) and unknown rotation $\mathbf{A}$ and unit vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$. He provided asymptotic tests and confidence regions for $\mathbf{A}$ in the cases in which either the sample size is large or $\mathbf{x}_i$ and $\mathbf{y}_i$ have concentrated Fisher distributions. Detailed consideration of a related problem in plate tectonics led Chang (1988) to the model in which the spherical random variables $\mathbf{x}_{ij}$ and $\mathbf{y}_{ik}$ $(1 \leqslant j \leqslant m_i,\ 1 \leqslant k \leqslant n_i,\ 1 \leqslant i \leqslant s)$ have certain independent concentrated wrapped multivariate normal distributions with respective means $\boldsymbol{\alpha}_{ij}$ and $\boldsymbol{\beta}_{ik}$, where $\boldsymbol{\alpha}_{ij}^{\mathrm{T}}\boldsymbol{\eta}_i = 0$ and $\boldsymbol{\beta}_{ik}^{\mathrm{T}}\mathbf{A}\boldsymbol{\eta}_i = 0$ for some unknown unit vectors $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n$. He obtained asymptotic confidence regions for the rotation $\mathbf{A}$.

Discrete time stationary Markov processes on the sphere have been considered by Accardi, Cabrera & Watson (1987). They proposed models based on the generalised linear model construction (2.12), so that $\mathbf{x}_t$ conditional on $\mathbf{x}_{t-1}$ has a Fisher or Watson distribution with natural parameter which is an affine function of $\mathbf{x}_{t-1}$. They used simulation to explore the behaviour of these processes. In the special case of S^1, the group structure and the existence of a cumulative distribution function provided Wehrly & Johnson (1980) with another method of constructing stationary Markov processes on the circle.

3 Inference

An essential requirement of inferential procedures used for directional data is that they are invariant under rotations so that conclusions do not depend on any coordinate system chosen. Furthermore, a coordinate-free description of such procedures often aids geometrical insight.

3.1 Testing Uniformity

A fundamental hypothesis about directional distributions is that of uniformity. Two key tests of uniformity are Rayleigh's test (which rejects uniformity if the resultant length $\|\mathbf{x}_1 + \ldots + \mathbf{x}_n\|$ of a sample is too large) and Watson's (1961) U^2 test (which is an adaptation for the circle of the Cramér–von Mises test). Numerous other tests of uniformity have been proposed. Many of these fit into Giné's (1975) general framework which we now describe.

This framework is based on key idea (iv) of § 1.1 and extends the embedding approach of § 1.2 by transforming directional problems into infinite-dimensional multivariate problems as follows. Naturally associated with the sphere S^{p-1} is the Hilbert space $L^2(S^{p-1})$ of square-integrable functions on S^{p-1}. Any suitable mapping $\mathbf{t}$ of S^{p-1} into $L^2(S^{p-1})$ transforms distributions on S^{p-1} into distributions on $L^2(S^{p-1})$. Since the Hilbert space $L^2(S^{p-1})$ behaves just like familiar Euclidean space, infinite-dimensional versions of ordinary (non-normal) multivariate methods can then be applied. Desirable properties of $\mathbf{t}$ are continuity and equivariance. By *equivariance* we mean that for every rotation $\mathbf{U}$ of S^{p-1}

$$\mathbf{t}(\mathbf{U}\mathbf{x}) = \mathbf{U}\mathbf{t}(\mathbf{x}), \quad \mathbf{x} \in S^{p-1},$$

where on the right-hand side $\mathbf{U}$ denotes the corresponding rotation of $L^2(S^{p-1})$ given by $f(.) \mapsto f(\mathbf{U}.)$. Note that, if $\mathbf{t}$ is equivariant, then $\|\mathbf{t}(\mathbf{x})\|$ is constant as $\mathbf{x}$ runs through S^{p-1}, so that $\mathbf{t}(\mathbf{x})$ is a point in an infinite-dimensional sphere.

One way of constructing such mappings $\mathbf{t}$ is the following method due to Giné (1975) and based on the eigenfunctions of the Laplacian on S^{p-1}; see also Jupp & Spurr (1985) for the Hilbert space approach given here.

Let E_k denote the space of eigenfunctions corresponding to the kth eigenvalue, for $k \geq 1$. Then there is a well-defined map $\mathbf{t}_k$ of S^{p-1} into E_k given by

$$\mathbf{t}_k(\mathbf{x}) = \sum_{i=1}^{n_k} f_i(\mathbf{x}) f_i, \tag{3.1}$$

where $\{f_i : 1 \leq i \leq n_k\}$ is any orthonormal base of E_k. If $\{a_k\}$ is a sequence which converges sufficiently rapidly to 0, an equivariant mapping $\mathbf{t}$ of S^{p-1} into $L^2(S^{p-1})$ can be defined by

$$\mathbf{x} \mapsto \mathbf{t}(\mathbf{x}) = \sum_{k=1}^{\infty} a_k \mathbf{t}_k(\mathbf{x}). \tag{3.2}$$

For example, on the circle, provided that

$$\sum_{k=1}^{\infty} a_k^2 < \infty,$$

the angle θ is transformed into the *function*

$$x \mapsto \sum_{k=1}^{\infty} a_k \cos k(x - \theta).$$

In the general case, under mild conditions on the sequence $\{a_k\}$, $\mathbf{t}$ is continuous. Then $\mathbf{t}$ gives rise to a mapping $\boldsymbol{\tau}$ of the set $P(S^{p-1})$ of all probability distributions on S^{p-1} into $L^2(S^{p-1})$ defined by

$$\boldsymbol{\tau}(\nu) = \int \mathbf{t}(\mathbf{x})\, d\nu(\mathbf{x}).$$

If all the coefficients a_k are non-zero then $\boldsymbol{\tau}$ is a one-to-one function (indeed, a topological embedding) and so gives rise to tests which are consistent against all alternatives. Another way of looking at the transformation $\mathbf{t}$ given by (3.2) is to regard it as a 'weighted characteristic function', sending a distribution on S^{p-1} into the sequence $\{a_k E[\mathbf{t}_k(\mathbf{x})]\}$ of its 'Fourier coefficients'.

Any suitable mapping $\mathbf{t}$ as above gives a test for uniformity which rejects uniformity for large values of the 'resultant length' $\|\mathbf{t}(\mathbf{x}_1) + \ldots + \mathbf{t}(\mathbf{x}_n)\|$ of the transformed observations, where $\|.\|$ denotes the L_2 norm on $L^2(S^{p-1})$. Thus it is an infinite-dimensional version of Rayleigh's test. Note that the equivariance of $\mathbf{t}$ ensures that the test is invariant under rotations. Note also that the test is based on the V-statistic

$$T_n = n^{-1} \|\mathbf{t}(\mathbf{x}_1) + \ldots + \mathbf{t}(\mathbf{x}_n)\|^2 = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{t}(\mathbf{x}_i), \mathbf{t}(\mathbf{x}_j) \rangle, \tag{3.3}$$

where $\langle , \rangle$ denotes the inner-product on $L^2(S^{p-1})$ given by

$$\langle f, g \rangle = \int f(x) g(x)\, d\omega(x)$$

with ω denoting the uniform probability measure on S^{p-1}. Under uniformity, the distribution of T_n tends as $n \to \infty$ to that of

$$\sum_{k=1}^{\infty} a_k^2 U_k,$$

where the U_k have independent χ^2 distributions on dim E_k degrees of freedom.

The construction of tests of uniformity using mappings $\mathbf{t}$ into $L^2(S^{p-1})$ as above generalises readily to sample spaces which are compact Riemannian manifolds. Giné's (1975) presentation of these tests in this general case is equivalent to that above but used Sobolev spaces of generalised functions on S^{p-1} with square-integrable derivatives of appropriate order (Giné, 1975, pp. 1246–1247). Hence these tests are sometimes called 'Sobolev tests'. In the case of sample spaces which are compact homogeneous spaces (being acted on transitively by a group G) Beran (1968) obtained these tests in the form

$$T_n = n^{-1} \int \left[\sum_{i=1}^{n} f(g\mathbf{x}_i) - n \right]^2 dg$$

as locally most powerful invariant tests of uniformity against the alternative of a density $f(g\mathbf{x})$, where g is an unknown element of G,

$$f(\mathbf{x}) = 1 + \sum_{k=1}^{\infty} a_k \sum_{i=1}^{n_k} f_i(\mathbf{x})$$

and $\{f_i : 1 \leq i \leq n_k\}$ are as in (3.1).

Various nonparametric tests which are in the spirit of the Sobolev tests for uniformity are considered in § 9.1.

Appropriate choices of the coefficients a_k in (3.2) yield tests with suitable properties of consistency, ease of computation, etc. For example, on the circle the Rayleigh test corresponds to taking $a_k = 0$ for $k > 0$ and is easily calculated, whereas Watson's U^2 is given by $a_k = (2\pi k)^{-1}$ and is consistent against all alternatives. Hermans & Rasson (1985) designed a test of this form to be consistent against all alternatives and to be more powerful against multimodal alternatives than Watson's U^2. More explicit forms of these 'Sobolev tests' for uniformity of directional and axial distributions on S^{p-1} have been considered by Prentice (1978), who also gave large-sample approximations to the null distributions.

Diggle, Fisher & Lee (1985) have carried out a power study of three such tests of uniformity: the Rayleigh test, Giné's (1975) F_n, and a third test which is based on the number of pairs of data points less than a certain distance apart. They found that, against equal mixtures of Fisher distributions with the modes not too far apart, the Rayleigh test was most powerful in small samples but was dominated by the F_n test in larger samples. Against Fisher mixtures with modes 180° apart, the third test was most powerful. For testing uniformity of axial data against Watson alternatives, Bingham's (1974) test of uniformity and Giné's (1975) G_n were found to have comparable power and to be preferable to the axial analogue of the third test.

3.2 *Exact Inference*

For estimation in the families considered in § 2, the standard theory of exponential families shows that maximum likelihood estimation is, in principle, straightforward. In particular, for a sample from the Fisher family (2.1), the maximum likelihood estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\kappa}$ of $\boldsymbol{\mu}$ and κ are given by

$$\hat{\boldsymbol{\mu}} = \|\bar{\mathbf{x}}\|^{-1}\bar{\mathbf{x}}, \quad A_p(\hat{\kappa}) = n^{-1}R,$$

where $\bar{\mathbf{x}}$ and R are respectively the sample mean and the resultant length of the observations and $A_p(\kappa) = I_{p/2}(\kappa)/I_{p/2-1}(\kappa)$, $I_p(.)$ being the modified Bessel function of the first kind and order p. Best & Fisher (1981) have pointed out that $\hat{\kappa}$ is a biased estimator of κ and have suggested alternative estimators based on Taylor expansion bias-reduction and on jack-knifing. Sufficiency considerations led Schou (1978) to suggest estimating κ by maximising the *marginal* likelihood based on the distribution of the resultant length R. Explicitly, this marginal likelihood estimator $\tilde{\kappa}$ is given by

$$\tilde{\kappa} = 0 \quad \text{if } R \leqslant n^{\frac{1}{2}}, \quad nA_p(\tilde{\kappa}) = RA_p(\tilde{\kappa}R) \quad \text{if } R > n^{\frac{1}{2}}.$$

In the usual case ($p = 3$) of distributions on S^2, simulations indicate that this marginal likelihood estimator is less biased than the usual maximum likelihood estimator κ. Furthermore, in the several sample case, with the number of samples tending to infinity, $\tilde{\kappa}$ is consistent whereas $\hat{\kappa}$ is not.

If the random unit vector $\mathbf{x}$ has a von Mises–Fisher distribution in the family (2.1) then in the tangent-normal decomposition (1.1),

$$\mathbf{x} = t\boldsymbol{\mu} + (1 - t^2)^{\frac{1}{2}}\boldsymbol{\xi}, \tag{3.4}$$

$\boldsymbol{\xi}$ is uniformly distributed on S^{p-1} and is independent of t. In particular, for the usual spherical case ($p = 3$), $1 - t$ has an exponential distribution truncated to $[0, 2]$ with mean approximately equal to κ^{-1}. In terms of spherical polar coordinates θ (colatitude) and ϕ (longitude) with respect to the mean direction ($\theta = 0$),

$$\begin{gathered} \phi \text{ and } \theta \text{ are independent}, \quad \phi \sim U(0, 2\pi), \\ 2\kappa(1 - \cos\theta) \sim \chi^2_2, \quad \kappa \to \infty. \end{gathered} \tag{3.5}$$

Thus, for moderately large κ, inference can be based on the sample versions $c_i = 1 - \mathbf{x}_i^{\mathrm{T}}\hat{\boldsymbol{\mu}}$ of t ($i = 1, \ldots, n$).

3.3 *Approximate Inference*

Large-sample inference based on first or second sample moments (centre of mass and moment of inertia, respectively) can be obtained from the asymptotic results discussed in § 4.1 and has been developed by Watson (1983a, Ch. 4,5; 1983b).

Watson (1983a, pp. 166–167) has pointed out that, provided $E(t^2) \neq p-1$ (in the notation of (3.4)), the dominant eigenvector of the second moment matrix

$$\mathbf{M}_n = n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}}$$

can be a useful estimator of the mean direction of a distribution on S^{p-1}. See Watson (1983a, Ch. 5; 1984a) for the use of eigenvalues and eigenvectors of $\mathbf{M}_n$ in large-sample inference.

The large-sample behaviour of Mackenzie's (1957) estimators of the rotation matrix $\mathbf{A}$ in the rotation model (2.14) was considered by Chang (1986). These estimate $\mathbf{A}$ by the corresponding Procrustes rotation, i.e. the matrix $\hat{\mathbf{A}}$ maximising

$$\sum_{i=1}^{n} \mathbf{y}_i^{\mathrm{T}} \mathbf{A} \mathbf{x}_i.$$

The behaviour of $\hat{\mathbf{A}}$ when $\mathbf{y}_i$ has a concentrated Fisher distribution with mean direction $\mathbf{A}\mathbf{x}_i$ was given by Rivest (1989a).

For concentrated distributions on the sphere the asymptotic results of § 4.2 can be used. For the Fisher family, (2.1) with $p = 3$, large concentration inference is based mainly on (3.5) together with the following. Let (θ'', ϕ'') denote spherical polar coordinates chosen so that the sample mean direction is at $(\pi/2, 0)$ and so that ϕ'' has range $[-\pi, \pi]$.

Then taking a wrapped normal approximation to the von Mises distribution of $\phi'' \mid \theta''$ yields

$$\phi''(\sin\theta'')^{\frac{1}{2}} \sim N(0, \kappa^{-1}), \quad \kappa \to \infty. \tag{3.6}$$

See Lewis & Fisher (1982).

3.4 Correlation Coefficients

Inference about dependence of random variables is often based on sample correlation coefficients. Because it is by no means obvious how to construct sensible rotation-invariant correlation coefficients for use with directional data, this problem has received much attention and numerous correlation coefficients have been proposed. These can conveniently be divided into three types. Two of these types are suitable for distributions on general manifolds and are based on the embedding approach of § 1.2. Population correlation coefficients of the first type are functions of the corrected covariance matrix $\mathbf{\Sigma}$ of the random p-dimensional directions $\mathbf{x}$ and $\mathbf{y}$. Those of the second type use functions of the uncorrected covariance matrix $\mathbf{\Sigma}^*$. Those of the third type are for the circular or linear case only and are based on uniform scores. The sample versions of these are discussed in § 9.1.

Coefficients of the first and second types make use of vector-valued functions $\mathbf{t}$ and $\mathbf{u}$ on S^{p-1}. Let

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}, \quad \mathbf{\Sigma}^* = \begin{pmatrix} \mathbf{\Sigma}_{11}^* & \mathbf{\Sigma}_{12}^* \\ \mathbf{\Sigma}_{21}^* & \mathbf{\Sigma}_{22}^* \end{pmatrix}$$

denote the corrected and uncorrected covariance matrices of $(\mathbf{t}(\mathbf{x}), \mathbf{u}(\mathbf{y}))$, partitioned conformally. It is convenient to assume that $\mathbf{t}$ and $\mathbf{u}$ are continuous and equivariant. Then various correlation coefficients can be obtained based on appropriate functions of the eigenvalues of either $(\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{21})^{\frac{1}{2}}$ (Downs, 1974; Mardia, 1975a) or $(\mathbf{\Sigma}_{12}^*\mathbf{\Sigma}_{21}^*)^{\frac{1}{2}}$ (Stephens, 1979). Alternatively, eigenvalues of $\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}$ (Mardia, 1976a; Johnson & Wehrly, 1977; Jupp & Mardia, 1980) or $\mathbf{\Sigma}_{12}^*\mathbf{\Sigma}_{22}^{*-1}\mathbf{\Sigma}_{21}^*\mathbf{\Sigma}_{11}^{*-1}$ (Fisher & Lee, 1983, 1986) can be used.

Table 2
Some correlation coefficients for directional distributions.

(a) *Coefficients based on the covariance matrix* $\mathbf{\Sigma}$	
$\mathrm{tr}\,[(\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{21})^{\frac{1}{2}}]/\{\mathrm{tr}\,(\mathbf{\Sigma}_{11})\,\mathrm{tr}\,(\mathbf{\Sigma}_{22})\}^{\frac{1}{2}}$	Downs (1974) Mardia (1975a)
Largest eigenvalue of $\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$	Johnson & Wehrly (1977)
$\mathrm{tr}\,(\hat{\mathbf{\Sigma}}_{11}^{-1}\mathbf{\Sigma}_{12}\hat{\mathbf{\Sigma}}_{22}^{-1}\mathbf{\Sigma}_{21})$, $\hat{\mathbf{\Sigma}}_{ii} = A'(\kappa_i)\boldsymbol{\mu}_i\boldsymbol{\mu}_i^{\mathrm{T}} + A(\kappa_i)\kappa_i^{-1}(\boldsymbol{I}_p - \boldsymbol{\mu}_i\boldsymbol{\mu}_i^{\mathrm{T}})$, where $\boldsymbol{\mu}_i = \lVert E\{\mathbf{x}_i\}\rVert^{-1}E\{\mathbf{x}_i\}$	Mardia & Puri (1978)
$\mathrm{tr}\,(\check{\mathbf{\Sigma}}_{11}^{-1}\mathbf{\Sigma}_{12}\check{\mathbf{\Sigma}}_{22}^{-1}\mathbf{\Sigma}_{21})$, $\check{\mathbf{\Sigma}}_{ii}$ obtained from $\mathbf{\Sigma}_{ii}$, special to $p = 2,3$	Mardia & Puri (1978)
$\mathrm{tr}\,(\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21})$	Mardia (1976a) Jupp & Mardia (1980)
(b) *Coefficients based on the uncorrected sums of products matrix* $\mathbf{\Sigma}^*$	
$\mathrm{tr}\,(\mathbf{\Sigma}_{12}^*)$	Epp, Tukey & Watson (1971)
$\mathrm{tr}\,[(\mathbf{\Sigma}_{12}^*\mathbf{\Sigma}_{21}^*)^{\frac{1}{2}}]$	Stephens (1979) after Mackenzie (1957)
Smallest singular value of $\mathbf{\Sigma}_{12}^*$, suitably normalised	Rivest (1982)
$\det\,[(\mathbf{\Sigma}_{11}^*)^{-\frac{1}{2}}\mathbf{\Sigma}_{12}^*(\mathbf{\Sigma}_{22}^*)^{-\frac{1}{2}}]$	Fisher & Lee (1983, 1986)

Related constructions are given in Mardia & Puri (1978) and Rivest (1982). See Table 2. Note that the equivariance of $\mathbf{t}$ and $\mathbf{u}$ ensures that these coefficients are invariant under separate rotations of $\mathbf{x}$ and $\mathbf{y}$. Jupp & Spurr (1985) used the machinery described in § 3.1 to extend these coefficients to the case where the functions $\mathbf{t}$ and $\mathbf{u}$ take values in a Hilbert space.

Some of the above coefficients, such as those of Fisher & Lee (1983, 1986) are signed correlation coefficients which aim to distinguish between positive and negative correlation of $\mathbf{x}$ and $\mathbf{y}$. They take their maximum and minimum values when $\mathbf{x} = \mathbf{R}\mathbf{y}$ with $\mathbf{R}$ a rotation or reflection respectively.

The coefficients of Mardia & Puri (1978) and Jupp & Mardia (1980) are closely connected to the score test of independence of $\mathbf{x}$ and $\mathbf{y}$ in the exponential family (2.10). Kent (1983b) has extended this idea to a more general context by introducing a correlation coefficient which measures the information gain on modelling $\mathbf{x}$ and $\mathbf{y}$ as dependent rather than independent. Stephens' (1979) coefficient is also the maximum of $E\{\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}\}$ over rotations $\mathbf{A}$; that is, it is $\mathrm{tr}\,(\hat{\mathbf{A}}E\{\mathbf{y}^{\mathrm{T}}\mathbf{x}\})$, where $\hat{\mathbf{A}}$ is the corresponding Procrustes rotation.

Note that, for the correlation coefficients of Jupp & Mardia, the nesting of various sub-hypotheses of the hypothesis of independence of $\mathbf{x}$ and $\mathbf{y}$ leads to corresponding decompositions of the coefficients and of the asymptotic chi-squared distributions of their sample versions.

4 Asymptotic and Approximate Results

Asymptotic properties of directional distributions fall into three categories, according to whether the sample size, the concentration, or the dimension tends to infinity.

4.1 Large-Sample Asymptotics

Large-sample asymptotics can be found by using the embedding approach and applying the central limit theorem. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be independent random variables on the sphere with an axially symmetric density of the form (2.6). Denote the sample mean by $\bar{\mathbf{x}}$. Rewrite the tangent-normal decomposition (1.1) as

$$\mathbf{x} = \mathbf{x}_{\boldsymbol{\mu}} + \mathbf{x}_{\perp}, \tag{4.1}$$

where

$$\mathbf{x}_{\boldsymbol{\mu}} = (\mathbf{x}^{\mathrm{T}}\boldsymbol{\mu})\boldsymbol{\mu} = t\boldsymbol{\mu}, \quad \mathbf{x}_{\perp} = (\mathbf{I} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu})\mathbf{x} = (1 - t^2)^{\frac{1}{2}}\boldsymbol{\xi},$$

so that $\mathbf{x}_{\boldsymbol{\mu}}$ and $\mathbf{x}_{\perp}$ are the components of $\mathbf{x}$ parallel to $\boldsymbol{\mu}$ and orthogonal to it. Similar notation will be used for the analogous decomposition of $\bar{\mathbf{x}}$. Then the central limit theorem and the symmetry of the distribution imply that, as $n \to \infty$,

$$n^{\frac{1}{2}}\bar{\mathbf{x}}_{\perp} \sim N(\mathbf{0}, ((1 - E[t^2])/(p-1))(\mathbf{I} - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}})), \tag{4.2}$$

$$n^{\frac{1}{2}}(\bar{\mathbf{x}}_{\boldsymbol{\mu}}^{\mathrm{T}}\boldsymbol{\mu} - E[t]) \sim N(0, \operatorname{var}(t)). \tag{4.3}$$

See Watson (1983a, p. 137; 1983b,c). It follows that

$$2n(p-1)E[t^2](1 - E[t]^2)^{-1}(1 - \boldsymbol{\mu}^{\mathrm{T}}\hat{\boldsymbol{\mu}}) \sim \chi^2_{p-1}, \quad n \to \infty,$$

where $\hat{\boldsymbol{\mu}} = \|\bar{\mathbf{x}}\|^{-1}\bar{\mathbf{x}}$. The special case of von Mises–Fisher distributions is considered in Watson (1983b) and extensions to distributions symmetric about a subspace are given in Watson (1983c). For confidence cones for $\boldsymbol{\mu}$ based on (4.2) and (4.3) see Watson (1983a, pp. 138–140) and Fisher & Lewis (1983). The distribution of $n\,\|\bar{\mathbf{x}}_{\perp}\|^2$ under local alternatives to the hypothesis of given mean direction was given by Watson (1983a, pp. 140–143).

4.2 Large-Concentration Asymptotics

Large-concentration asymptotics for the von Mises–Fisher family (2.1) can be obtained by straight-forward expansion of the density about the mean or as examples of the general results of Jørgensen (1987) on dispersion models. See Watson (1983a, pp. 157–165). In terms of the tangent-normal decomposition (4.1) one obtains

$$\kappa^{\frac{1}{2}}\mathbf{x}_{\perp} \sim N(\mathbf{0}, \boldsymbol{I} - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}), \quad \kappa \to \infty.$$

It follows that

$$2\kappa(1 - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{x}) \sim \chi^2_{p-1}, \quad \kappa \to \infty, \tag{4.4}$$

generalising (3.5). If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent observations from (2.1) with sample mean $\bar{\mathbf{x}}$, then it follows from (4.4) and the 'analysis of variance' decomposition

$$2\kappa n(1 - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{x}}) = 2\kappa n(1 - \|\bar{\mathbf{x}}\|) + 2\kappa n(\|\bar{\mathbf{x}}\| - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{x}})$$

that

$$2\kappa(n - R) = 2\kappa n(1 - \|\bar{\mathbf{x}}\|) \sim \chi^2_{(n-1)(p-1)} \quad \kappa \to \infty, \tag{4.5}$$

$$2\kappa n(\|\bar{\mathbf{x}}\| - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{x}}) \sim \chi^2_{(p-1)} \quad \kappa \to \infty,$$

and these two statistics are asymptotically independent. These results form the basis of various F-tests, e.g. Watson (1956) and Watson & Williams (1956). Watson (1983a, pp. 157–165; 1984b) found the non-null distributions of many of these F-statistics.

Another consequence of (4.5) is that, for large κ and for t moderately small compared with n,

$$(n - t - 1)(t + R_{n-t} - R_n)/\{t(n - t - R_{n-t})\} \sim F_{(p-1)t,(p-1)(n-t-1)}, \tag{4.6}$$

where $R_m = \|\mathbf{x}_1 + \ldots + \mathbf{x}_m\|$. Fisher & Willcox (1978) derived (4.6) and showed by simulation that (for $p=2$, 3 and $t=1$) this approximation to an F-distribution is reasonable for $\kappa \geqslant 3{\cdot}0$ and $n \geqslant 5$.

The high-concentration asymptotic distribution of the term $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ in the log-density of Bingham distributions (2.8) has been investigated by Bingham & Chang (1988).

Note that as the concentration of a directional distribution increases, the corresponding random unit vector is 'distributed on a smaller portion of S^{p-1}' and the embedding, intrinsic and wrapped approaches become indistinguishable.

4.3 High-Dimensional Asymptotics

High-dimensional asymptotics are also available but, so far, have been almost entirely of theoretical interest, although a possible application to large-sample asymptotics of certain permutation distributions was given in Watson (1988a). Let $\mathbf{x}$ be a random vector in S^{p-1} having a distribution which is symmetrical about some q-dimensional subspace V of R^p, so that its density is of the form $f(\mathbf{x}) = c_p^{-1} g(\mathbf{x}_V)$, where $\mathbf{x}_V$ denotes the projection of $\mathbf{x}$ onto V, $g(.)$ is a given function on V and c_p is a norming constant. Watson (1983d) proved that, as $p \to \infty$ with V fixed,

$$p^{\frac{1}{2}}\mathbf{x}_V \sim N(\mathbf{0}, \mathbf{I}_q). \tag{4.7}$$

The case in which $\mathbf{x}$ has a uniform distribution is due to Stam (1982), who proved also the asymptotic $N(\mathbf{0}, \boldsymbol{I})$ distribution of $\{p^{\frac{1}{2}}\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_j : 1 \leqslant i < j \leqslant r\}$, where $\mathbf{x}_1, \ldots, \mathbf{x}_r$ are independent and uniformly distributed on S^{p-1}. Watson (1983d, 1988) gave generalisations of (4.7) for von Mises–Fisher and Watson distributions and for uniform distributions on Stiefel manifolds. In particular, he showed that, if $\mathbf{x}$ has a von Mises–Fisher distribution on S^{p-1} with mean direction $\boldsymbol{\mu}$ in V and with concentration $p^{\frac{1}{2}}\kappa$, then as $p \to \infty$

$$p^{\frac{1}{2}}\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x} \sim N(\kappa, 1),$$

$$p^{\frac{1}{2}}[\mathbf{x}_V - (\boldsymbol{\mu}^{\mathrm{T}}\mathbf{x})\boldsymbol{\mu}] \sim N(\mathbf{0}, \boldsymbol{I}_q - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}})$$

and are asymptotically independent.

5 Probabilistic and Distributional Aspects

Various characterisations of directional distributions have been found. In particular, Kent, Mardia & Rao (1979) showed that the uniform distribution on S^{p-1} is characterized by the independence of the mean direction $\|\bar{\mathbf{x}}\|^{-1}\bar{\mathbf{x}}$ and the mean resultant length $\|\bar{\mathbf{x}}\|$. See also Bingham (1978). M.S. Bingham & Mardia (1975) give a rigorous proof that the Fisher family (2.1) is characterized by the maximum likelihood estimate of $\boldsymbol{\mu}$ being the sample mean direction. Maximum-entropy characterisations of various exponential families used in directional statistics are given in Mardia (1975c). Reuter, see p. 416 of Kendall (1974b), and Gordon & Hudson (1977) have shown that the von Mises distributions are also the distributions at first hitting time on the circle of particles started at the centre and undergoing two-dimensional Brownian motion with drift.

Kent (1983a) has shown the identifiability of finite mixtures of distributions with polynomial log-densities on (products of) spheres, Stiefel manifolds or Euclidean spaces.

Poincaré's theorem on the 'large time' limit of wrapped distributions on the circle has been extended to spheres and general symmetric spaces by Jupp (1984).

Hartman & Watson (1974) showed that the von Mises–Fisher distributions (2.1) are mixtures over the stopping time of Brownian motion distributions. Kent (1977) proved that the von Mises–Fisher distributions are infinitely divisible by showing the infinite divisibility of these mixing distributions. Thus the 'mth roots' of the von Mises–Fisher distributions can also be expressed as stopping-time mixtures of Brownian motion distributions. Lewis (1975) used another approach to show the infinite divisibility of von Mises distributions (on the circle) for sufficiently small values of the concentration parameter κ.

Random walks on spheres and general Riemannian manifolds were studied by Roberts & Ursell (1960). Roberts & Winch (1984) considered various random walks on the rotation group so(3). In particular, they considered the cases (i) each rotation is through a fixed angle about an arbitrary axis, (ii) the angle of rotation is arbitrary but the axis of rotation makes a given angle Θ with a fixed axis. If $\Theta = \pi/2$ in (ii) then the set of such rotations can be identified (via the image of the north pole) with S^2 and the distributions obtained on so(3) correspond to those obtained from random walks on the sphere.

6 Outliers and Goodness-of-Fit Tests

6.1 *Outliers*

Outliers in directional data have been the subject of much attention. Tests of discordancy of observations from a von Mises–Fisher distribution (2.1) fall into three main groups: those based on likelihood-ratio tests; those which exploit (3.5) (with $p = 3$), converting the problem into one of testing for discordancy of an observation from an exponential distribution; and those which consider as a potential outlier the observation most influential on the mean resultant length. Tests of the third type were proposed by Mardia (1975a, p. 390) and various tests of the first and second types were introduced by Collett (1980) and Fisher, Lewis & Willcox (1981) in the circular and spherical ($p = 3$) cases respectively. The latter two papers report on simulation comparisons of various such tests for a single outlier. Both of these studies found consistently good performance from a test of the third type. For the case of several outliers, a test of the second type was given by Kimber (1985) and tests of the third type based on (4.6) were proposed by Fisher, Lewis & Willcox (1981). For the Watson distribution (2.9), Best & Fisher (1986) obtained both formal and graphical tests of discordancy by using the appropriate analogue of (3.5) with $\cos\theta$ replaced by $\cos^2\theta$.

Graphical methods for the detection of outliers are discussed in § 13.

6.2 *Goodness-of-Fit*

Goodness-of-fit tests for directional data are of four types: adaptations of tests of uniformity; tests exploiting symmetry; tests against a specified larger model; and tests based on density estimates. Tests of the first type have been developed only for the circle. They use the fitted cumulative distribution $\hat{F}$ to transform the data to $\hat{F}(\mathbf{x}_1), \ldots, \hat{F}(\mathbf{x}_n)$, which are then tested for uniformity. For example, Lockhart & Stephens (1985) discussed the use of Watson's U^2 statistic for testing goodness-of-fit to von Mises distributions and gave a table of critical points.

Tests of the second type are based on the tangent-normal decomposition (3.4). In the case of von Mises–Fisher distributions, uniformity of $\boldsymbol{\xi}$ on S^{p-1} and exponentiality of $1 - \mathbf{x}^T\boldsymbol{\mu}$ can be tested by standard means. In particular, Stephens' (1974) modifications of Kuiper's test and the Kolmogorov–Smirnov test can be applied respectively to the longitude and colatitude taken with respect to the sample mean as 'north pole'. The independence in (3.5) can be tested using (3.6). That is, $\phi''(\sin\theta'')^{\frac{1}{2}}$ are tested for normality, where (θ'', ϕ'') are spherical polar coordinates chosen so that the sample mean direction is at $(\pi/2, 0)$ and so that ϕ'' has range $[-\pi, \pi]$. Details are given by Fisher &

Best (1984) (together with a power study) and by Fisher, Lewis & Embleton (1987, pp. 122–125). Similarly, Best & Fisher (1986) use the tangent-normal decomposition (1.1) to obtain both formal and graphical tests of goodness-of-fit of a Watson distribution. See also Fisher, Lewis & Embleton (1987, pp. 168–170).

Tests of the third type have been mainly score tests for submodels of exponential families. These were introduced by Cox (1975) for the von Mises case. This was generalised by Mardia, Holmes & Kent (1984) to a goodness-of-fit test for the von Mises–Fisher family (2.1) based on the score test against a Fisher–Bingham (2.2) alternative. Other goodness-of-fit tests based on the score test are given in Kent (1982). Rivest (1986) modified the score test of a von Mises–Fisher distribution against the alternative of a member of Kent's (1982) family to get a statistic which is asymptotically chi-squared for large samples or large concentrations.

Tests of the fourth type, comparing a fitted density in the specified class to a suitable non-parametric density estimate, have been introduced by Bowman (1989).

7 Robustness and Robust Methods

Outliers in spherical data tend to have little effect on the maximum likelihood estimate of the mean direction of a von Mises–Fisher distribution (2.1) but can have considerable influence on the estimate of the concentration. Influence functions provide a quantitative measure of robustness. Wehrly & Shine (1981) found the influence functions of the mean and median directions of circular distributions and concluded that the former is quite robust. Watson (1986a) calculated the influence functions of the mean direction and the dominant eigenvector of the scatter matrix for distributions on S^{p-1} with densities of the form (2.6). He pointed out that, for sufficiently concentrated distributions, the latter estimator can be more robust than the mean direction. Ko & Guttorp (1988) considered *standardised* influence functions of estimators for spherical distributions. They showed that, for a very wide class of families of distributions on S^{p-1}, the mean direction has infinite standardised gross error sensitivity, i.e. the asymptotic effect of a small contamination can be very large compared with the dispersion. The maximum likelihood estimator of κ in the von Mises–Fisher family (2.1) also has infinite standardised gross error sensitivity.

Robust estimation of the mean direction of a von Mises–Fisher distribution has been based mainly on two approaches: the median direction; and the dominant eigenvector of the scatter matrix. The circular median was introduced by Mardia (1972, pp. 28,31) and considered by R. Edwards in an unpublished Ph.D. thesis. Some asymptotic properties were investigated by Ducharme & Milasevic (1987b). Fisher (1985) introduced the spherical median $\mathbf{m}$ of a distribution on S^{p-1} as the point minimising the expected arc-length to a random point $\mathbf{x}$, that is the point $\mathbf{a}$ in S^{p-1} minimising $E[\cos^{-1}(\mathbf{x}^{\mathrm{T}}\mathbf{a})]$. He investigated properties of the sample version and proposed associated test procedures. In contrast to Fisher's intrinsic approach, the embedding approach used by Ducharme & Milasevic (1987a) considers $\tilde{\boldsymbol{\mu}}$ defined by $\tilde{\boldsymbol{\mu}} = \|\tilde{\mathbf{m}}\|^{-1}\tilde{\mathbf{m}}$, where $\tilde{\mathbf{m}}$ is the spatial median in R^p, that is the point $\tilde{\mathbf{m}}$ in R^p minimising $E(\|\mathbf{x}-\mathbf{a}\|)$. A natural generalisation of the sample median is given by M-estimators of the mean direction $\boldsymbol{\mu}$. These estimate $\boldsymbol{\mu}$ on the basis of observations $\mathbf{x}_1, \ldots, \mathbf{x}_n$ on S^{p-1} by minimising

$$\sum_{i=1}^{n} \rho(t(\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\mu}; \kappa)),$$

where $\rho(.)$ and $t(.)$ are suitable functions (and if κ is unknown it is replaced by a suitable robust estimate). On the circle, such M-estimators were considered by Lenth (1981a), using $t(u;\kappa) = \operatorname{sign}(u)\{2\kappa(1-u)\}^{\frac{1}{2}}$ and taking ρ to be Huber's (1964) function defined by

$$\rho(t) = \begin{cases} t^2/2 & |t| \leqslant c, \\ c\,|t| - c^2/2 & |t| > c, \end{cases}$$

or Andrews' (1974) function

$$\rho(t) = \begin{cases} -c^2 \cos(t/c) & |t| \leqslant c\pi, \\ 0 & |t| > c\pi, \end{cases}$$

for a suitable value of c.

In view of (3.5), inference about the concentration parameter κ of a Fisher distribution is closely connected with inference about the mean of an exponential distribution. Thus familiar robust methods can be transferred to the spherical context. In particular, Fisher (1982b) suggested the use of the L-estimator

$$\hat{\kappa}_R^{-1} = \sum_{i=1}^{n} l_i c_{(i)}$$

of κ^{-1}, where the $c_{(i)}$ are order statistics of $1 - \mathbf{x}_i^{\mathrm{T}}\hat{\boldsymbol{\mu}}$ and l_i are suitable weights. He also considered a Winsorised version which was modified by Kimber (1985) to reduce its bias. A robust method of testing the equality of the concentration parameters of several von Mises–Fisher (2.1) or Watson (2.9) distributions was suggested by Fisher (1986a,b). His test is based on the ($\kappa \to \infty$) asymptotic $(p-1)$-variate normality of the tangential part $t\boldsymbol{\xi}$ of $\mathbf{x}$ (see (3.4)) and is carried out as a one-way analysis of variance on the absolute values of the components of the sample analogues $\mathbf{x}_i - (\mathbf{x}_i^{\mathrm{T}}\hat{\boldsymbol{\mu}})\hat{\boldsymbol{\mu}}$ of $t\boldsymbol{\xi}$.

8 Simulation

Distributions on the circle can easily be simulated by applying the wrapping approach to distributions on the line. In particular, the uniform distribution is obtained by reducing mod 2π the uniform distribution on $[0, 2\pi]$. For simulating a von Mises distribution ((2.1) with $p = 2$), Best & Fisher (1979) gave a simple and efficient algorithm which uses the acceptance-rejection method with a wrapped Cauchy envelope.

Most simulation methods on spheres exploit the symmetry of the distribution considered. In particular, the uniform distribution on S^{p-1} is easily simulated by using its invariance under rotation. If $\mathbf{z}$ has any rotationally invariant distribution, such as a $N(0, \mathbf{I}_p)$ distribution, then $\|\mathbf{z}\|^{-1}\mathbf{z}$ is uniformly distributed on S^{p-1}. Faster methods have been given by Sibuya (1962), Marsaglia (1972) and Tashiro (1977). See also Watson (1983a, pp. 68–70).

For axially symmetric distributions simulation can be based on the fact that, in the tangent-normal decomposition (2.7), $\boldsymbol{\xi}$ is uniformly distributed on S^{p-2}. The Fisher case ((2.1) with $p = 3$) is particularly tractable as $1 - t$ has a truncated exponential distribution (where $t = \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}$ is the 'normal part' of $\mathbf{x}$). Details are given in Fisher, Lewis & Willcox (1981) and Fisher, Lewis & Embleton (1987, p. 59). For simulation of the Watson distribution (2.9) see Best & Fisher (1986) and Fisher, Lewis & Embleton (1987, p. 59). In the general axially symmetric case Ulrich (1984) used envelope acceptance-rejection to generate the 'normal part' t. Simulation of Fisher–Bingham distributions (2.2) is considered in Wood (1987).

It is interesting to note that two well-established techniques for simulation of ordinary real-valued variates are based on properties of directional distributions. Firstly, one way of simulating the uniform distribution on $[0, 1]$ is to use the random number generator of

Wichmann & Hill (1982). This is an application of the 'central limit theorem' on the circle (Mardia, 1972, p. 88), which states that if $\theta_1, \ldots, \theta_n$ are identically distributed angles with a non-lattice distribution on S^1, then $\theta_1 + \ldots + \theta_n \pmod{2\pi}$ has a limiting uniform distribution on $[0, 2\pi]$, as $n \to \infty$. Secondly, the Box–Muller algorithm for generating normal variates relies on the uniform distribution on the circle.

9 Nonparametric and Bootstrap Methods

9.1 Nonparametric Methods

Because it is not always appropriate to use one of the standard parametric models discussed in § 2, various nonparametric methods for directional statistics have been developed. These can conveniently be divided into methods for one-dimensional sample spaces (the circle and the line), methods for higher dimensional spheres (and general Riemannian manifolds), and methods based on large-sample or high-concentration approximations. For a more detailed survey of nonparametric methods in directional statistics up to 1984 see Rao Jammalamadaka (1984).

One-dimensional sample spaces. The one-dimensional sample spaces, the circle and the line, are unique in that they can be ordered in a natural way (unique up to sign). This ordering enables distribution functions and rank statistics to be defined. Thus adaptations of the usual nonparametric procedures are available for one-sample tests, two-sample tests, and tests of independence.

(i) One-sample tests can be obtained from tests of uniformity, as described at the beginning of § 6.2.

(ii) Distribution-free two-sample tests on the circle can be derived from rotation-invariant tests of uniformity by means of uniform scores. Given two samples $(x_1, \ldots, x_m)$ and $(y_1, \ldots, y_n)$ of angles on the unit circle, *uniform scores* are defined as follows. Choose an orientation and an origin of the circle and arrange the combined sample of $m + n$ observations in numerical order. Let $(r_1, \ldots, r_m)$ denote the ranks of the first sample in the combined sample. Then the uniform scores of the first sample are the angles $(2\pi r_1[m+n]^{-1}, \ldots, 2\pi r_m[m+n]^{-1})$. Testing the uniform scores for uniformity provides the associated two-sample test. For example, the Rayleigh test of uniformity yields the two-sample test of Wheeler & Watson (1964).

Distribution-free two-sample tests on the circle can also be derived from ordinary two-sample rank tests on the line. Given a rank test T on the line, the corresponding test statistic for circular data is

$$\max T(r_1, \ldots, r_n),$$

the maximum being over all choices of origin and orientation of the circle. The circular test obtained in this way from the Mann–Whitney test was introduced by Batschelet (1965) and studied by Eplett (1979, 1982).

Another class of distribution-free two-sample tests on the circle consists of those based on the number of observations from one sample which fall in the gap between two specified observations from the other sample; see Rao & Mardia (1980). Persson (1979) has shown that Watson's (1962) two-sample U^2 statistic has this form and can be regarded as a circular analogue of the Mann–Whitney statistic.

(iii) Distribution-free tests of independence of circular variables can be obtained from suitable correlation coefficients by means of uniform scores. Relabel the observations

$(x_1, y_1), \ldots, (x_n, y_n)$ so that the ranks of the x_i and the corresponding y_i are $1, \ldots, n$ and $r_1, \ldots, r_n$ respectively. Applying any rotation-invariant correlation coefficient to the pairs $(i2\pi n^{-1}, r_i 2\pi n^{-1})$ yields a distribution-free test of independence. In particular, Jupp & Spurr (1985) showed that various previously-known correlation coefficients for circular data can be obtained by applying the general Sobolev tests of independence described below to the uniform scores. A similar construction works in the linear-circular case. For example, Mardia's (1976a) nonparametric linear-circular correlation coefficient results from taking $t(x) = x$ and $\mathbf{u}(\theta) = (\cos\theta, \sin\theta)$ for $(x, \theta) \in R \times S^1$ and using the uncorrected covariance matrix S_{12}^* instead of S_{12} in (9.1).

Another class of circular correlation coefficients obtained from uniform scores consists of those based on

$$\bar{R}_{\pm}^2 = \left\| n^{-1} \sum_{i=1}^{n} t([i \mp r_i] 2\pi n^{-1}) \right\|^2,$$

for some suitable mapping $t: S^1 \to L^2(S^1)$. Mardia (1975a, p. 360) introduced such correlation coefficients by considering $\max(\bar{R}_+^2, \bar{R}_-^2)$, whereas Fisher & Lee (1982) suggested the signed correlation coefficient $\bar{R}_+^2 - \bar{R}_-^2$ as a circular analogue of Spearman's rho.

Signed correlation coefficients of a different kind were introduced by Fisher & Lee (1981, 1982). These analogues of Kendall's tau are U-statistics based (in the bivariate circular case) on the number of triples (x_i, y_i), (x_j, y_j), (x_k, y_k) of pairs of angles for which the ordering of (x_i, x_j, x_k) is the same as that of (y_i, y_j, y_k).

Higher-dimensional sample spaces. For general spheres (and other manifolds) the only distribution-free procedures developed so far are permutation tests based on the infinite-dimensional embedding approach of § 3.1. Wellner (1979) introduced two-sample 'Sobolev tests' which measure the 'distance' between the samples $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ by

$$\left\| m^{-1} \sum_{i=1}^{m} \mathbf{t}(\mathbf{x}_i) - n^{-1} \sum_{j=1}^{n} \mathbf{t}(\mathbf{y}_j) \right\|^2,$$

cf. (3.3). In the same spirit, Jupp & Spurr (1983) proposed tests of symmetry under the action of a group G. Their test statistics have the form

$$\left\| m^{-1} \sum_{i=1}^{m} \left\{ \mathbf{t}(\mathbf{x}_i) - \int \mathbf{t}(g\mathbf{x}_i)\, dg \right\} \right\|^2,$$

where $g\mathbf{x}_i$ denotes the transform of $\mathbf{x}_i$ by the element g of G. The same approach was used by Jupp & Spurr (1985) to provide tests of independence. If $\mathbf{t}$ and $\mathbf{u}$ are mappings of the type described in § 3.1 then the sample correlation coefficients defined by

$$n \operatorname{tr}(\mathbf{S}_{12}\mathbf{S}_{12}^{\mathrm{T}}), \tag{9.1}$$

where $\mathbf{S}_{12}$ denotes the (infinite-dimensional) sample covariance matrix of $\mathbf{t}(\mathbf{x})$ and $\mathbf{u}(\mathbf{y})$, can be used to test independence of the spherical random variables $\mathbf{x}$ and $\mathbf{y}$.

Many of the tests on the circle given above have the typical 'non-parametric' behaviour of being unchanged by small changes of the data. For higher dimensional spheres a correlation coefficient and a two-sample test with this property were proposed by Jupp (1987).

Approximate methods. Approximately distribution-free methods for large samples can be obtained by applying the central limit theorem to first and second sample moments and by exploiting simplifications due to any assumed symmetry. See Watson (1983a, pp. 136–140) and Prentice (1984, 1986).

9.2 Bootstrap Methods

For small-sample problems in which there are no grounds for adopting a parametric model, traditional nonparametric methods are increasingly being replaced by bootstrapping. Watson (1983e) and Fisher, Lewis & Embleton (1987) have called attention to the use of bootstrap methods for directional data. A simulation study comparing bootstrap confidence cones for the mean direction $\boldsymbol{\mu}$ of a spherical distribution with various parametric competitors was given by Ducharme et al. (1985). By use of asymptotically pivotal statistics, Fisher & Hall (1989a) obtained a variety of bootstrap confidence regions for $\boldsymbol{\mu}$ with good coverage properties.

10 Density Estimation

One way of obtaining a grasp of the message given by a data set is to estimate the underlying density and to produce a corresponding contour plot.

Kernel density estimation on the sphere using a von Mises–Fisher kernel has been discussed by Watson (1983a, pp. 9, 37; 1983e) and implemented with a computer program by Diggle & Fisher (1985). Hall, Watson & Cabrera (1987) consider two classes of kernel density estimators on S^{p-1}. Given observations $\mathbf{x}_1, \ldots, \mathbf{x}_n$, these have the forms

$$\hat{f}_K(\mathbf{x}; \kappa) = n^{-1} c_0(\kappa) \sum_{i=1}^{n} K(\kappa \mathbf{x}^{\mathrm{T}} \mathbf{x}_i), \tag{10.1}$$

$$\hat{f}_L(\mathbf{x}; \kappa) = n^{-1} d_0(\kappa) \sum_{i=1}^{n} L(\kappa(1 - \mathbf{x}^{\mathrm{T}} \mathbf{x}_i)), \tag{10.2}$$

where $K(.)$ and $L(.)$ are known kernel functions and κ is an unknown concentration parameter. The kernel most used in practice is the von Mises–Fisher kernel which belongs to both these classes and is obtained by taking $K(t) = e^t$ or $L(t) = e^{-t}$. The authors investigated the bias and variance of these estimators, as well as their expected squared-error and Kullback–Leibler losses. They showed that all estimators of the form (10.1) are asymptotically equivalent for large samples and they gave an asymptotically optimal kernel of the type (10.2).

A different construction of density estimators was given by Hendriks (1987) for quite general Riemannian manifolds. It is based on the infinite-dimensional embedding approach of § 3.1. For $T = 1, 2, \ldots$, let $\mathbf{t}_T$ denote the function $\mathbf{t}$ from S^{p-1} into $L^2(S^{p-1})$ given by (3.2) with $a_k = 1$ for $k < T$ and $a_k = 0$ for $k \geqslant T$. Then these density estimators $\hat{f}_T(.)$ are defined by

$$\hat{f}_T(.) = n^{-1} \sum_{i=1}^{n} \mathbf{t}_T(\mathbf{x}_i),$$

or equivalently by

$$\hat{f}_T(\mathbf{x}) = n^{-1} \sum_{i=1}^{n} \langle \mathbf{t}_T(\mathbf{x}), \mathbf{t}_T(\mathbf{x}_i) \rangle$$

where $\langle\, ,\, \rangle$ denotes the inner-product on $L^2(S^{p-1})$. Hendricks gave bounds in terms of T for the expected squared-error and supremum-error losses of these estimators.

11 Curve Fitting and Smoothing

Because of its geophysical importance, regression of a spherical variable on a scalar predictor (such as time) has been the subject of much attention. Perhaps the simplest parametric family of regression functions for regressing a spherical variable on a scalar (or an unrestricted vector) is

$$\mathbf{f}(t; \mathbf{A}, c) = \mathbf{A}(\cos ct, \sin ct, \mathbf{0})^{\mathrm{T}}, \tag{11.1}$$

where the parameter $\mathbf{A}$ is a rotation matrix and c is the constant speed of movement along the corresponding great circle path. The fundamental problem with fitting this family is that the residual sum of squares does not attain its lower bound as the speed of wrapping increases. This is seen most dramatically in regression of an angle θ in S^1 on 'time' t. Let angles θ_i be observed at times t_i. As the angles θ_i and $\theta_i + 2k\pi$ are equivalent, one can get an arbitrarily close fit to the data by increasing the speed. Consequently, almost all the work on curve fitting and smoothing for directional data has been nonparametric.

Apart from a 'kernel-type' smoothing algorithm given by Watson (1985), most of the methods proposed for nonparametric spherical regression have been concerned with constructing suitable sphere-valued analogues of cubic splines in the plane. Two of these are due to Fisher & Lewis (1985). One of their proposals is to use piecewise smooth functions $\mathbf{f}(.)$ satisfying $\|\mathbf{f}^{(3)}(t)\| = \text{constant}$ (by analogy with $\mathbf{f}^{(3)}(t) = \text{constant}$ for cubic splines in R^3) as well as $\|\mathbf{f}(t)\| = 1$. The other uses portions of loxodromes (curves of constant 'compass bearing' with respect to some pole). Jupp & Kent (1987) took an intrinsic geometrical approach to the construction of smoothing splines on the sphere. For fitting data $\{(t_i, \mathbf{v}_i) : 1 \leqslant i \leqslant n\}$ with $\|\mathbf{v}_i\| = 1$, they suggested minimising

$$\sum_{i=1}^{n} w_i \cos^{-1}\{\mathbf{f}(t_i)^{\mathrm{T}}\mathbf{v}_i\} + \lambda \int \{[\mathbf{I} - \mathbf{f}(t)\mathbf{f}(t)^{\mathrm{T}}]\mathbf{f}^{(2)}(t)\}^2 \, dt \tag{11.2}$$

(by analogy with minimising

$$\sum_{i=1}^{n} w_i \|\mathbf{f}(t_i) - \mathbf{v}_i\|^2 + \lambda \int \{\mathbf{f}^{(2)}(t)\}^2 \, dt \tag{11.3}$$

for cubic splines in R^3). Here $w_1, \ldots, w_n$ are weights and λ is a smoothing parameter. The first term in (11.2) and (11.3) penalizes lack of fit; the second term penalizes curvature of the fitted path. As a more easily implemented alternative, they use an extension of the wrapping approach of § 1.2 to 'unwrap' the sphere onto the plane (using a given path rather than a given point), thus converting a spherical problem into a standard planar one. They also comment on several techniques used previously.

For nonparametric regression of a rotation matrix on a scalar Prentice (1987) has taken the wrapping approach and used the exponential map to transform splines in R^3 to 'splines' in so(3). Nonparametric regression of an angular variate on a planar variate can be interpreted as smoothing a unit vector field (a field of directions at points on the plane). Mendoza (1986) has shown how thin plate splines can be used for this purpose.

Another variation on the theme of splines was given by Watson's (1985) construction of interpolating curves in the circular case. His elegant response to the problem of fitting (11.1) is to use curves minimising

$$\int \{\mathbf{f}''(t)\}^2 \, dt + \nu \int \{\mathbf{f}'(t)\}^2 \, dt,$$

for an appropriate value of the parameter ν, thus penalising speed as well as curvature.

12 Non-Standard Applications

Applications of directional statistics to various problems in the natural sciences (in particular, astronomy, biology, and earth sciences) have now become standard. There have also been some interesting non-trivial applications in other areas. For example, the circular structural model

$$\mathbf{x}_i = \mathbf{a} + \rho(\cos \tau_i, \sin \tau_i)^{\mathrm{T}} + \boldsymbol{\varepsilon}_i \quad (i = 1, \ldots, n)$$

in which $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$ are independent isotropic bivariate normal errors, has been used in the case of uniformly distributed τ_i by Mardia & Holmes (1980) and Anderson (1981) to model the positions $\mathbf{x}_i$ of stones in megalithic stone 'circles'. Berman (1983) has pointed out that the case in which $\tau_i = \theta_0 + \theta_i$ with $\theta_1, \ldots, \theta_n$ known is a linear model. Berman & Griffiths (1985) considered the model

$$\mathbf{x}_i = \mathbf{a} + R_i(\cos \zeta_i, \sin \zeta_i)^{\mathrm{T}} \quad (i = 1, \ldots, n)$$

in which R_i and ζ_i are independent with respectively a gamma and a von Mises distribution with concentration κ and mean direction τ_i.

Cruz-Orive et al. (1985) applied directional statistics to the stereology of anisotropic structures (such as capillaries in skeletal muscle) by considering the case in which the line elements in the phase of interest have a Watson distribution (2.9).

Various problems arising outside the directional context can usefully be recast in directional form. For example, Kent, Briden & Mardia (1983) considered the problem of finding linear segments among the means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n$ of a sequence of multivariate random variables. One approach to this is to note that collinearity of the $\boldsymbol{\mu}_i$ is equivalent to equality of the unit vectors $\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$, $1 \leqslant i, j \leqslant n$.

13 Informal Methods and Exploratory Data Analysis

Important developments have occurred in exploratory data analysis and informal graphical methods for directional data. Exploratory data analysis of spherical data is largely based on:

(i) plots of suitable projections of either the raw data or of contour or shade plots of estimated densities,
(ii) eigenvalues of the scatter matrix.

See Watson (1983a, pp. 27–39) and Fisher, Lewis & Embleton (1987, pp. 35–49).

Probability plots for spherical data based on (3.5) and (3.6) were introduced by Lewis & Fisher (1982). These are

(i) 'colatitude plots' of $1 - \cos \theta'_{(i)}$ against $-\log \{1 - (i - \frac{1}{2})n^{-1}\}$,
(ii) 'longitude plots' of $\phi'_{(i)}$ against $(i - \frac{1}{2})n^{-1}$,
(iii) plots of the ordered values of $\phi''_i(\sin \theta''_i)^{\frac{1}{2}}$ against standard normal quantiles,

where n is the sample size, (θ', ϕ') denote spherical polar coordinates with the sample mean direction as pole, and (θ'', ϕ'') are as in (3.6). For the Fisher family these plots should be approximately linear with slopes approximately κ^{-1}, 1, and $\kappa^{-\frac{1}{2}}$ respectively and should pass near the origin. Thus they provide graphical tests of goodness of fit, quick estimates of the concentration parameter, and a method of detecting outliers. Similar plots for the Watson family (2.9) were given by Best & Fisher (1986). See also Fisher, Lewis & Embleton (1987, pp. 114–115, 117–125, 164–165, 170–175, 182–183, 204, 208).

Residuals for directional data have been constructed by Jupp (1988) and Rivest (1989a) using the wrapping approach of § 1.2. When points $\hat{\mathbf{y}}_1, \ldots, \hat{\mathbf{y}}_n$ are fitted to observations

$\mathbf{y}_1, \ldots, \mathbf{y}_n$ on S^{p-1}, the corresponding crude residuals are the tangent vectors $\mathbf{y}_i - (\mathbf{y}_i^{\mathrm{T}}\hat{\mathbf{y}}_i)\hat{\mathbf{y}}_i$ to S^{p-1} at $\hat{\mathbf{y}}_i$.

14 Non-Spherical Sample Spaces

In the discussion so far we have concentrated on the circular and spherical cases but other manifolds do occur as sample spaces in directional statistics. Perhaps the most important of these from the practical point of view are the rotation groups so(p) of R^p. Most of the known results on rotation groups generalize readily to Stiefel and Grassmann manifolds. The Stiefel manifold $V_{r,p}$ of orthonormal r-frames in p-space can be regarded as the set of $r \times p$ matrices $\mathbf{X}$ satisfying $\mathbf{X}\mathbf{X}^{\mathrm{T}} = \boldsymbol{I}_p$. The Grassmann manifold $G_{r,p}$ of r-dimensional subspaces of p-space can be considered as the set of $p \times p$ matrices $\mathbf{X}$ of rank r satisfying $\mathbf{X} = \mathbf{X}^{\mathrm{T}} = \mathbf{X}\mathbf{X}^{\mathrm{T}}$. In particular, $V_{1,p} = S^{p-1}$, $V_{p,p} = O(p)$ and $G_{1,p} = RP^{p-1}$, $(p-1)$-dimensional real projective space. As RP^{p-1} is obtained from S^{p-1} by identifying antipodal points $\mathbf{x}$ and $-\mathbf{x}$, axial data can be regarded as data on RP^{p-1}. There is a useful function from $V_{r,p}$ to $G_{r,p}$ given by $\mathbf{X} \mapsto \mathbf{X}^{\mathrm{T}}\mathbf{X}$ which associates with each frame of r vectors the subspace it spans. Observations on Stiefel manifolds include the specifications of elliptical cometary orbits by their perihelion and normal directions (Mardia, 1975a; Jupp & Mardia, 1979) and similar 'orbits' in vectorcardiography (Downs, 1972; Prentice, 1986).

The basic family of distributions on Stiefel manifolds consists of the matrix Fisher distributions introduced by Khatri & Mardia (1977) with probability density functions

$$f(\mathbf{X}; \mathbf{A}) = a(\mathbf{A})^{-1} \exp\{\operatorname{tr}(\mathbf{X}^{\mathrm{T}}\mathbf{A})\}, \tag{14.1}$$

where $\mathbf{A}$ is an $r \times p$ matrix parameter. This is an exponential family with canonical statistic the inclusion of $V_{r,p}$ into the vector space of all $r \times p$ matrices. Taking this inclusion as the function $\mathbf{t}$ in (3.2) yields the test of uniformity on a Stiefel manifold given by Mardia & Khatri (1977). An analogue on $V_{r,p}$ of the Bingham family (2.8) is the exponential family introduced by Prentice (1982) with canonical statistic the composite of the above function from $V_{r,p}$ to $G_{r,p}$ with the inclusion of $G_{r,p}$ into the vector space of all symmetric $p \times p$ matrices. The probability density functions are

$$f(\mathbf{X}; \mathbf{A}) = a(\mathbf{A})^{-1} \exp\{\operatorname{tr}(\mathbf{A}\mathbf{X} \otimes \mathbf{X}^{\mathrm{T}})\}, \tag{14.2}$$

where $\mathbf{A}$ is a symmetric $rp \times rp$ matrix parameter. This family extends the matrix Bingham family of Khatri & Mardia (1977). Explicit maximum likelihood estimation in the families (14.1) and (14.2) is considered by Jupp & Mardia (1979) and Prentice (1982). Prentice proposes subfamilies of (14.2) with suitable axial symmetry as models for the distribution of X-shaped objects and considers a matrix version of the Fisher–Bingham family (2.2) as a model for T-shaped objects.

If the Euclidean inner product on R^p is replaced by the indefinite symmetric bilinear form $*$ given by

$$\mathbf{x} * \mathbf{y} = x_1 y_1 - x_2 y_2 - \ldots - x_p y_p$$

for $\mathbf{x} = (x_1, \ldots, x_p)$, $\mathbf{y} = (y_1, \ldots, y_p)$, the analogue of S^{p-1} is the unit hyperboloid H_p in R^p defined by $H_p = \{\mathbf{x} \in R^p : \mathbf{x} * \mathbf{x} = 1, x_1 > 0\}$. Since the function $(x_1, \ldots, x_p) \mapsto (x_1 x_2, \ldots, x_1 x_p)$ identifies H_p with R^{p-1} and takes the distinguished point $(1, 0, \ldots, 0)$ to the origin, distributions in H_p can be useful for modelling random $(p-1)$-vectors, especially where zero has a special role, e.g. (for $p = 3$) the wind directions and strengths considered by Jensen (1981). Barndorff-Nielsen (1978) introduced a family of distributions on H_p analogous to the von Mises–Fisher family (2.1). The probability density functions have the form

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = a(\kappa)^{-1} \exp\{\kappa \boldsymbol{\mu} * \mathbf{x}\}$$

with respect to a measure invariant under the group of transformations preserving $*$. The parameters κ and $\boldsymbol{\mu}$ satisfy $\kappa > 0$ and $\boldsymbol{\mu} \in H_p$. The properties of these distributions have been investigated by Jensen (1981).

Transformations of sample spaces play a useful role. Stephens (1982) has exploited the mapping $(x_1, \ldots, x_p) \mapsto (x_1^{\frac{1}{2}}, \ldots, x_p^{\frac{1}{2}})$ of the standard $(p-1)$-simplex into S^{p-1} to transform compositional data into directional data. See also Mardia (1976b). The following transformation of rotational data into axial data on S^3 is one of the principal tools for handling rotations. There is a canonical 2 to 1 map of S^3 onto so(3) given in coordinate terms by

$$\pm\mathbf{u} \mapsto \mathbf{M}(\mathbf{u}) = \begin{pmatrix} u_1^2 + u_2^2 - u_3^2 - u_4^2 & -2(u_1u_4 - u_2u_3) & 2(u_1u_3 + u_2u_4) \\ 2(u_1u_4 + u_2u_3) & u_1^2 + u_3^2 - u_2^2 - u_4^2 & -2(u_1u_2 - u_3u_4) \\ -2(u_1u_3 - u_2u_4) & 2(u_1u_2 + u_3u_4) & u_1^2 + u_4^2 - u_2^2 - u_3^2 \end{pmatrix}, \quad (14.3)$$

where $\mathbf{u} = (u_1, \ldots, u_4) \in S^3$. Since $\mathbf{M}(-\mathbf{u}) = \mathbf{M}(\mathbf{u})$, the function sending $\mathbf{M}(\mathbf{u})$ to the axis given by $\pm\mathbf{u}$ transforms rotational data into axial data on S^3. This transformation has been exploited by Moran (1975) and Prentice (1978, 1986, 1987). Prentice (1986) gave a neat description of (14.3) in terms of quaternions and showed that $\mathbf{M}(\mathbf{x})$ has a matrix Fisher distribution if and only if $\mathbf{x}$ has a Bingham distribution on S^3.

In a deep study inspired originally by alignments of archaeological features Kendall (1984) has considered the *shapes* of k-tuples of points in R^p. Two k-tuples are considered to have the same shape if they can be transformed into one another by translations, rotations and changes of scale. The sets Σ_2^k of shapes of k-tuples of points in the plane are of particular interest. The manifold Σ_2^k can be identified with complex projective space CP^{k-1}. In particular, the set of shapes of (labelled) triangles in the plane can be identified with the sphere S^2, for example by associating to the triangle ABC the point $(1+r^2)^{-1}(1-r^2, 2r\cos\phi, 2r\sin\phi)$ of S^2, where the complex number $2^{-1}|AB|3^{\frac{1}{2}}re^{i\phi}$ represents the line joining the midpoint of AB to C. Thus probability distributions of the vertices of a triangle give rise to distributions on S^2. Further work on shapes of triangles has appeared in Mardia, Edwards & Puri (1977), Mardia (1980) and Kendall (1985). Mardia (1989a,b) has put forward the thesis that one can simply use the statistics for the Fisher distribution to analyse Kendall's shape sphere for triangles. See also Mardia & Dryden (1989). Watson (1983g, 1986b) pointed out that a sequence $\{\mathbf{C}_n : n = 1, 2, \ldots\}$ of independent identically-distributed 3×3 matrices gives rise to a Markov process on the set of triangles in the plane, in which the nth transition takes the triangle $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ to the triangle $(\mathbf{C}_n\mathbf{x}_1, \mathbf{C}_n\mathbf{x}_2, \mathbf{C}_n\mathbf{x}_3)$. In the case of circulant matrices, he obtained the limiting distribution of the corresponding shapes.

A very large class of distributions on the p-dimensional torus $T^p = (S^1)^p$ can be obtained from spectral densities of stationary random fields on the lattice Z^p. If a spectral density is positive and integrable, it can be normalized to give a probability density on T^p. It follows from Mardia's (1988a) results on the spectral densities of scalar-valued symmetric stationary Gaussian Markov random fields that the corresponding probability densities have the form

$$f(\boldsymbol{\theta}) = a(\boldsymbol{\beta})\left\{1 - \sum_{\mathbf{n}} \beta_{\mathbf{n}} \cos(\mathbf{n}^{\mathrm{T}}\boldsymbol{\theta})\right\}^{-1} \quad \boldsymbol{\theta} \in T^p,$$

where $\mathbf{n}$ runs through Z^p and

$$E(x_{\mathbf{r}} \mid x_{\mathbf{s}}\ \ \mathbf{s} \neq \mathbf{r}) = \mu + \sum_{\mathbf{n}} (\beta_{\mathbf{n}} - \mu) x_{\mathbf{n}}.$$

15 Closing Remarks

There are various areas of statistics other than those discussed above for which a directional version has been developed. For example, work has been done on Bayesian analysis (Mardia & El-Atoum, 1976; Bagchi & Guttman, 1988), design of experiments (Laycock, 1975) and sequential analysis (Gadsden & Kanji, 1980, 1982). Other areas such as spherical time series analysis, higher order asymptotics and the use of sample spaces other than spheres and rotation groups have seen comparatively little activity for directional data. Two notable gaps are the paucity of theory outside the usual framework of independent identically distributed observations and the lack of a commercial computer package for the analysis of directional data. The former is serious because in geology, for example, observed directions are hardly ever randomly selected; see C.R. Rao (1975). Future practical problems should provide the stimulus for further worthwhile developments in the topic.

Bibliographical note

The references below are not restricted to those discussed in the text but are intended to include all the papers in the statistical literature on directional statistics in the period 1975–1988.

Acknowledgements

In preparing this paper we have been assisted by Kikuchi's (1982) list of abstracts of articles on directional statistics in the period 1972–1981. During the final stages of writing, Professor Mardia received support from NSF grant no. DMS 8803207 to Professor G.S. Watson, Princeton University. We are grateful to the referees for their helpful comments.

References

Abeysasekara, S. & Collett, D. (1982). On the estimation of the parameters of the von Mises distribution. *Comm. Statist.* A **11,** 2083–2090.

Accardi, L., Cabrera, J. & Watson, G.S. (1987). Some stationary Markov processes in discrete time for unit vectors. *Metron* **45,** 115–133.

Ackermann, H. (1985). Distribution-free tolerance regions for circular data. *EDV Med. Biol.* **16,** 97–99.

Anderson, D.A. (1981). The circular structural model. *J. R. Statist. Soc. B* **43,** 131–141.

Andrews, D.F. (1974). A robust method for multiple linear regression. *Technometrics* **16,** 523–531.

Arnaiz Tovar, G. & Ruiz-Rivas, C. (1986) Outliers in circular data: a Bayesian approach. *Questiio* **10,** 1–6.

Arsham, H. (1988). Kuiper's *P*-value as a measuring tool and decision procedure for the goodness-of-fit test. *J. Appl. Statist.* **15,** 131–135.

Baba, Y. (1981). Statistics of angular data. Wrapped normal distribution model (In Japanese). *Proc. Inst. Statist. Math.* **28,** 41–54.

Bagchi, P. & Guttman, I. (1988). Theoretical considerations of the multivariate von Mises–Fisher distribution. *J. Appl. Statist.* **15,** 149–169.

Bai, Z.D., Rao, C.R. & Zhao, L.C. (1988). Kernel estimators of density function of directional data. *J. Mult. Anal.* **27,** 24–39.

Barndorff-Nielsen, O.E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.* **5,** 151–157.

Barndorff-Nielsen, O.E., Blœsild, P., Jensen, J.L. & Jørgensen, B. (1982). Exponential transformation models. *Proc. R. Soc. Lond.* A **379,** 41–65.

Bartels, R. (1984). Estimation in a bidirectional mixture of von Mises distributions. *Biometrics* **40,** 777–784.

Batschelet, E. (1965). *Statistical Methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms.* Washington: Am. Inst. Bio. Sciences.

Batschelet, E. (1981). *Circular Statistics in Biology.* London: Academic Press.

Beran, R. (1968). Testing for uniformity on a homogeneous space. *J. Appl. Prob.* **5,** 177–195.

Beran, R. (1979). Exponential models for directional data. *Ann. Statist.* **7,** 1162–1178.

Berman, M. (1983). Estimating the parameters of a circle when angular differences are known. *Appl. Statist.* **32,** 1–6.

Berman, M. & Griffiths, D. (1985). Incorporating angular information into models for stone circle data. *Appl. Statist.* **34,** 237–245.

Berman, M. & Culpin, D. (1986). The statistical behaviour of some least squares estimators of the centre and radius of a circle. *J.R. Statist. Soc.* B **48,** 183–196.

Best, D.J. & Fisher, N.I. (1979). Efficient simulation of the von Mises distribution. *Appl. Statist.* **24,** 152–157.
Best, D.J. & Fisher, N.I. (1981). The bias of the maximum likelihood estimators of the von Mises–Fisher concentration parameters. *Comm. Statist.* B **10,** 493–502.
Best, D.J. & Fisher, N.I. (1986). Goodness of fit and discordancy tests for samples from the Watson distribution on the sphere. *Aust. J. Statist.* **28,** 13–31.
Bingham, C. (1974). An antipodally symmetric distribution on the sphere. *Ann. Statist.* **2,** 1201–1225.
Bingham, C. (1980). Distribution on the sphere. In *R.A. Fisher: An Appreciation, Lecture Notes in Statistics I,* Ed. S. Fienberg et al., pp. 171–181. New York: Springer-Verlag.
Bingham, C. & Chang, T. (1988). The use of the Bingham distribution in spherical regression inference. Technical report 514, School of Statistics, Univ. of Minnesota.
Bingham, C. & Mardia, K.V. (1978). A small circle distribution on the sphere. *Biometrika* **65,** 379–389.
Bingham, M.S. (1978). A characterization of the uniform distribution on the circle in the analysis of directional data. *J. Appl. Prob.* **15,** 852–857.
Bingham, M.S. & Mardia, K.V. (1975). Maximum likelihood characterization of the von Mises distribution. In *Statistical Distributions in Scientific Work,* **3,** Ed. G.P. Patil, S. Kotz and J.K. Ord, pp. 387–398. Dordrecht: Reidel.
Blœsild, P. (1979). Conditioning with conic sections in the two-dimensional normal distribution. *Ann. Statist.* **7,** 659–670.
Bowman, A.W. (1989). Density based tests of distributional shape. Manuscript.
Byth, K. (1982). On kernel methods of estimating marginal radial and angular probability density functions. *Biom. J.* **24,** 49–58.
Chang, T. (1986). Spherical regression. *Ann. Statist.* **14,** 907–924.
Chang, T. (1987). On the statistical properties of estimated rotations. *J. Geophys. Res.* **92,** B7, 6319–6329.
Chang, T. (1988). Estimating the relative rotation of two tectonic plates from boundary crossings. *J. Am. Statist. Assoc.* **83,** 1178–1183.
Chang, T. (1989). Spherical regression with errors in variables. *Ann. Statist.* **17,** 293–306.
Chou, R.J. & Hwang, T.C. (1985). Correlation between unit vectors. *Chinese J. Math.* **13,** 137–152.
Clark, R.M. (1983). Estimation of parameters in the marginal Fisher distribution. *Aust. J. Statist.* **25,** 227–237.
Clark, R.M. (1985). A FORTRAN program for constrained sequence-slotting based on minimum combined path length. *Comp. Geosci.* **11,** 605–617.
Clark, R.M. (1988). An evaluation by simulation of alternative estimators for the marginal Fisher distribution. *J. Appl. Statist.* **15,** 235–246.
Clark, R.M. & Morrison, B.J. (1983). A Normal approximation to the Fisher distribution. *Aust. J. Statist.* **25,** 96–104.
Collett, D. (1980). Outliers in circular data. *Appl. Statist.* **29,** 50–57.
Collett, D. & Lewis, T. (1981). Discriminating between the von Mises and wrapped Normal distributions. *Aust. J. Statist.* **23,** 73–79.
Cox, D.R. (1975). Discussion of paper by K.V. Mardia. *J.R. Statist. Soc.* B **37,** 380–381.
Cressie, N. (1977). On some properties of the scan statistic on the circle and the line. *J. Appl. Prob.* **14,** 272–283.
Cruz-Orive, L.M., Hoppeler, H., Mathieu, O. & Weibel, E.R. (1985). Stereological analysis of anisotropic structures using directional statistics. *Appl. Statist.* **34,** 14–32.
Darling, D.A. (1983). On the asymptotic distribution of Watson's statistic. *Ann. Statist.* **11,** 1263–1266.
Daubenspeck, J.A. & Ogden, R.D. (1978). Estimation of response slopes in respiratory control using directional statistics. *J. Appl. Physiol.* **45,** 823–829.
Degerine, S. (1977). Tests on the dispersion parameter in von Mises distributions. In *Recent Developments in Statistics, Proc. EMS, Grenoble 1976,* Ed. J.R. Barra et al., pp. 403–408. Amsterdam: North Holland.
Degerine, S. (1978). Sur la complétion des structures de von Mises. *C.R. Acad. Sci. Paris* A **287,** 29–31.
Degerine, S. (1979a). Lois de von Mises et lois liées. *Ann. Inst. Henri Poincaré* B **15,** 63–77.
Degerine, S. (1979b). Tests optimaux sur les paramètres des lois de von Mises. *Ann. Inst. Henri Poincaré* B **15,** 375–392.
de Waal, D.J. (1979). On the normalizing constant for the Bingham–von Mises–Fisher matrix distribution. *S. Afr. J. Statist.* **13,** 103–112.
Diggle, P.J. & Fisher, N.I. (1985). SPHERE A contouring program for spherical data. *Comp. Geosci.* **11,** 725–766.
Diggle, P.J., Fisher, N.I. & Lee, A.J. (1985). A comparison of tests of uniformity for spherical data. *Aust. J. Statist.* **27,** 53–69.
Dobson, A.J. (1978). Simple approximations to the von Mises concentration statistic. *Appl. Statist.* **27,** 345–346.
Downs, T.D. (1972). Orientation statistics. *Biometrika* **59,** 665–676.
Downs, T.D. (1974). Rotational angular correlation. In *Biorhythms and Human Reproduction,* Ed. M. Ferin, F. Halberg and L. van der Wiele, pp. 97–104. New York: Wiley.
Ducharme, G.R., Jhun, M., Romano, J.P. & Truong, K.N. (1985). Bootstrap confidence cones for directional data. *Biometrika* **72,** 637–645.
Ducharme, G.R. & Milasevic, P. (1987a). Spatial median and directional data. *Biometrika* **74,** 212–215.
Ducharme, G.R. & Milasevic, P. (1987b). Some asymptotic properties of the circular median. *Comm. Statist.* A **16,** 659–664.

Dudley, R.M., Perkins, P.C., & Giné, M.E. (1975). Statistical tests for preferred orientation. *J. Geol.* **83,** 685–705.
Dufour, J.M. & Roy, R. (1976). On spectral estimation for a homogeneous random process on the circle. *Stoch. Proc. Appl.* **4,** 107–120.
Eben, K. (1983). Classification into two von Mises distributions with unknown mean. *Aplikace Matematiky* **28,** 230–237.
Eplett, W.J.R. (1979). The small sample distribution of a Mann–Whitney type statistic for circular data. *Ann. Statist.* **7,** 446–453.
Eplett, W.J.R. (1982). Two Mann–Whitney type rank tests. *J.R. Statist. Soc.* B **44.** 270–286.
Epp, R.J., Tukey, J.W. & Watson, G.S. (1971). Testing unit vectors for correlation. *J. Geophys. Res.* **76,** 8480–8483.
Fisher, N.I. (1982a). Goodness-of-fit and outlier detection procedures for samples from Fisher's distribution on the sphere. *Contemporary Math.* **9,** 377–380.
Fisher, N.I. (1982b). Robust estimation of the concentration parameter of Fisher's distribution on the sphere. *Appl. Statist.* **31,** 152–154.
Fisher, N.I. (1985). Spherical medians. *J.R. Statist. Soc.* B **47,** 342–348.
Fisher, N.I. (1986a). Robust comparison of dispersions for samples of directional data. *Aust. J. Statist.* **28,** 213–219.
Fisher, N.I. (1986b). Robust tests for comparing the dispersions of several Fisher or Watson distributions on the sphere. *Geophys. J.R. Astr. Soc.* **85,** 563–572.
Fisher, N.I. (1987). Problems with the current definition of circular standard deviation of wind direction. *J. Climate Appl. Meteorol.* **26,** 1522–1529.
Fisher, N.I. & Best, D.J. (1984). Goodness-of-fit tests for Fisher's distribution on the sphere. *Aust. J. Statist.* **26,** 142–150.
Fisher, N.I. & Hall, P. (1989a). Bootstrap confidence regions for directional data. *J. Am. Statist. Assoc.* To appear.
Fisher, N.I. & Hall, P. (1989b). New statistical methods for directional data I: bootstrap comparison of mean directions and the fold test in palaeomagnetism. *Geophys. J.R. Astr. Soc.* To appear.
Fisher, N.I., Huntingdon, J.F., Jackett, D.R., Willcox, M.E. & Creasey, J.W. (1985). Spatial analysis of two-dimensional orientation data. *J. Math. Geol.* **17,** 177–194.
Fisher, N.I. & Lee, A.J. (1981). Non-parametric measures of angular-linear association. *Biometrika* **68,** 629–636.
Fisher, N.I. & Lee, A.J. (1982). Non-parametric measures of angular-angular association. *Biometrika* **69,** 315–321.
Fisher, N.I. & Lee, A.J. (1983). A correlation coefficient for circular data. *Biometrika* **70,** 327–332.
Fisher, N.I. & Lee, A.J. (1986). Correlation coefficients for random variables on a sphere or hypersphere. *Biometrika* **73,** 159–164.
Fisher, N.I. & Lewis, T. (1983). Estimating the common mean direction of several circular or spherical distributions with differing dispersions. *Biometrika* **70,** 333–341.
Fisher, N.I. & Lewis, T. (1985). A note on spherical splines. *J.R. Statist. Soc.* B **47,** 482–488.
Fisher, N.I., Lewis, T. & Embleton, B.J.J. (1987). *Statistical Analysis of Spherical Data.* Cambridge University Press.
Fisher, N.I., Lewis, T. & Willcox, M.E. (1981). Tests of discordancy for samples from Fisher's distribution on the sphere. *Appl. Statist.* **30,** 230–237.
Fisher, N.I. & Powell, C.McA. (1989). Statistical analysis of two-dimensional palaeocurrent data: methods and examples. *Aust. J. Earth Sciences* **36,** 91–107.
Fisher, N.I. & Willcox, M.E. (1978). A useful decomposition of the resultant length for samples from the von Mises–Fisher distributions. *Comm. Statist.* B **7,** 257–267.
Fraser, D.A.S., Guttmann, I. & Styan, G.P.H. (1976). Serial correlation and distributions on the sphere. *Comm. Statist.* A **1,** 97–118.
Fraser, M.D., Hsu, Y.-S. & Walker, J.J. (1981). Identifiability of finite mixtures of von Mises distributions. *Ann. Statist.* **9,** 1130–1131.
Fraser, D.A.S. & Massam, H. (1988). Location inference on spheres and cylinders. *J. Statist. Plan. Inf.* **18,** 195–201.
Gadsden, R.J. & Kanji, G.K. (1980). Sequential analysis for angular data. *Statistician* **30,** 119–129.
Gadsden, R.J. & Kanji, G.K. (1982). Sequential analysis applied to circular data. *Sequential Anal.* **1,** 305–314.
Gates, D.J. & Westcott, M. (1980). Further bounds for the distribution of minimum interpoint distance on a sphere. *Biometrika* **67,** 466–469.
Gidskehaug, A. (1976). Statistics on a sphere. *Geophys. J.* **45,** 657–676.
Giné, M.E. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. *Ann. Statist.* **3,** 1243–1266.
Gordon, L. & Hudson, M. (1977). A characterization of the von Mises distribution. *Ann. Statist.* **5,** 813–814.
Gupta, D. (1976). Some mixtures of von Mises distributions. *S. Afr. Statist. J.* **10,** 69–76.
Gupta, D. (1977). Some mixtures of von Mises distributions II. *S. Afr. Statist. J.* **11,** 181–186.
Guttorp, P. & Lockhart, R.A. (1989). Finding the location of a signal: a Bayesian analysis. *J. Am. Statist. Assoc.* **83,** 322–330.

Hall, P., Watson, G.S. & Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika* **74,** 751–762.

Harrison, D., Kanji, G.K. & Gadsden, R.J. (1986). Analysis of variance for circular data. *J. Appl. Statist.* **13,** 123–138.

Harrison, D. & Kanji, G.K. (1988). The development of analysis of variance for circular data. *J. Appl. Statist.* **15,** 197–224.

Hartman, P. & Watson, G.S. (1974). "Normal" distribution functions on spheres & the modified Bessel functions. *Ann. Prob.* **2,** 593–607.

Harvey, P.K. & Ferguson, C.C. (1976). On testing orientation data for goodness of fit to a von Mises distribution. *Comp. Geosci.* **2,** 261–268.

Hendriks, H. (1987). Non-parametric estimation of a probability density on a Riemannian manifold using Fourier expansions. Research report, Dept. of Maths., Catholic Univ., Nijmegen.

Hermans, M. & Rasson, J.P. (1985). A new Sobolev test for uniformity on the circle. *Biometrika* **72,** 698–702.

Hill, G.W. (1976). New approximations to the von Mises distribution. *Biometrika* **63,** 673–676.

Hill, G.W. (1977). Algorithm 518: Incomplete Bessel function $I_0(x)$; the von Mises distribution [S 14]. *ACM Trans. Math. Software* **3,** 279–284.

Hill, G.W. (1981). Evaluation and inversion of the ratios of modified Bessel functions $I_1(x)/I_0(x)$ and $I_{1\cdot5}(x)/I_{0\cdot5}(x)$. *ACM Trans. Math. Software* **7,** 199–208.

Holmquist, B. (1985). Correlation measures for circular data. In *Contributions to Probability and Statistics in Honour of Gunnar Blom,* Ed. J. Lanke and G. Lindgren, pp. 157–168. Dept. Math. Statist., Lund University.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35,** 73–101.

Ibero, M. (1975). Déviation de la loi empirique d'une variable bidimensionnelle. *C.R. Acad Sci. Paris* **A281,** 1059–1062.

Ibero, M. (1976). Un test d'ajustement bidimensionnel independent de la loi. *C.R. Acad. Sci. Paris* **A283,** 387–390.

Jammalamadaka, S.R. (= J.S. Rao) & Sarma, Y.R. (1988). A correlation coefficient for angular data. In *Statistical Theory and Data Analysis II, Proc. Second Pacific Area Stat. Conference, Tokyo*; Ed. K. Matusita, pp. 349–364. Amsterdam: Elsevier (North Holland).

Jensen, J.L. (1981). On the hyperboloid distribution. *Scand. J. Statist.* **8,** 193–206.

Jirina, M. (1978). A biased roulette. *Ann. Inst. Henri Poincaré* B **14,** 1–23.

Johnson, R.A. & Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *J.R. Statist. Soc.* B **39,** 222–229.

Johnson, R.A. & Wehrly, T. (1978). Some angular–linear distributions and related regression models. *J. Am. Statist. Assoc.* **73,** 602–606.

Jones, C.E. (1983). A note on the use of directional statistics in weighted Euclidean distance multidimensional scaling. *Psychometrika* **48,** 473–476.

Jørgensen, B. (1987). Small dispersion asymptotics. *Revista Brasileira de Probabilidade e Estatistica* **1,** 59–90.

Jupp, P.E. (1984). A Poincaré limit theorem for wrapped probability distributions on compact symmetric spaces. *Math. Proc. Camb. Phil. Soc.* **95,** 329–334.

Jupp, P.E. (1987). A non-parametric correlation coefficient and a two-sample test for random vectors or directions. *Biometrika* **74,** 887–890.

Jupp, P.E. (1988). Residuals for directional data. *J. Appl. Statist.* **15,** 137–147.

Jupp, P.E. & Kent, J.T. (1987). Fitting smooth paths to spherical data. *Appl. Statist.* **36,** 34–46.

Jupp, P.E. & Mardia, K.V. (1979). Maximum likelihood estimation for the matrix von Mises–Fisher and Bingham distributions. *Ann. Statist.* **7,** 599–606.

Jupp, P.E. & Mardia, K.V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika* **67,** 163–173.

Jupp, P.E. & Spurr, B.D. (1983). Sobolev tests for symmetry of directional data. *Ann. Statist.* **11,** 1225–1231.

Jupp, P.E. & Spurr, B.D. (1985). Sobolev tests for independence of directions. *Ann. Statist.* **13,** 1140–1155.

Jupp, P.E. & Spurr, B.D. (1989). Statistical estimation of a shock centre: Slate Islands astrobleme. *J. Math. Geol.* **21,** 191–198.

Jupp, P.E., Spurr, B.D., Nichols, G.J. & Hirst, J.P.P. (1987). Statistical estimation of the apex of a sedimentary distributary system. *J. Math. Geol.* **19,** 319–333.

Kanatani, K. (1984). Distribution of directional data and fabric tensors. *Int. J. Eng. Sci.* **22,** 149–164.

Keilson, J., Petrondas, D., Sumita, U. & Wellner, J. (1983). Significance points for some tests of uniformity on the sphere. *J. Statist. Computat. Simul.* **17,** 159–172.

Kelker, D.G. & Langenberg, C.W. (1982). A mathematical model for orientation data from macroscopic conical folds. *J. Math. Geol.* **14,** 289–307.

Kelker, D.G. & Langenberg, C.W. (1987). A mathematical model for orientation data from macroscopic elliptical conical folds. *J. Math. Geol.* **19,** 729–743.

Kendall, D.G. (1974a). Hunting quanta. *Phil. Trans. R. Soc.* A **376,** 231–266.

Kendall, D.G. (1974b). Pole-seeking Brownian motion and bird navigation (with discussion). *J.R. Statist. Soc.* B **36,** 365–417.

Kendall, D.G. (1984). Shape manifolds, Procrustean metrics, and complex projective space. *Bull. Lond. Math. Soc.* **16,** 81–121.

Kendall, D.G. (1985). Exact distributions for shapes of random triangles in convex sets. *Adv. Appl. Prob.* **17**, 308–329.

Kendall, D.G. & Young, G.A. (1984). Indirectional statistics and the significance of an asymmetry discovered by Birch. *Mon. Not. R. Astr. Soc.* **207**, 637–647.

Kent, J.T. (1977). The infinite divisibility of the von Mises–Fisher distribution for all values of the parameter in all dimensions. *Proc. Lond. Math. Soc.* **35**, 359–384.

Kent, J.T. (1978a). Limiting behaviour of the von Mises–Fisher distribution. *Math. Proc. Camb. Phil. Soc.* **84**, 531–536.

Kent, J.T. (1978b). Some probabilistic properties of Bessel functions. *Ann. Prob.* **6**, 760–770.

Kent, J.T. (1982). The Fisher–Bingham distribution on the sphere. *J.R. Statist. Soc.* B **44**, 71–80.

Kent, J.T. (1983a). Identifiability of finite mixtures for directional data. *Ann. Statist.* **11**, 954–988.

Kent, J.T. (1983b). Information gain and a general measure of correlation. *Biometrika* **70**, 163–173.

Kent, J.T. (1987). Asymptotic expansions for the Bingham distribution. *Appl. Statist.* **36**, 139–144.

Kent, J.T., Briden, J.C. & Mardia, K.V. (1983). Linear and planar structure in ordered multivariate data as applied to progressive demagnetization of palaeomagnetic remanence. *Geophys. J.R. Astr. Soc.* **75**, 593–621.

Kent, J.T., Mardia, K.V., & Rao, J.S. (1979). A characterization of the uniform distribution on the circle. *Ann. Statist.* **7**, 882–889.

Kent, J.T. & Tyler, D.E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Statist.* **15**, 247–254.

Khan, A.H. & Yaqub, M. (1980). Distribution of a distance function. *Ann. Inst. Statist. Math.* **32**, 247–253.

Khatri, C.G. & Mardia, K.V. (1977). The von Mises–Fisher distribution in orientation statistics. *J.R. Statist. Soc.* B **39**, 95–106.

Kikuchi, D.A. (1982). Directional data abstracts: 1972–1981. Technical Report 265, Dept. of Statistics, Ohio State Univ., Columbus.

Kimber, A.C. (1985). A note on the detection and accommodation of outliers relative to Fisher's distribution on the sphere. *Appl. Statist.* **34**, 169–172.

Ko, D. & Guttorp, P. (1988). Robustness of estimators for directional data. *Ann. Statist.* **16**, 609–618.

Kokic, P.N. (1987). On tests of uniformity for randomly distributed arcs on a circle. *Aust. J. Statist.* **29**, 179–187.

Lawson, A. (1988a). On tests for spatial trend in a non-homogeneous Poisson process. *J. Appl. Statist.* **15**, 225–234.

Lawson, A. (1988b). Fitting the von Mises distribution using GLIM. *J. Appl. Statist.* **15**, 255–260.

Laycock, P.J. (1975). Optimal design: regression models for directions. *Biometrika* **62**, 305–311.

Lee, J.C. (1980). A circular sign test for uniformity. In *Recent Developments in Statistical Inference and Data Analysis*: *Proceedings of the International Conference in Statistics in Tokyo,* Ed. K. Matusita, pp. 191–195.

Lenth, R.V. (1981a). Robust measures of location for directional data. *Technometrics* **23**, 77–81.

Lenth, R.V. (1981b). On finding the source of a signal. *Technometrics* **23**, 149–154.

Lewis, T. (1975). Probability functions which are proportional to characteristic functions and the infinite divisibility of the von Mises distribution. In *Perspectives in Probability and Statistics,* Ed. J. Gani, pp. 19–28. Sheffield: Applied Probability Trust.

Lewis, T. (1976). On the infinite divisibility of the von Mises distribution. *J. Aust. Math. Soc.* **22**, 332–342.

Lewis, T. & Fisher, N.I. (1982). Graphical methods for investigating the fit of a Fisher distribution to spherical data. *Geophys. J.R. Astr. Soc.* **69**, 1–13.

Liddell, I.G. & Ord, J.K. (1978). Linear-circular correlation coefficients: some further results. *Biometrika* **65**, 448–450.

Lo, J.T.-H. (1977). Exponential Fourier densities and optimal estimation and detection on the circle. *IEEE Trans. Info. Theory* **IT-23**, 110–116.

Lo, J.T.-H. & Eshleman, L.R. (1977). Exponential Fourier densities on S^2 and optimal estimation and detection for directional processes. *IEEE Trans. Info. Theory* **IT-23**, 321–336.

Lo, J.T.-H. & Eshleman, L.R. (1979a). Exponential Fourier densities and optimal estimation for axial processes. *IEEE Trans. Info. Theory* **IT-25**, 463–470.

Lo, J.T.-H. & Eshleman, L.R. (1979b). Exponential Fourier densities on SO(3) and optimal estimation and detection for rotational processes. *SIAM J. Appl. Math.* **36**, 73–82.

Lockhart, R.A. & Stephens, M.A. (1985). Tests of fit for the von Mises distribution. *Biometrika* **72**, 647–652.

Lwin, T. (1975). On von Mises directions. *Ann. Inst. Statist. Math.* **27**, 79–85.

Mackenzie, J.K. (1957). The estimation of an orientation relationship. *Ann. Cryst.* **10**, 61–62.

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

Mardia, K.V. (1975a). Statistics of directional data (with discussion). *J.R. Statist. Soc.* B **37**, 349–393.

Mardia, K.V. (1975b). Distribution theory for the von Mises–Fisher distribution and its application. In *Statistical Distributions in Scientific Work,* **1**, Ed. G.P. Patil, S. Kotz & J.K. Ord, pp. 113–130. Dordrecht: Reidel.

Mardia, K.V. (1975c). Characterizations of directional distributions. In *Statistical Distributions in Scientific Work,* **3**, Ed. G.P. Patil, S. Kotz and J.K. Ord, pp. 365–385. Dordrecht: Reidel.

Mardia, K.V. (1976a). Linear-circular correlation coefficients and rhythmometry. *Biometrika* **63**, 403–405.

Mardia, K.V. (1976b). Discussion of paper by V. Barnett. *J.R. Statist. Soc.* A **139**, 346–347.

Mardia, K.V. (1977). Distributions on Stiefel and Grassman manifolds and their applications. *Adv. Appl. Prob.* **9,** 435–436.
Mardia, K.V. (1979). Discussion of paper by A.P. Dawid. *J.R. Statist. Soc.* B **41,** 25.
Mardia, K.V. (1980). Discussion of paper by S. Broadbent. *J.R. Statist. Soc.* A **143,** 137.
Mardia, K.V. (1981a). Directional statistics in geosciences. *Comm. Statist.* A **10,** 1523–1543.
Mardia, K.V. (1981b). Recent directional distributions with applications. In *Statistical Distributions in Scientific Work,* Ed. C. Taillie et al., pp. 1–19. Dordrecht: Reidel.
Mardia, K.V. (1981c). The evolution of directional models in geosciences since Fisher. In *Down-to-Earth Statistics,* Ed. D.F. Merriam, pp. 39–46. Syracuse University.
Mardia, K.V. (1988a). Multi-dimensional multivariate Gaussian Markov random fields. *J. Mult. Anal.* **24,** 265–284.
Mardia, K.V. (1988b). Directional data analysis: an overview. *J. Appl. Statist.* **15,** 115–122.
Mardia, K.V. (1989a). Some contributions to shape analysis. Discussion of paper by D.G. Kendall. *Statist. Sci.* **4,** 108–111.
Mardia, K.V. (1989b). Shape analysis of triangles through directional techniques. *J.R. Statist. Soc.* B **51**. To appear.
Mardia, K.V., Bogle, S. & Edwards, R. (1983). The statistics of response slope. *J. Appl. Physiol.* **54,** 309–313; **55,** 1044.
Mardia, K.V. & Dryden, I. (1989). The statistical analysis of shape data. *Biometrika* **76,** 271–281.
Mardia, K.V. & Edwards, R. (1982). Weighted distributions and rotating caps. *Biometrika* **69,** 323–330.
Mardia, K.V., Edwards, R. & Puri, M.L. (1977). Analysis of central place theory. *Bull. Int. Statist. Inst.* **47,** 93–110.
Mardia, K.V. & El-Atoum, S.A.M. (1976). Bayesian inference for the von Mises–Fisher distribution. *Biometrika* **63,** 203–205.
Mardia, K.V. & Gadsden, R.J. (1977). A circle of best fit for spherical data and areas of vulcanism. *Appl. Statist.* **26,** 238–245.
Mardia, K.V. & Holmes, D. (1980). A statistical analysis of megalithic data under elliptic pattern. *J.R. Statist. Soc.* A **143,** 293–302.
Mardia, K.V., Holmes, D. & Kent, J.T. (1984). A goodness-of-fit test for the von Mises–Fisher distribution. *J.R. Statist. Soc.* B **46,** 72–78.
Mardia, K.V. & Khatri, C.G. (1977). Uniform distribution on a Stiefel manifold. *J. Mult. Anal.* **7,** 468–473.
Mardia, K.V. & Puri, M.L. (1978). A robust spherical correlation coefficient against scale. *Biometrika* **65,** 391–396.
Mardia, K.V. & Sutton, T.W. (1975). On the modes of a mixture of two von Mises distributions. *Biometrika* **62,** 699–701.
Mardia, K.V. & Sutton, T.W. (1978). A model for cylindrical variables with applications. *J.R. Statist. Soc.* B **40,** 229–233.
Mardia, K.V. & Zemroch, P.J. (1975a). Algorithm AS 80: Spherical statistics. *Appl. Statist.* **24,** 144–146.
Mardia, K.V. & Zemroch, P.J. (1975b). Algorithm AS 81: Circular statistics. *Appl. Statist.* **24,** 147–150.
Mardia, K.V. & Zemroch, P.J. (1975c). Algorithm AS 86: The von Mises distribution function. *Appl. Statist.* **24,** 268–272.
Mardia, K.V. & Zemroch, P.J. (1977). Table of maximum likelihood estimates for the Bingham distribution. *J. Statist. Comp. Simul.* **6,** 29–34.
Marsaglia, G. (1972). Choosing a point from the surface of a sphere. *Ann. Math. Statist.* **43,** 645–646.
McFadden, P.L. (1980a). The best estimate of Fisher's precision parameter *k*. *Geophys. J.R. Astr. Soc.* **60,** 397–407.
McFadden, P.L. (1980b). Determination of the angle in a Fisher distribution which will be exceeded with a given probability. *Geophys. J.R. Astr. Soc.* **60,** 391–396.
McFadden, P.L. (1980c). Simple graphical methods for estimating the confidence region about the orientation of the intersection of two planes: Discussion. *Can. J. Earth Sci.* **17,** 1111–1113.
McFadden, P.L. & Lowes, F.J. (1981). The discrimination of mean directions drawn from Fisher distributions. *Geophys. J.R. Astr. Soc.* **67,** 19–33.
Mendoza, C.E. (1986). Smoothing unit vector fields. *J. Math. Geol.* **18,** 307–322.
Mirvaliev, M. (1978). Rejection of outlying results of angular measurements. *Theory Prob. Appl.* **23,** 814–819.
Moffitt, S. (1978). Discriminant-analysis for directional data (Abstract). *Biometrics* **34,** 162–163.
Moore, B.F. (1980). A modification of Rayleigh's test for vector data. *Biometrika* **67,** 175–180.
Moran, P.A.P. (1975). Quaternions, Haar measure and the estimation of a palaeomagnetic rotation. In *Perspectives in Probability and Statistics,* Ed. J. Gani, pp. 295–301. Sheffield: Applied Probability Trust.
Moran, P.A.P. (1979). The closest pair of *N* random points on a sphere. *Biometrika* **66,** 158–162.
Morris, J.E. & Laycock, P.J. (1974). Discriminant analysis of directional data. *Biometrika* **61,** 335–341.
Nikulin, V.N. (1986). Stability of the characterization of the von Mises distribution (in Russian). In *Stability Problems for Stochastic Models, Varna, 1985,* pp. 73–84. Moscow: Vsesoyuz. Nauch.-Inst. Sistem. Issled.
Onstott, T.C. (1980). Applications of the Bingham distribution function in palaeomagnetic studies. *J. Geophys. Res.* **85**(B3), 1500–1510.
Papakonstantinou, V. (1984). The distribution of the length of *N* random unit vectors for a von Mises population. *Statist. Prob. Letters* **2,** 111–115.

Pelz, W. & Good, I.J. (1976). Approximating the lower tail-areas of the Kolmogorov–Smirnov one-sample statistic. *J.R. Statistic. Soc.* B **38**, 152–156.

Persson, T. (1979). A new way to obtain Watson's U^2. *Scand. J. Statist.* **6**, 119–122.

Prentice, M.J. (1978). On invariant tests of uniformity for directions and orientations. *Ann. Statist.* **6**, 169–176.

Prentice, M.J. (1982). Antipodally symmetric distributions for orientation statistics. *J. Statist. Plan. Inf.* **6**, 205–214.

Prentice, M.J. (1984). A distribution-free method of interval estimation for unsigned directional data. *Biometrika* **52,** 147–154.

Prentice, M.J. (1986). Orientation statistics without parametric assumptions. *J.R. Statist. Soc.* B **48**, 214–222.

Prentice, M.J. (1987). Fitting smooth paths to rotation data. *Appl. Statist.* **36,** 325–331.

Prentice, M.J. (1988). Using sample principal axes to estimate small circles and ellipses on the sphere. *J. Appl. Statist.* **15,** 171–177.

Prentice, M.J. (1989). Spherical regression on matched pairs of orientation statistics. *J.R. Statist. Soc.* B **51,** 241–248.

Puri, M.L. & Rao, J.S. (1977). Problems of association for bivariate circular data and a new test of independence. In *Multivariate Analysis IV,* Ed. P.R. Krishnaiah, pp. 513–522. Amsterdam: North Holland.

Rao, C.R. (1975). Some problems of sample surveys. *Suppl. Appl. Prob.* **7,** 50–61.

Rao, J.S. (1976). Some tests based on arc length for the circle. *Sankhyā* B **38,** 651–661.

Rao Jammalamadaka, S. (1984). Nonparametric methods in directional data analysis. In *Handbook of Statistics,* **4**: *Nonparametric Methods,* Ed. P.R. Krishnaiah and P.K. Sen, pp. 755–770. Amsterdam: Elsevier.

Rao, J.S. & Mardia, K.V. (1980). Pitman efficiencies of some two-sample nonparametric tests. In *Recent Developments in Statistical Inference and Data Analysis*: *Proceedings of the International Conference in Statistics in Tokyo,* Ed. K. Matusita, pp. 248–254. Amsterdam: North Holland.

Rao, J.S. & Yoon, Y. (1983). Comparison of the limiting efficiencies of two chi-square type tests for the circle. *J. Ind. Statist. Soc.* **21,** 19–26.

Rivest, L.-P. (1982). Some statistical methods for bivariate circular data. *J.R. Statist. Soc.* B **44,** 81–90.

Rivest, L.-P. (1984a). Symmetric distributions for dependent unit vectors. *Ann. Statist.* **12,** 1050–1057.

Rivest, L.-P. (1984b). On the information matrix for symmetric distributions on the hypersphere. *Ann. Statist.* **12,** 1085–1089.

Rivest, L.-P. (1986). Modified Kent's statistics for testing the goodness-of-fit of the Fisher distribution in small concentrated samples. *Statist. Prob. Letters* **4,** 1–4.

Rivest, L.-P. (1988). The bivariate Fisher–von Mises distribution. *Comm. Statist.* A **17,** 461–483.

Rivest, L.-P. (1989a). Spherical regression for concentrated Fisher–von Mises distributions. *Ann. Statist.* **17,** 307–317.

Rivest, L.-P. (1989b). Small sample inference for spherical regression. *Ann. Statist.* To appear.

Roberts, P.H. & Ursell, H.D. (1960). Random walk on a sphere and on a Riemannian manifold. *Phil. Trans. R. Soc. Lond.* A **252,** 317–356.

Roberts, P.H. & Winch, D.E. (1984). On random rotations. *Adv. Appl. Prob.* **16,** 638–655.

Roy, R. (1973). Estimation of the covariance function of a homogeneous process on the sphere. *Ann. Statist.* **1,** 780–785.

Roy, R. (1976). Spectral analysis for a random process on the sphere. *Ann. Inst. Statist. Math.* **28,** 91–97.

Saw, J.G. (1978). A family of distributions on the m-sphere and some hypothesis tests. *Biometrika* **65,** 69–73.

Saw, J.G. (1983). Dependent unit vectors. *Biometrika* **70,** 665–671.

Saw, J.G. (1984). Ultraspherical polynomials and statistics on the m-sphere. *J. Mult. Anal.* **14,** 105–113.

Schäben, H. (1986). SPHERE A contouring program for spherical data—Comment. *Comp. Geosci.* **12,** 729.

Schatte, P. (1983). On sums modulo 2π of independent random variables. *Math. Nachr.* **110,** 243–262.

Schou, G. (1978). Estimation of the concentration parameter in von Mises–Fisher distributions. *Biometrika* **65,** 369–377.

Selkirk, K.E. (1982). Statistics on a circle. *Teaching Statist.* **4,** 89–92.

Sibuya, M. (1962). A method for generating uniformly distributed points on n-dimensional spheres. *Ann. Inst. Statist. Math.* **14,** 81–85.

Silverman, B.W. (1978). Distances on circles and spheres. *J. Appl. Prob.* **15,** 136–143.

Spurr, B.D. (1981). On estimating the parameters in mixtures of circular normal distributions. *J. Math. Geol.* **13,** 163–173.

Stam, A.J. (1982). Limit theorems for uniform distributions on spheres in high dimensional euclidean spaces. *J. Appl. Prob.* **19,** 221–229.

Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. *J. Am. Statist. Assoc.* **69,** 730–737.

Stephens, M.A. (1975a). A new test for the modal vector of the Fisher distribution. *Biometrika* **62,** 171–174.

Stephens, M.A. (1975b). Axial and bimodal data on a sphere. In *Proc. Conf. at Dalhousie Univ. Halifax, Nova Scotia,* Ed. R.P. Gupta, pp. 311–323. Amsterdam: North Holland.

Stephens, M.A. (1977). Whitworth runs on a circle. *Ann. Inst. Statist. Math.* **29,** 287–294.

Stephens, M.A. (1979). Vector correlation. *Biometrika* **66,** 41–48.

Stephens, M.A. (1982). Use of the von Mises distribution to analyse continuous proportions. *Biometrika* **69,** 197–203.

Tanasi, C. (1982). A problem of random anisotropic walk for a new circular model. *Atti Accad. Sci. Lett. Arti Palermo* **3,** 59–92.

Tashiro, Y. (1977). On methods for generating uniform random points on the surface of a sphere. *Ann. Inst. Statist. Math.* **29,** 295–300.
Tyler, D. (1987). Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika* **74,** 579–589.
Ulrich, G. (1984). Computer generation of distributions on the m-sphere. *Appl. Statist.* **33,** 158–163.
Upton, G.J.G. (1973). Single-sample tests for the von Mises distribution. *Biometrika* **60,** 87–99.
Upton, G.J.G. (1974). New approximations to the distribution of certain angular statistics. *Biometrika* **61,** 369–373.
Upton, G.J.G. (1976). More multisample tests for the von Mises distribution. *J. Am. Statist. Assoc.* **71,** 675–678.
Upton, G.J.G. (1986). Approximate confidence intervals for the mean direction of a von Mises distribution. *Biometrika* **73,** 525–527.
Upton, G.J.G. & Fingleton, B. (1989). *Spatial Data Analysis by Example,* **2,** *Categorical and Directional Data.* Chichester: Wiley.
Watson, G.S. (1956). Analysis of dispersion on a sphere. *Mon. Not. R. Astr. Soc. Geophys. Suppl.* **7,** 153–159.
Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48,** 109–114.
Watson, G.S. (1962). Goodness-of-fit tests on a circle. II. *Biometrika* **49,** 57–63.
Watson, G.S. (1970). Orientation statistics in the earth sciences. *Bull. Geol. Inst. Univ. Uppsala* **2,** 73–89.
Watson, G.S. (1977). Optimal invariant tests for uniformity. In *Studies in Probability and Statistics,* Ed. E.J. Williams, pp. 121–127. Amsterdam: North Holland.
Watson, G.S. (1982a). Distributions on the circle and sphere. In *Essays in Statistical Science,* Ed. J. Gani and E.J. Hannan, pp. 265–280. *J. Appl. Prob. Special Vol.* **19A.**
Watson, G.S. (1982b). The estimation of palaeomagnetic pole position. In *Stochastics and Probability. Essays in Honour of C.R. Rao,* Ed. J. Kallianpur, et al., pp. 703–712. Amsterdam: North Holland.
Watson, G.S. (1982c). Three aspects of the statistics of directions. In *Lecture Notes Contemporary Math.* **9,** 187–205.
Watson, G.S. (1983a). *Statistics on Spheres. University of Arkansas Lecture Notes in the Mathematical Sciences* **6,** New York: Wiley.
Watson, G.S. (1983b). Large sample theory of the Langevin distribution. *J. Statist. Plan. Inf.* **8,** 245–256.
Watson, G.S. (1983c). Large sample theory for distributions on the hypersphere with rotational symmetries. *Ann. Inst. Statist. Math.* **35,** A, 303–319.
Watson, G.S. (1983d). Limit theorems on high dimensional spheres and Stiefel manifolds. In *Studies in Econometrics, Time Series and Multivariate Statistics,* Ed. T. Amemiya, L. Goodman and S. Karlin, pp. 559–570. New York: Academic Press.
Watson, G.S. (1983e). The computer simulation treatment of directional data. In *Proceedings of the Geological Conference, Kharagpur, India, December 1983, Indian J. Earth Sci.,* 19–23.
Watson, G.S. (1983f). Optimal and robust estimation on the sphere. *Bull. 44th Session Int. Statist. Inst.* **2,** 816–818.
Watson, G.S. (1983g). Random triangles. In *Proceedings of the Second International Conference on Stochastic Geometry, Aarhus,* Ed. E.B. Jensen and H.J.G. Gundersen, pp. 183–199. Dept. Theoretical Statistics, University of Aarhus.
Watson, G.S. (1984a). The calculation of confidence regions for eigenvectors. *Aust. J. Statist.* **26,** 272–276.
Watson, G.S. (1984b). The theory of concentrated Langevin distributions. *J. Mult. Anal.* **14,** 74–82.
Watson, G.S. (1985). Interpolation and smoothing of directed and undirected line data. In *Multivariate Analysis VI,* Ed. P.R. Krishnaiah, pp. 613–625. New York: Academic Press.
Watson, G.S. (1986a). Some estimation theory on the sphere. *Ann. Inst. Statist. Math.* **38,** 263–275.
Watson, G.S. (1986b). The shapes of a random sequence of triangles. *Adv. Appl. Prob.* **18,** 156–169.
Watson, G.S. (1988). The Langevin distribution on high dimensional spheres. *J. Appl. Statist.* **15,** 123–130.
Watson, G.S. (1989). The statistics of rotations. In *Probability Measures on Groups IX, Proceedings, Oberwohlfach 1988, Lecture Notes in Mathematics,* Ed. H. Heyer. To appear. Berlin: Springer Verlag.
Watson, G.S. & Williams, E.J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika* **43,** 344–352.
Wehrly, T. & Johnson, R.A. (1980). Bivariate models for dependence of angular observations and a related Markov process. *Biometrika* **67,** 255–256.
Wehrly, T. & Shine, E.P. (1981). Influence curves for directional data. *Biometrika* **68,** 634–635.
Wellner, J.A. (1979). Permutation tests for directional data. *Ann. Statist.* **7,** 929–943.
Wheeler, S. & Watson, G.S. (1964). A distribution-free two-sample test on a circle. *Biometrika* **51,** 256–257.
White, G.C. (1985). Optimal location of towers for triangulation studies using biotelemetry. *J. Wildl. Manag.* **49,** 190–196.
Wichmann, B.A. & Hill, I.D. (1982). Algorithm AS183: An efficient and portable pseudo-random number generator. *Appl. Statist.* **31,** 188–190.
Wilkie, D. (1983). Rayleigh test for randomness of circular data. *Appl. Statist.* **32,** 311–312.
Wood, A. (1982). A bimodal distribution for the sphere. *Appl. Statist.* **31,** 52–58.
Wood, A. (1987). The simulation of spherical distributions in the Fisher–Bingham family. *Comm. Statist.* B **16,** 885–898.
Wood, A. (1988). Some notes on the Fisher–Bingham family on the sphere. *Comm. Statist.* A **17,** 3881–3897.
Yaqub, M. & Khan, A.H. (1980). On the distribution of a distance function on the sphere. *Metrika* **27,** 145–151.

Yfantis, E.A. & Borgman, L.E. (1982). An extension of the von Mises distribution. *Comm. Statist.* A **11,** 1695–1706.
Yoshimura, I. (1978). On a test of homogeneity hypothesis for directional data. *Sankhyā* A **40,** 310–312.
Zar, J.H. (1976). Two-sample and multi-sample testing of circular data. *Behavior Res. Meth. Instr.* **8,** 329–330.

Résumé

De nombreux articles sur la théorie et la pratique de la statistique directionnelle ont paru depuis la revue du sujet par Mardia (1975a). Cet article présente une approche cohérente de ces dévelopments. L'attention est portée sur les rapports entre la théorie sous-jacente et quelques idées clefs: familles exponentielles, modèles de transformation, décompositions tangentes-normales, transformations en problèmes multidimensionnels, et le théorème limite centrale. De plus on identifie trois points de vue fondamentaux de la statistique directionelle, dans lesquels l'espace d'échantillonage de base, la sphère, est considéré respectivement comme étant un sous-ensemble de l'espace euclidien, un objet en lui-même et une entité approximée par une surface tangente. On considère des méthodes paramétriques, non-paramétriques et informelles. La discussion se fonde surtout sur le cas où les observations proviennent du cercle ou de la sphère, mais une section sur les espaces d'échantillonage non-sphériques est incluse.

[*Received October* 1988, *accepted June* 1989]