Statistics of Directional Data

By K. V. MARDIA

Leeds University

[Read before the ROYAL STATISTICAL SOCIETY at a meeting organized by the RESEARCH SECTION on Wednesday, March 19th, 1975, Professor R. L. PLACKETT in the Chair]

SUMMARY

Directional data analysis is emerging as an important area of statistics. Within the past two decades, various new techniques have appeared, mostly to meet the needs of scientific workers dealing with directional data. The paper first introduces the two basic models for the multi-dimensional case known as the von Mises–Fisher distribution and the Bingham distribution. Their sampling distribution theory depends heavily on the isotropic case and some developments are discussed. An optimum property of an important test for the von Mises–Fisher case is established. A non-parametric test is proposed for the hypothesis of independence for observations on a torus. In addition to some numerical examples on the preceding topics, five case studies are given which illuminate the power of this new methodology. The case studies are concerned with cancer research, origins of comets, arrival times of patients, navigational problems and biological rhythms. Some unsolved problems are also indicated.

Keywords: ARRIVAL TIMES; BINGHAM DISTRIBUTION; BIOLOGICAL RHYTHMS; BIRD NAVIGATION; CANCER CELLS; CHARACTERIZATION; CORRELATION ON TORUS; DIRECTIONAL DATA; INDEPENDENCE; ORIGIN OF COMETS; RANDOM WALK; UNIFORM SCORES; VON MISES–FISHER DISTRIBUTION

1. INTRODUCTION

THERE are various statistical problems which arise in the analysis of data when the observations are directions. Directional data are often met in astronomy, biology, geology, medicine and meteorology, such as in investigating the origins of comets, solving bird navigational problems, interpreting paleomagnetic currents, assessing variation in the onset of leukaemia, analysing wind directions, etc.

The directions are regarded as points on the circumference of a circle in two dimensions or on the surface of a sphere in three dimensions. In general, directions may be visualized as points on the surface of a hypersphere but observed directions are obviously angular measurements.

The subject has recently been receiving increasing attention, but the field is as old as the subject of mathematical statistics itself. Indeed, the theory of errors was developed by Gauss primarily to analyse certain directional measurements in astronomy. The breakthrough in the subject is marked by a pioneering paper of R. A. Fisher which appeared only two decades ago (Fisher, 1953); his work was motivated by a paleomagnetic problem posed by a geophysicist, J. Hospers. Since then, thanks mostly to G. S. Watson and M. A. Stephens, the development of the subject has been rapid. It is interesting to note that Karl Pearson was involved with a bird-migration problem leading to the isotropic random walk on a circle as early as 1905. Also, von Mises introduced an important circular distribution in 1918 to study the deviation of atomic weights from integral values. For further historical notes, we refer to Mardia (1972a, pp. xvii–xix).

The main aim of this paper is to cover certain new topics which either illuminate the methodology or raise new data-analytic problems. However, the paper does not aim to give a review of the subject as a whole since this has already been covered by Mardia (1972a) for the circular and the spherical cases. Some new mathematical results are given, but considerations of space have made it necessary to present their proofs in other papers.

Some important distributions on a p-dimensional hypersphere are introduced in Section 2. These include the von Mises–Fisher distribution which has certain characterizations analogous to those of the linear normal distribution. Some other desirable characterizations lead to the Brownian motion distribution on a hypersphere. However, the von Mises–Fisher distribution leads to tractable maximum likelihood estimates and sampling distributions in problems of hypothesis testing, whereas the Brownian motion distribution does not. Therefore, the von Mises–Fisher distribution is widely studied and used by statisticians. We discuss some information theoretic characterizations of basic models, and provide an appropriate model for two correlated vectors each having a von Mises–Fisher distribution. Section 3 gives some important results on sampling distribution theory for the von Mises–Fisher population, which leans heavily on the isotropic random walk on a hypersphere. The method of derivation is briefly outlined. An inference problem related to the population mean direction for the von Mises–Fisher distribution is treated in Section 4. Sections 2–4 have resulted from an attempt to unify some of the parametric work; Watson and Williams (1956) were the first to attempt such a unification for the von Mises–Fisher distribution. Section 5 deals with some problems associated with the Bingham distribution which is appropriate when the data are axial. Section 6 looks into the concept of dependence for two circular variables and proposes a non-parametric test to test for independence. Section 7 gives some additional case studies while Section 8 indicates some unsolved problems and outlines some problems not discussed here.

2. BASIC MODELS

Let $\mathbf{l}^{\mathrm{T}} = (l_1, \ldots, l_p)$ be a unit random vector taking values on the surface of a p-dimensional hypersphere S_p of unit radius and having its centre at the origin. The vector $\mathbf{l}$ can be viewed as a vector of direction cosines. Some important directional distributions are as follows. Their discussion for $p = 2$ and $p = 3$ will be found in Mardia (1972a) and for general p in Mardia (1975a, b).

2.1. *The von Mises–Fisher Distribution*

A random vector $\mathbf{l}$ is said to have a p-variate von Mises–Fisher distribution if its probability density function (p.d.f.) is given by

$$c_p(\kappa)\exp(\kappa\boldsymbol{\mu}^{\mathrm{T}}\mathbf{l}), \quad \kappa > 0, \quad \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu} = 1, \quad \mathbf{l} \in S_p, \tag{2.1}$$

where κ is the concentration parameter, $\boldsymbol{\mu}$ is the population mean-direction vector (i.e. $E(\mathbf{l}) = \rho\boldsymbol{\mu}$, $\rho > 0$), and

$$c_p(\kappa) = \kappa^{\frac{1}{2}p-1}/\{(2\pi)^{\frac{1}{2}p} I_{\frac{1}{2}p-1}(\kappa)\}, \tag{2.2}$$

with $I_r(\kappa)$ denoting the modified Bessel function of the first kind and order r. If $\mathbf{l}$ has its p.d.f. of the form (2.1.), we will say that it is distributed as $M_p(\boldsymbol{\mu}, \kappa)$. For $p = 2$ and $p = 3$, (2.1) reduces to p.d.f.'s of the von Mises and the Fisher distributions respectively. These particular cases explain the nomenclature for $p > 3$ but the distribution (2.1) was first introduced by Watson and Williams (1956).

For $\kappa = 0$, the distribution reduces to the uniform distribution on S_p with p.d.f.

$$f(\boldsymbol{l}) = c_p, \quad \boldsymbol{l} \in S_p, \tag{2.3}$$

where

$$c_p = c_p(0) = \Gamma(\tfrac{1}{2}p)/(2\pi^{\frac{1}{2}p}). \tag{2.4}$$

For $\kappa > 0$, the distribution has a mode at $\boldsymbol{l} = \boldsymbol{\mu}$. The larger the value of κ, the greater is the clustering around the mean-direction vector. For this reason, it is used as a model when the data are suspected to be unimodal.

It is sometimes convenient to consider the density of $\boldsymbol{l}$ in terms of the spherical polar co-ordinates $\boldsymbol{\theta}^{\mathrm{T}} = (\theta_1, \ldots, \theta_{p-1})$ with the help of the transformation

$$\boldsymbol{l} = \mathbf{u}(\boldsymbol{\theta}), \quad 0 < \theta_i \leqslant \pi, \quad i = 1, \ldots, p-2, \quad 0 < \theta_{p-1} \leqslant 2\pi, \tag{2.5}$$

where

$$\mathbf{u}(\boldsymbol{\theta}) = (u_1(\boldsymbol{\theta}), \ldots, u_p(\boldsymbol{\theta}))^{\mathrm{T}},$$

$$u_j(\boldsymbol{\theta}) = \cos\theta_j \prod_{i=0}^{j-1} \sin\theta_i, \quad j = 1, \ldots, p, \quad \sin\theta_0 = \cos\theta_p = 1. \tag{2.6}$$

The density (2.1) then reduces to

$$g(\boldsymbol{\theta};\, \boldsymbol{\mu}_0, \kappa) = c_p(\kappa)\exp\{\kappa \mathbf{u}^{\mathrm{T}}(\boldsymbol{\mu}_0)\,\mathbf{u}(\boldsymbol{\theta})\}\, a_p(\boldsymbol{\theta}), \tag{2.7}$$

where $\boldsymbol{\mu}_0 = (\mu_{0,1}, \ldots, \mu_{0,p-1})^{\mathrm{T}}$ denotes the spherical polar co-ordinates of $\boldsymbol{\mu}$, and

$$a_p(\boldsymbol{\theta}) = \prod_{j=2}^{p-1} \sin^{p-j}\theta_{j-1}. \tag{2.8}$$

For $p = 2$, we will say that θ_1 is distributed as $M(\mu_0, \kappa)$.

2.2. *The Bingham Distribution*

Another important model arises when the observations are axial so the p.d.f. of $\boldsymbol{l}$ satisfies the antipodal-symmetry property

$$f(\boldsymbol{l}) = f(-\boldsymbol{l}).$$

How to construct an appropriate model? First, consider a simple construction of the von Mises–Fisher distribution. Let $\mathbf{x}$ be distributed as $N_p(\boldsymbol{\mu}, \kappa^{-1}\mathbf{I})$ with $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu} = 1$. Then the conditional distribution of $\mathbf{x}$ given $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$ is $M_p(\boldsymbol{\mu}, \kappa)$. Since, for the conditional distribution to be axial, there should be no terms involving $\boldsymbol{\mu}$, the mean vector of the normal distribution, we now assume that $\mathbf{x}$ is $N_p(\mathbf{0}, \boldsymbol{\Sigma})$. Then the conditional distribution of $\mathbf{x}$ given $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$ leads to a p.d.f. of the following form:

$$f(\boldsymbol{l};\, \boldsymbol{\mu}, \boldsymbol{\kappa}) = \text{const} \times \exp\{\mathrm{tr}\,(\boldsymbol{\kappa}\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\mu})\}, \quad \boldsymbol{l} \in S_p, \tag{2.9}$$

where $\boldsymbol{\mu}$ now denotes an orthogonal matrix, $\boldsymbol{\kappa}$ is a diagonal matrix of constants and the normalizing constant depends only on $\boldsymbol{\kappa}$. Since $\mathrm{tr}\,(\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\mu}) = 1$, the sum of the parameters κ_i is arbitrary, and it is usual to take $\kappa_p = 0$. Different values of $\boldsymbol{\kappa}$ in (2.9) give the uniform distribution, symmetric and asymmetric girdle distributions and bimodal distributions. The density for the general case is due to Bingham (1964) who investigated the distribution extensively for statistical applications for $p = 3$. For $p = 2$, the distribution is called the bimodal distribution of von Mises type.

2.3. *Information Theoretic Characterizations*

Various characterizations of directional distributions are discussed in Bingham and Mardia (1975) and Mardia (1975a). One of the most important characterizations of these distributions involves maximizing entropy under certain constraints. The proof is obtained from the following result of Mardia (1975a) which turns out to be an extension of Theorem 13.2.1 of Kagan *et al.* (1973, p. 409).

Suppose that distributions defined over a space S are to be represented by densities relative to some familiar measure such as Lebesgue, Haar, etc. Let $t_1, \ldots, t_q$ represent q given real valued functions over S such that no linear combination of $t_1, \ldots, t_q$ is constant. If, for a density f,

(i) S^* is the support of $f(s)$ where $s \in S^*$, $S^* \subset S$,
(ii) $E\{t_i(s)\} = a_i$, fixed, $i = 1, \ldots, q$, and
(iii) the entropy is maximized,

then the p.d.f. is of the form

$$f(s) = \exp\left\{b_0 + \sum_{i=1}^{q} b_i t_i(s)\right\}, \quad s \in S^*, \tag{2.10}$$

provided there exist $b_0, b_1, \ldots$ such that (2.10) satisfies (i) and (ii). Further, if there exists such a density then it is unique.

Thus the maximum entropy distributions with the fixed mean vector $E(\mathbf{l})$ and fixed "moment of inertia" $E(\mathbf{l}\mathbf{l}^{\mathrm{T}})$ are the von Mises–Fisher and the Bingham distributions respectively. The characterizations for the von Mises distribution and the Fisher distribution were first given by Mardia (1972, pp. 65–66) and Rao (1969, pp. 141–142) respectively. If both $E(\mathbf{l})$ and $E(\mathbf{l}\mathbf{l}^{\mathrm{T}})$ are fixed, we can write down the p.d.f. but it has not so far received any attention.

2.4. *A Generalized von Mises–Fisher Distribution*

We now use this result in constructing a suitable distribution when two unit random vectors $\mathbf{l}_1$ and $\mathbf{l}_2$ are correlated. Obviously, we should specify $E(\mathbf{l}_1)$, $E(\mathbf{l}_2)$, $E(\mathbf{l}_1 \mathbf{l}_2^{\mathrm{T}})$ so that the maximum entropy density from (2.10) is

$$\text{const} \times \exp(\mathbf{a}_1^{\mathrm{T}} \mathbf{l}_1 + \mathbf{a}_2^{\mathrm{T}} \mathbf{l}_2 + \operatorname{tr} \mathbf{A}\, \mathbf{l}_1 \mathbf{l}_2^{\mathrm{T}}), \quad \mathbf{l}_1, \mathbf{l}_2 \in S_p. \tag{2.11}$$

We show that the marginal distributions of $\mathbf{l}_1$ and $\mathbf{l}_2$ are not of the von Mises–Fisher form except for trivial cases. Without any loss of generality, let us take $p = 2$. In this case, the p.d.f. can be written as

$$f(\theta, \phi) = C \exp\{\kappa_1 \cos(\theta - \mu) + \kappa_2 \cos(\phi - \nu) + a \cos\theta \cos\phi$$
$$+ b \sin\theta \cos\phi + c \cos\theta \sin\phi + d \sin\theta \sin\phi\}, \tag{2.12}$$

where $0 < \theta$, $\phi \leqslant 2\pi$. The marginal p.d.f. of ϕ is found to be

$$C \exp\{\kappa_2 \cos(\phi - \nu)\}\, I_0(Q^{\frac{1}{2}}), \tag{2.13}$$

where

$$Q = \kappa_1^2 + 2\kappa_1(a \cos\mu + b \sin\mu) \cos\phi + 2\kappa_1(c \cos\mu + d \sin\mu) \sin\phi$$
$$+ (a^2 + b^2) \cos^2\phi + (c^2 + d^2) \sin^2\phi + (ac + bd) \cos\phi \sin\phi. \tag{2.14}$$

Hence, the p.d.f. of ϕ is of the von Mises form if and only if Q is a constant. For $\kappa_1 \neq 0$, we have (i)

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = K\begin{pmatrix} \sin\mu & -\cos\mu \\ -\sin\mu & \cos\mu \end{pmatrix},$$

where K is a constant and (ii) the rows of $\mathbf{A}$ are orthogonal. Hence, we have $K = 0$. For $K = 0$, the variables θ and ϕ are independently distributed. If $\kappa_1 = 0$ then either $\mathbf{A} = \mathbf{0}$ or const. $\times\, \mathbf{A}$ is orthogonal, and so θ is not von Mises although it can be uniform. Thus, our assertion is established. However, the model (2.12) possesses some important properties (Mardia, 1975a). Let us take $\mathbf{A} = \rho(\kappa_1 \kappa_2)^{\frac{1}{2}} \mathbf{B}$ in (2.12) where $\mathbf{B}$ is an orthogonal matrix. The parameter ρ then behaves as a measure of circular dependence. As $\kappa_1, \kappa_2 \to \infty$, $(\kappa_1^{\frac{1}{2}}\theta, \kappa_2^{\frac{1}{2}}\phi)$ tends to a bivariate normal with limiting correlation depending on ρ. For $\rho = 0$, θ and ϕ are independently distributed and each has a von Mises distribution. When one or both of the variables are uniformly distributed, ρ has no meaning. The question of a measure of circular dependence will be discussed further in Section 6.

Two other methods of constructing directional distributions consist of considering the distribution of $\mathbf{l}$ to be either a conditional distribution or the marginal distribution of an appropriate multinormal distribution (real or complex). We have already used the first of these methods in Section 2.2.

3. Isotropic Random Walk and Sampling from von Mises–Fisher Populations

3.1. *Introduction*

Let $\mathbf{l}_1, \ldots, \mathbf{l}_n$ be a random sample from $M_p(\boldsymbol{\mu}, \kappa)$. Inference problems for $\boldsymbol{\mu}$ and κ depend on the sufficient statistics $\sum \mathbf{l}_i^{\mathrm{T}}$ which we shall denote in the following different ways.

$$\sum_{i=1}^{n} \mathbf{l}_i^{\mathrm{T}} = (R_{x_1}, \ldots, R_{x_p}) = \mathbf{R}_x^{\mathrm{T}} = R\mathbf{l}_0^{\mathrm{T}} = R\mathbf{u}^{\mathrm{T}}(\bar{\mathbf{x}}_0), \tag{3.1}$$

Following the terminology of mechanics, $R_{x_1}, \ldots, R_{x_p}$ are described as $x_1, \ldots, x_p$-components of $\mathbf{l}_i$, $i = 1, \ldots, n$. The length and direction of the resultant are given by R and $\bar{\mathbf{x}}_0$ respectively, which are described as the sample resultant length and the sample mean-direction vector respectively.

In most cases, the sampling distribution of R and $\bar{\mathbf{x}}_0$ from the von Mises–Fisher population can be derived with the help of the corresponding distributions for the uniform case. For example, if $f_\kappa(\mathbf{R}_x; \boldsymbol{\mu})$ denotes the p.d.f. of $\mathbf{R}_x$ for the von Mises–Fisher case, then, by a simple argument,

$$f_\kappa(\mathbf{R}_x; \boldsymbol{\mu}) = \text{const} \times \exp(\kappa \boldsymbol{\mu}^{\mathrm{T}} \mathbf{R}_x)\, f_0(\mathbf{R}_x),$$

where $f_0(\mathbf{R}_x)$ denotes the p.d.f. of $\mathbf{R}_x$ for the uniform case. In fact, the distribution theory for the uniform case is comparatively well developed as it is related to the problem of the isotropic random walk on a hypersphere. Various results can be derived for the uniform case by using a characteristic function method which we shall describe in Section 3.2.

3.2. *The Characteristic Function Method and the Isotropic Case*

Let $\mathbf{l}$ be a unit random vector having a singular density. Suppose that the characteristic function (c.f.) of $\mathbf{l}$ is

$$\phi(\mathbf{t}) = E\{\exp(i\mathbf{t}^{\mathrm{T}}\mathbf{l})\}, \quad \mathbf{t}^{\mathrm{T}} = (t_1, \ldots, t_p). \tag{3.2}$$

Then the p.d.f. of $\mathbf{R}_x$ can be obtained from the inversion theorem and so on transforming $\mathbf{t}$ to the spherical polar co-ordinates ρ and $\boldsymbol{\Phi}^{\mathrm{T}} = (\Phi_1, \ldots, \Phi_{p-1})$ with the help of the transformation $\mathbf{t} = \rho\mathbf{u}(\boldsymbol{\Phi})$, we can find the p.d.f. of $(R, \bar{\mathbf{x}}_0)$.

For the isotropic random walk with n equal steps starting from the origin in p-dimensions, the direction-vectors $\mathbf{l}_1, \ldots, \mathbf{l}_n$ of the n-steps can be regarded as a random sample from the population with p.d.f. (2.3). Hence, it is found that (i) R and $\bar{\mathbf{x}}_0$ are independently distributed, (ii) $\bar{\mathbf{x}}_0$ is uniformly distributed on S_p and (iii) the p.d.f. of R is $h_n(R)$ which is the value of (3.3) at $\beta_1 = \ldots = \beta_n = 1$. If the successive steps are of different length $\beta_1, \ldots, \beta_n$, then the p.d.f. of R is

$$h_n(R;\, \beta_1, \ldots, \beta_n) = (2\pi)^{-p}\, c_p^{n-1}\, R^{p-1} \int_0^\infty \rho^{p-1} \left\{ c_p(i\rho R) \prod_{j=1}^{n} c_p(i\rho\beta_j) \right\}^{-1} d\rho, \tag{3.3}$$

where

$$c_p(i\rho) = \rho^{\frac{1}{2}p-1} / \{(2\pi)^{\frac{1}{2}p}\, J_{\frac{1}{2}p-1}(\rho)\}. \tag{3.4}$$

The distribution of R given by (3.3) was first derived by G. N. Watson (1948, p. 421) who extended a method of Kluyver (1906) for $p = 2$; the method involves highly specialized Bessel function theory. For $p = 3$, the distribution has been simplified by Mardia (1972a, pp. 238–240) on relating the integral (3.3) to a c.f. of some uniform variables on a line. The distribution for $p = 2$ and $p = 3$ has received the attention of various workers (see Mardia, 1972a, pp. 96, 240).

3.3. *Distributional Problems for the von Mises–Fisher Case*

3.3.1. *Single sample problems*

We now quote some results from Mardia (1975b) which are obtained in general from Section 3.2. on using the process described in Section 3.1. Let us assume that $\mathbf{l}_1, \ldots, \mathbf{l}_n$ is a random sample from $M_p(\boldsymbol{\mu}, \kappa)$.

(i) $$\bar{\mathbf{x}}_0 \mid R \sim M_p(\boldsymbol{\mu}, \kappa R). \tag{3.5}$$

The distribution of $\bar{\mathbf{x}}_0$ has not received sufficient attention although it plays an important role.

(ii) The p.d.f. of R_{x_1} is

$$f(R_{x_1}) = \pi^{-1}\, c_p^n(\kappa) \exp(\kappa\mu R_{x_1})\, b(R_{x_1};\, \lambda), \quad -\infty < R_{x_1} < \infty, \tag{3.6}$$

where

$$b(R_{x_1};\, \lambda) = \int_0^\infty \cos t R_{x_1}\, c_p^{-n}\{i(t^2 - \lambda^2)^{\frac{1}{2}}\} dt, \quad \mu = \cos\mu_{0,1}, \quad \lambda = \kappa \sin\mu_{0,1}. \tag{3.7}$$

(iii) The conditional p.d.f. of $R \mid R_{x_1}$ is

$$\pi c_p^{-n} [c_{p-1}\{\lambda(R^2 - R_{x_1}^2)^{\frac{1}{2}}\}]^{-1}\, h_n(R)\, (R^2 - R_{x_1}^2)^{\frac{1}{2}(p-3)} / b(R_{x_1};\, \lambda), \quad 0 < R_{x_1}^2 < R^2. \tag{3.8}$$

Intuitively, it is not clear why the distribution should depend only on λ.

3.3.2. *Multi-sample problem*

Let $R_1, \ldots, R_q$ denote the resultants for q independent random samples of sizes $n_1, \ldots, n_q$ drawn from $M_p(\boldsymbol{\mu}, \kappa)$. Let $n = \sum n_i$. Suppose that R is the resultant for the combined sample. Extending the argument of Mardia (1972a, pp. 243–244) it is found that (Mardia, 1975b)

$$f\kappa(\mathbf{R}^* \mid R) = h_q(R;\, \mathbf{R}^*) \prod_{i=1}^{q} h_{n_i}(R_i)/h_n(R), \tag{3.9}$$

where $h_q(R;\, \mathbf{R}^*)$ and $h_n(R)$ are given at (3.3) and $\mathbf{R}^* = (R_1, \ldots, R_q)^{\mathrm{T}}$. Hence the density of $\mathbf{R}^* \mid R$ does not depend on κ. This fact leads to its applications for inference on $\boldsymbol{\mu}$ when κ is unknown.

It should be noted that the result (3.9) was derived for $p = 3$ and $q = 2$ by Fisher (1953). For $p = 2$, the result is due to J. S. Rao (1969). For $p = 3$, Mardia (1972a, pp. 242–244) provides the exact solution to the problem and also gives historical points. For $q = 2$, (3.9) can be further simplified (Mardia, 1975b).

4. AN INFERENCE PROBLEM FOR SAMPLES FROM VON MISES–FISHER POPULATIONS

Suppose that $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_n$ is a random sample from $M_p(\boldsymbol{\mu}, \kappa)$. Consider the problem of testing

$$H_0\colon \boldsymbol{\mu}_0 = \mathbf{0} \quad \text{against} \quad H_1\colon \boldsymbol{\mu}_0 \neq \mathbf{0}, \tag{4.1}$$

where κ is unknown. Since R_{x_1} is a complete sufficient statistic for κ when H_0 is true and $R_{x_1}, \ldots, R_{x_p}$ are jointly complete and sufficient for $\kappa, \boldsymbol{\mu}_0$ when H_1 is true, the use of the conditional distribution of $\mathbf{R}_x = (R_{x_1}, \ldots, R_{x_p})^{\mathrm{T}} \mid R_{x_1}$ can only lead to similar tests (see Lehmann, 1959, p. 130). It can be seen that there is no *UMP* similar test.

A UMP invariant test. Let us modify the hypotheses (4.1) to

$$H_0\colon \boldsymbol{\mu} = \mathbf{e}_1, \quad H_1\colon \boldsymbol{\mu} \neq \mathbf{e}_1 \tag{4.2}$$

where $\mathbf{e}_1 = (\pm 1, 0, 0, \ldots, 0)^{\mathrm{T}}$. The problem remains invariant under the orthogonal transformations $\boldsymbol{l}^* = \mathbf{A}\boldsymbol{l}$, where $\mathbf{A}$ is an orthogonal matrix with the first column as $\mathbf{e}_1$. The condition on $\mathbf{A}$ ensures that R_{x_1} is invariant under the transformation. A function of $\mathbf{R}_x$ invariant under the transformation is the resultant length

$$R = \left\{\left(\sum_{i=1}^{n} \boldsymbol{l}_i\right)^{\mathrm{T}} \left(\sum_{i=1}^{n} \boldsymbol{l}_i\right)\right\}^{\frac{1}{2}}.$$

Indeed, following the argument of Lehmann (1959, p. 297), it is seen that R is a maximal invariant. Hence the invariant tests should be based on $R \mid R_{x_1}$. We now show that the region

$$\{R > K \mid R_{x_1}\} \tag{4.3}$$

provides a *UMP* invariant test. By the Neyman–Pearson lemma, the best invariant critical region is given by

$$f(R \mid R_{x_1};\, \boldsymbol{\mu}, \kappa)/f(R \mid R_{x_1};\, \mathbf{e}_1, \kappa) \geqslant K, \tag{4.4}$$

where the p.d.f. of $R \mid R_{x_1}$ is given by (3.8). It is seen that (4.4) reduces to

$$I_{\frac{1}{2}p-3}\{\lambda(R^2 - R_{x_1}^2)^{\frac{1}{2}}\}/b(R_{x_1};\, \lambda) \geqslant K, \quad \lambda \neq 0.$$

Since $b(R_{x_1};\lambda)$ is a factor in the p.d.f. of R_{x_1} given by (3.6) it is positive. Further, for $Z>0$, $I_{\frac{1}{2}p-3}(Z)$ is a monotonically increasing function of Z. Thus the best invariant critical region for all values of λ is given by (4.3). The result follows.

It should be noted that the power of the test contains the "non-centrality parameter" λ. The test was recommended by Watson and Williams (1956) on intuitive grounds. Under the null hypothesis, the p.d.f. of $R\,|\,R_{x_1}$ is an even function of R_{x_1} so that the above test is strictly a test for a prescribed axis rather than of a fixed direction.

Confidence cone. Let R_0 be the observed value of R, and let $R_{x_1,\alpha}>0$ and δ satisfy the equations

$$P(R>R_0\,|\,R_{x_1}=R_{x_1,\alpha} \text{ and lies in the same direction as } \boldsymbol{\mu}_0)=\alpha,$$

$$R_{x_1,\alpha}=R_0\cos\delta,\quad 0<\delta\leqslant\tfrac{1}{2}\pi. \tag{4.5}$$

Let us assume that $\boldsymbol{\mu}_0=\mathbf{0}$. Then $R_{x_1}=R\cos\bar{x}_{0,1}$ where $\bar{x}_{0,1}$ denotes the angle between the true mean direction and the sample mean direction. Consequently, the probability that the true mean direction lies within a cone with vertex at the origin, axis as the sample mean direction and semi-vertical angle δ is $1-\alpha$ provided that $0<\bar{x}_{0,1}<\frac{1}{2}\pi$. Hence, the confidence coefficient is not $1-\alpha$ but $1-\alpha'$ such that

$$1-\alpha'=(1-\alpha)\,p^*, \tag{4 6}$$

where

$$p^*=P(0<\bar{x}_{0,1}<\tfrac{1}{2}\pi),$$

when the underlying population is $M_p(\mathbf{0},\kappa)$.

Exact values of p^* can be calculated from (3.6) since it is equivalent to $P(R_{x_1}>0)$ when $\mu=1$ and $\lambda=0$ in (3.6). For large κ, $p^*\to 1$ so that $\alpha'\to\alpha$. Table 1 gives a

TABLE 1

Confidence cone for $\boldsymbol{\mu}$: *comparison of actual confidence coefficient* $1-\alpha'$, *and pseudo-confidence coefficient* $1-\alpha$

p	$1-\alpha$	$n\to$	10	10	20	20	40	40
2		$\kappa\to$	0·74	1·08	0·53	0·74	0·37	0·53
	0·95		0·90	0·94	0·90	0·94	0·90	0·94
	0·99		0·94	0·98	0·94	0·98	0·94	0·98
3		$\kappa\to$	0·91	1·30	0·64	0·91	0·45	0·64
	0·95		0·90	0·94	0·90	0·94	0·90	0·94
	0·99		0·94	0·98	0·94	0·98	0·94	0·98

comparison of the actual confidence coefficient $1-\alpha'$ and the "pseudo-confidence coefficient" $1-\alpha$ by using the values of p^* from Pearson and Hartley (1972, pp. 127–131). Thus for moderately large values of n or κ, α' is approximately equal to α. Of course, we can construct an exact two-sided cone since then the factor $p^*=1$.

An approximate confidence cone can be obtained using (3.5). Let $\hat{\kappa}$ be the maximum likelihood estimator (m.l.e.) of κ and let $\kappa' = R\hat{\kappa}$. Suppose that θ_1 is distributed as the first marginal variable of $M_p(\boldsymbol{\mu}, \kappa')$. That is, the p.d.f. of θ_1 is

$$f(\theta_1) = \{c_p(\kappa')/c_p\} \exp(\kappa' \cos\theta_1), \quad 0 < \theta < \pi.$$

We can obtain a δ' such that

$$P(\pi - \delta' < \theta_1 < \pi) = \alpha.$$

An approximate $(1-\alpha)$-confidence cone for $\boldsymbol{\mu}$ is the cone with vertex at the origin, the axis as the sample mean direction and the semi-vertical angle δ'.

A question arises whether the true mean direction should point towards the sample mean direction or could it point in the opposite direction? If so, how to determine which one is the appropriate cone?

Mardia (1975b) deals with optimum properties of various other tests including multi-sample tests depending on (3.9).

5. Sampling from the Bingham Distribution

5.1. *Distribution of* $\mathbf{T}$ *and* $\boldsymbol{\tau}$ *for the Isotropic Case*

Let us assume that $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_n$ is a random sample from the Bingham population with p.d.f. (2.9). Inference problems for this case are found to depend on the matrix of sum of squares and products

$$\mathbf{T} = \sum_{i=1}^{n} \boldsymbol{l}_i \boldsymbol{l}_i^{\mathrm{T}} \tag{5.1}$$

and the eigenvalues $\tau_1, \ldots, \tau_p$ of $\mathbf{T}$. As in the von Mises–Fisher case, the sampling distribution theory for the Bingham population depends heavily on the distribution theory for the uniform case. However, the exact distributions of $\mathbf{T}$ and $\boldsymbol{\tau}^{\mathrm{T}} = (\tau_1, \ldots, \tau_p)$ for the uniform case are unknown and this fact has somewhat hampered further progress (cf. Bingham, 1964; Anderson and Stephens, 1972). We briefly outline the solution to this problem without giving rigorous detail. For simplicity, let us assume $p = 3$ so that the p.d.f. of $\boldsymbol{l}$ is given by

$$f(\boldsymbol{l};\, \boldsymbol{\mu}, \boldsymbol{\kappa}) = \{4\pi d(\boldsymbol{\kappa})\}^{-1} \exp\left\{\sum_{i=1}^{3} \kappa_i (\boldsymbol{l}^{\mathrm{T}} \boldsymbol{\mu}_i)^2\right\}, \tag{5.2}$$

where $d(\boldsymbol{\kappa})$ is a confluent hypergeometric function (Mardia, 1972).

We first obtain the distribution of $\mathbf{T}$ by the c.f. method. Let $\mathbf{U} = (u_{ij})$ be a symmetric matrix of order 3×3 with $\operatorname{tr}(\mathbf{U}) = 1$. The c.f. of $\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}$ is given by

$$\phi(\mathbf{U}) = (4\pi)^{-1} \int_{S_3} \exp(i\boldsymbol{l}^{\mathrm{T}} \mathbf{U} \boldsymbol{l})\, d\boldsymbol{l}. \tag{5.3}$$

There exists an orthogonal matrix $\mathbf{C}$ such that $\mathbf{U} = \mathbf{C}^{\mathrm{T}} \boldsymbol{\Lambda} \mathbf{C}$ where $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$. Substituting the value of $\mathbf{U}$ in (5.3) and transforming $\boldsymbol{l}$ to $\boldsymbol{l}^* = \mathbf{C}\boldsymbol{l}$, it is found from (5.2) that $\phi(\mathbf{U}) = d(i\boldsymbol{\Lambda})$. Hence, the c.f. of $\mathbf{T}$ is known, which, when used in the inversion theorem, gives the p.d.f. of $\mathbf{T}$ to be

$$f(\mathbf{T}) = 4(2\pi)^{-3} \int_{\Sigma\lambda_i = 1} \left[\int_{\mathrm{O}(3)} \exp\{-i \operatorname{tr}(\mathbf{T}\mathbf{C}^{\mathrm{T}} \boldsymbol{\Lambda} \mathbf{C})\}\, d\mathbf{C}\right] d^n(i\boldsymbol{\Lambda}) \prod_{i<j} (\lambda_i - \lambda_j)\, d\boldsymbol{\Lambda}, \tag{5.4}$$

where the inner integral is over the group of orthogonal matrices of order 3 denoted by O(3) and the outer integral is taken over λ_1, λ_2 and λ_3 with $\lambda_1 + \lambda_2 + \lambda_3 = 1\ \lambda_1 > \lambda_2 > \lambda_3$.

In deriving (5.4) we have used the decomposition

$$d\mathbf{U} = \prod_{i<j} (\lambda_i - \lambda_j)\, d\mathbf{C}\, d\boldsymbol{\lambda};$$

the differential form $d\mathbf{C}$ representing invariant Haar measure on O(3) (see Kshirsagar, 1972, pp. 517–518). Using equation (60) of James (1964), (5.4) reduces to

$$f(\mathbf{T}) = 4(2\pi)^{-3} \sum_{k=0}^{\infty} \sum_{m} \frac{(-i)^k C_m(\mathbf{T})}{k!\, C_m(\mathbf{I}_3)} \int_{\Sigma\lambda_i=1} C_m(\boldsymbol{\Lambda})\, d^n(i\boldsymbol{\Lambda}) \prod_{i<j} (\lambda_i - \lambda_j)\, d\boldsymbol{\Lambda}, \tag{5.5}$$

where the inner summation is taken over all partitions m of k into p or fewer parts, $\mathbf{I}_3$ is the 3×3 identity matrix and $C_m(\cdot)$ is as defined in James (1964).

Let us transform $\mathbf{T}$ in (5.5) with the help of $\mathbf{T} = \mathbf{D}^{\mathrm{T}} \boldsymbol{\tau} \mathbf{D}$, where $\mathbf{D}$ is an orthogonal matrix. Using the Jacobian given in Kshirsagar (1972, Example 83, p. 517), it is found that $\mathbf{D}$ and $\boldsymbol{\tau}$ are independently distributed, and the p.d.f. of $\boldsymbol{\tau}$ is given by

$$4\pi^{-1} \prod_{i<j} (\tau_i - \tau_j) \sum_{k=0}^{\infty} \sum_{m} \frac{(-i)^k C_m(\boldsymbol{\tau})}{k!\, C_m(\mathbf{I}_3)} \int_{\Sigma\lambda_i=1} C_m(\boldsymbol{\Lambda})\, d^n(i\boldsymbol{\Lambda}) \prod_{i<j} (\lambda_i - \lambda_j)\, d\boldsymbol{\Lambda}, \tag{5.6}$$

where $\tau_1 > \tau_2 > \tau_3 > 0$ and $\tau_1 + \tau_2 + \tau_3 = 1$. Hence, the exact distribution of $\boldsymbol{\tau}$ is known but, as in the normal case, (5.6) is not manageable. An approximation can be obtained following Bingham (1972). The asymptotic distribution of $\boldsymbol{\tau}$ has already been obtained by Anderson and Stephens (1972), using an elegant argument.

5.2. *Critical Values of a Test of Uniformity*

To test the hypothesis of uniformity against the alternative hypothesis of a Bingham distribution, we use the criterion (see Mardia, 1972a, p. 276)

$$U = S_u/n = \frac{15}{2} \sum_{i=1}^{3} (\bar{\tau}_i - \tfrac{1}{3})^2, \quad \bar{\tau}_i = \tau_i/n,$$

where the null hypothesis is rejected for large values of U. It is known that S_u is asymptotically distributed as χ^2_5. The exact distribution of U can be formally written from (5.6) but it leads to a formidable numerical integration problem and we therefore utilized a Monte Carlo method. Using 10,000 trials, we obtained simulated critical values $S_n(\alpha)$ of the test, for each value of α in Table 2, and $n = 6, 7, 8, 10, 12, 15, 30, 40, 60, 100$. Using the points $\{n^{-1}, S_n(\alpha)\}$ together with a few points from the chi-squared approximation for large n, we obtained smooth curves for each α on a computer, with the help of a general procedure which ensures consistency in the shapes of the various curves for differing α. The resulting critical values are shown in Table 2, and these should be accurate to within one unit in the second decimal place. For $n > 100$, the chi-squared approximation is adequate.

Example. From 30 measurements of direction on the c-axis of calcite grains from the Taconic mountains of New York (Bingham, 1964; Mardia, 1972a, p. 226), it is found that

$$\tau_1 = 4{\cdot}1764, \quad \tau_2 = 10{\cdot}9639, \quad \tau_3 = 14{\cdot}8595.$$

We have here $U = 0{\cdot}487$. Using Table 2, we note that the hypothesis is (almost) rejected at the 1 per cent level; the 1 per cent critical value is 0.49. The chi-squared approximation gives the 1 per cent value as $0{\cdot}503$.

TABLE 2

Critical values of the U-test for uniformity

n	$\alpha \to$ 0·001	0·005	0·01	0·025	0·05	0·075	0·1
6	3·21	2·59	2·35	1·99	1·74	1·59	1·47
7	2·82	2·27	2·05	1·74	1·51	1·37	1·27
8	2·50	2·01	1·81	1·53	1·33	1·21	1·12
9	2·24	1·80	1·62	1·37	1·19	1·08	0·99
10	2·03	1·63	1·47	1·24	1·07	0·97	0·90
11	1·85	1·48	1·34	1·13	0·98	0·88	0·82
12	1·70	1·36	1·23	1·04	0·90	0·81	0·75
14	1·46	1·17	1·05	0·89	0·77	0·70	0·64
15	1·36	1·09	0·99	0·83	0·72	0·65	0·60
16	1·28	1·02	0·92	0·78	0·67	0·61	0·56
20	1·02	0·82	0·74	0·62	0·54	0·49	0·45
30	0·68	0·55	0·49	0·42	0·36	0·33	0·30
50	0·41	0·33	0·29	0·25	0·22	0·20	0·18
60	0·34	0·27	0·25	0·21	0·18	0·16	0·15
80	0·25	0·20	0·18	0·16	0·13	0·12	0·11
100	0·20	0·16	0·15	0·12	0·11	0·10	0·09
χ^2_5	20·52	16·75	15·09	12·83	11·07	10·01	9·24

6. DEPENDENCE OF VARIABLES ON A TORUS

6.1. *Rotational Dependence*

Let θ and ϕ be two circular random variables. The standard linear correlation coefficient is supposed to measure the degree of linearity. In fact, two bivariate variables x and y are perfectly correlated if the whole mass is situated on a line. Keeping the algebraic structure of the circle in mind, we may say that θ and ϕ are perfectly correlated if the whole mass is concentrated on

$$(l\theta \pm m\phi + \psi) \bmod 2\pi = 0, \tag{6.1}$$

where l and m are positive integers and ψ is an angular quantity. Relation (6.1) can be understood easily on considering the case of $l = m = 1$. In this case, (6.1) is true if and only if

$$\theta \equiv (\psi + \phi) \bmod 2\pi \tag{6.2}$$

or

$$\theta \equiv (\psi - \phi) \bmod 2\pi. \tag{6.3}$$

In both cases, we can "rotate" ϕ anti-clockwise or clockwise by an angle ψ to match it with θ. The cases (6.2) and (6.3) have also been recently utilized by Downs (1974) in developing his measure of circular correlation. When (6.2) holds, θ and ϕ are said to be positively correlated whereas if (6.3) holds θ and ϕ are negatively correlated.

6.2. *A Circular Rank Correlation Coefficient*

Definition. Let (θ_i, ϕ_i), $i = 1, \ldots, n$, be a random sample on (θ, ϕ) and let

$$\theta_i^* \equiv l\theta_i \bmod 2\pi, \quad \phi_i^* \equiv m\phi_i \bmod 2\pi,$$

where l and m are assumed *known*. Let us assume that the linear ranks of the θ_i^*'s are $1, \ldots, n$ and those of the ϕ_i^*'s are $r_1, \ldots, r_n$. Replace the angles (θ_i^*, ϕ_i^*) by the uniform scores $(2\pi i/n, 2\pi r_i/n)$. Now for (6.3) to hold, the resultant length of the points $(\theta_i^* - \phi_i^*) \bmod 2\pi$ or $(\theta_i^* + \phi_i^*) \bmod 2\pi$, $i = 1, \ldots, n$, should be one for perfect dependence. Thus, we could define a rank correlation coefficient on the circle by

$$r_0 = \max(\bar{R}_1^2, \bar{R}_2^2), \tag{6.4}$$

where

$$n^2 \bar{R}_1^2 = \left[\sum_{i=1}^{n} \cos\{2\pi(i-r_i)/n\}\right]^2 + \left[\sum_{i=1}^{n} \sin\{2\pi(i-r_i)/n\}\right]^2$$

and

$$n^2 \bar{R}_2^2 = \left[\sum_{i=1}^{n} \cos\{2\pi(i+r_i)/n\}\right]^2 + \left[\sum_{i=1}^{n} \sin\{2\pi(i+r_i)/n\}\right]^2.$$

Obviously, $0 \leqslant r_0 \leqslant 1$. Further, r_0 is invariant under changes of zero direction in either θ or ϕ. We have $\bar{R}_1^2 = 1$ for "positive" perfect dependence while $\bar{R}_2^2 = 1$ for "negative" perfect dependence. When the variables are uncorrelated, $\bar{R}_1^2 = \bar{R}_2^2 \simeq 0$ asymptotically.

Null distribution of r_0. Table 3 gives selected critical values of r_0 for $n = 5\ (1)\ 10$. For $n > 10$, the following approximation is found adequate. It can be seen that

$$\bar{R}_1^2 = (T_{cc} + T_{ss})^2 + (T_{cs} - T_{sc})^2, \quad \bar{R}_2^2 = (T_{cc} - T_{ss})^2 + (T_{cs} + T_{sc})^2,$$

where

$$T_{cc} = \frac{1}{n} \sum \cos\frac{2\pi i}{n} \cos\frac{2\pi r_i}{n}, \quad T_{cs} = \frac{1}{n} \sum \cos\frac{2\pi i}{n} \sin\frac{2\pi r_i}{n}$$

TABLE 3

Critical values of the r_0-test (upper entry); lower entry: exact level

n	$\alpha \to$	0·001	0·01	0·05	0·10
5					1·0000
					0·0833
6			1·0000		0·6944
			0·0167		0·1167
7		1·0000	0·7964	0·6160	0·5223
		0·0028	0·0222	0·0611	0·1000
8		1·0000	0·7286	0·5335	0·4321
		0·0004	0·0115	0·0528	0·0980
9		0·8987	0·6273	0·4474	0·37†
		0·0005	0·0096	0·0515	
10		0·84†	0·59†	0·41†	0·33†
$-\log_e\{1-\sqrt{(1-\alpha)}\}$		7·6007	5·2958	3·6761	2·9697

† Values calculated from the approximation.

and T_{ss}, T_{sc} are similarly defined. Since the T's are linear rank statistics, we can obtain their moments using a result of Hajek and Sidak (1967, pp. 57–58). It is found that the vector $\frac{1}{2}(n-1)^{-\frac{1}{2}}(T_{cc}, T_{ss}, T_{cs}, T_{sc})^{\mathrm{T}}$ is asymptotically $N_4(\mathbf{0}, \mathbf{I})$. Hence $2(n-1)\bar{R}_1^2$ and $2(n-1)\bar{R}_2^2$ are independently distributed asymptotically as χ_2^2. As a consequence, for large n, $U = 2(n-1)r_0$ has the p.d.f.

$$f(u) = \exp(-\tfrac{1}{2}u)\{1-\exp(-\tfrac{1}{2}u)\}, \quad u>0,$$

so the upper percentage point $r_{0,\alpha}$ of r_0 is given by

$$r_{0,\alpha} = -(n-1)^{-1}\log_e\{1-(1-\alpha)^{\frac{1}{2}}\}.$$

The values of the logarithmic terms for some selected values of α are given in Table 3.

Example 1. In a medical experiment, various measurements were taken on 10 medical students several times daily for a period of several weeks (Downs, 1974). The estimated peak times for two successive measurements of diastolic blood pressure (converted into angles) were

$$\theta:\ 30°\ 15°\ 11°\ \ 4°\ 348°\ 347°\ 341°\ 333°\ 332°\ 285°$$

$$\phi:\ 25°\ \ 5°\ 349°\ 358°\ 340°\ 347°\ 345°\ 331°\ 329°\ 287°.$$

It is found that $\bar{R}_1^2 = 0{\cdot}731$, $\bar{R}_2^2 = 0{\cdot}004$, $r_0 = 0{\cdot}731$. The 5 per cent value of r_0 is 0·41 and therefore the null hypothesis of independence is rejected. However, if we *incorrectly* apply linear techniques, we find $r = 0{\cdot}53$ which leads to acceptance of the hypothesis at the 5 per cent level of significance under the assumption of normality. It is plausible to expect a dependence on medical grounds. The dependence is positive as reflected by the high value of $\bar{R}_1^2$. Also if we look at the differences in $i-r_i$, we get the same impression.

Example 2. The wind directions at Gorleston on February 1st to 8th, 1968, at 1300 hours each day were

$$220°\ 250°\ 280°\ 330°\ 180°\ 220°\ 140°\ 60°.$$

Is there any serial correlation between these readings? A study for all the data for 1968 suggests that the parent distribution is bimodal. Let $\theta_1, \theta_2, \ldots$ be the angles doubled. By considering the pairs (θ_1, θ_2), (θ_2, θ_3), etc., it is found that $\bar{R}_1^2 = 0{\cdot}2180$, $\bar{R}_2^2 = 0{\cdot}1897$. The 10 per cent value of r_0 for $n = 7$ from Table 3 is 0·5233. Hence, we accept the null hypothesis. If the angles are not doubled, $\bar{R}_1^2 = 0{\cdot}2108$, $\bar{R}_2^2 = 0{\cdot}1217$.

7. CASE STUDIES

We have given some numerical examples to illustrate specific applications in previous sections. We now investigate some broader problems from various fields. For calculations of basic circular and spherical statistics, algorithms of Mardia and Zemroch (1975a, b) were used.

7.1. *Analysis of a Cancer Cell Data*

In studies of cancer (*Rous sarcoma* virus), it is of interest to investigate the orientation of cancer cells (P. Crouch, R. Weiss and H. Goldstein, 1975, unpublished). For this virus, it is possible to regard every cell as a directed line. The cells are plated onto a glass plate, which has a number of parallel equispaced grooves cut into it. The orientation of the cells can then be measured with respect to this grid. It is known that there is a tendency for the cells to align themselves with these grooves, i.e. the grooves have a physical effect on the orientation of the cells. Therefore, it is of interest to study the effect of the grid spacing x (measured as number of grooves per inch) on these orientations. As a control experiment, the cells are plated onto the reverse smooth side of the plate. Orientations are measured as before as it is possible to see the grid through the glass, but obviously now there should be no physical alignment effect.

To investigate the above problems, Mr H. Goldstein of the National Children's Bureau and Dr R. Weiss of the Imperial Cancer Research Institute conducted an experiment to measure orientations of cancer cells for seven values of x (see Table 4).

TABLE 4

Circular statistics for a set of cancer cell data

Serial no.	No. of lines per inch/ 1,000 = $10^{-3}x$	Grooved side			Smooth side		
		$\bar{x}_0^* \pm \delta$	$\bar{R}$	$\hat{\kappa}$	$\bar{x}_0^* \pm \delta$	$\bar{R}$	κ
1	1·25	$-0{\cdot}43° \pm 1{\cdot}5°$‡	0·979	23·57	$-72{\cdot}9° \pm 39°$	0·188†	0·38
2	3·0	$-0{\cdot}43° \pm 2{\cdot}5°$	0·965	14·67	$80{\cdot}35° \pm 62°$	0·138†	0·28
3	5·0	$-1{\cdot}19° \pm 10°$	0·591	1·48	$-15{\cdot}36° \pm 46°$	0·169†	0·34
4	7·5	$-1{\cdot}20° \pm 4°$	0·919	6·45	$44{\cdot}74° \pm 62°$	0·139†	0·28
5	10·0	$2{\cdot}91° \pm 7°$	0·739	2·28	$-60{\cdot}76 \pm 84°$	0·058†	0·12
6	15·0	$-49{\cdot}56° \pm 48°$	0·163†	0·33	$1{\cdot}43° \pm 78°$	0·102†	0·21
7	30·0	$4{\cdot}37° \pm 12°$	0·502	1·16	$-79{\cdot}29° \pm 58°$	0·145†	0·29

† Uniformity hypothesis accepted at the 1 per cent level of significance.
‡ The 95 per cent confidence interval for μ_0^*.

Denote these values of x by x_i, $i = 1, \ldots, 7$ in ascending order. For each x, the orientation of 40 cancer cells was measured independently in turn (i) on the grooved side and (ii) on the smooth side of such a plate. The cells were not labelled so that the resulting angles are not paired. After doubling the angles, various circular statistics were obtained and are displayed in Table 4. Each value of $\bar{x}_0^*$ corresponds to the direction after dividing the mean direction of the doubled angles in the range $(-180°, 180°)$ by two. Corresponding 95 per cent confidence intervals for $\bar{x}_0^*$ are also shown. For testing the hypothesis of uniformity against a von Mises distribution for the doubled angles for each value of x, note that the 5 per cent critical value of $\bar{R}$ is 0·273 and the 1 per cent value of $\bar{R}$ is 0·336. It is interesting to see that the hypothesis of uniformity is accepted for the smooth sides for all values of x so that orientations on a smooth plate are not *informative* for cancer cells. In contrast, except for $x = 15{,}000$, the hypothesis of uniformity is rejected for the grooved plates. Let the doubled angles on the grooved plate with spacing x_i be a sample from $M(2\mu_{0\ i}^*, \kappa_i)$. It is seen that for each value of i where the distribution was found to be non-uniform, the hypothesis $\mu_{0,i}^* = 0$ is accepted. That is, the mean orientations $\mu_{0,i}^*$ do not depend on the grid spacing x—as was anticipated prior to the experiment. This is clear from the values of $\bar{x}_{0,i}$. The relationship between x_i and $\hat{\kappa}_i$, $i = 1, \ldots, 7$, is oscillatory, dampening down as $x \to \infty$. In view of the small number of points (x_i) available, instead of using a Fourier series, a linear regression of $\log_e \kappa$ on $\log_e x$ was tried. Fig. 1 shows the plot of κ against x with the corresponding fitted curve being

$$\kappa = 96727x^{-1{\cdot}1730}. \tag{7.1}$$

As $x \to \infty$, we have $\kappa \to 0$. It was expected from physical considerations that we would approach a uniform distribution as the grooves became closer. A uniform distribution was also expected as the grooves became infinitely wide apart ($x \to 0$) as of course then there would be so few grooves that their effect would become negligible. However, there are no data available for $x < 1{,}250$ to pursue this point further.

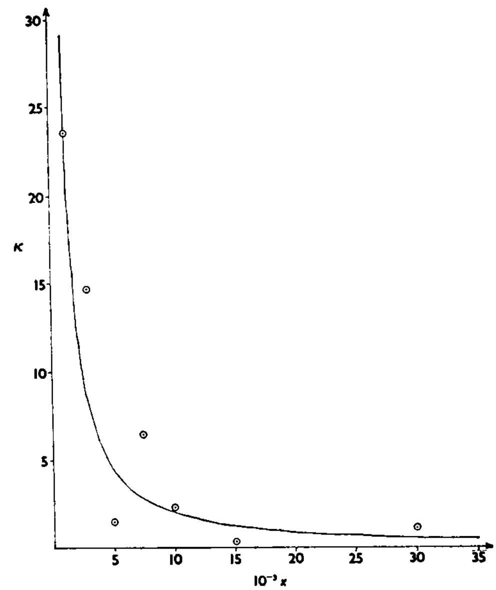

FIG. 1. Relation between the concentration parameter κ and x, the number of grooves per inch/1,000. Fitted curve and the observed points.

An alternative estimation procedure would be to regard the ith sample as drawn from $M(2\mu_0, ax_i^{-b})$ and obtain the maximum likelihood estimators of the parameters by using a numerical procedure.

7.2. *Analysis of Long-period Comet Data*

Tyror (1957) analysed the distribution of the directions of the perihelia of 448 long-period comets. From the accretion theory of Lyttleton (1953, 1961), one expects that

(i) the distribution should be non-uniform, and
(ii) the perihelion points should exhibit a preference for lying near the galactic plane.

Following Tyror (1957), we use a right-handed Cartesian system with the origin at the sun and OX and OY in the plane of the ecliptic (with OX in direction of $\phi = 0$), i.e.

$$l = \sin\theta\cos\phi, \quad m = \sin\theta\sin\phi, \quad n = \cos\theta.$$

In his notation, $\lambda = \phi$ and $\beta = 90° - \theta$. The various statistics in the notation of Mardia (1972) are as follows.

$$R = 63{\cdot}18, \quad (\bar{l}_0, \bar{m}_0, \bar{n}_0) = (-0{\cdot}0541, -0{\cdot}3316, 0{\cdot}9419), \quad \bar{x}_0 = 19{\cdot}63°, \quad \bar{y}_0 = 260{\cdot}73°,$$

$$\mathbf{T}/n = \begin{bmatrix} 0{\cdot}289988 & -0{\cdot}012393 & 0{\cdot}030863 \\ -0{\cdot}012393 & 0{\cdot}389926 & -0{\cdot}011153 \\ 0{\cdot}030863 & -0{\cdot}011153 & 0{\cdot}320086 \end{bmatrix}.$$

The matrix **T** differs only slightly from that given by Tyror (1957). (From our private correspondence, it emerges that the difference could be due to the use of ungrouped data by Tyror.) We find that the eigenvalues of $\mathbf{T}/n$ are

$$\bar{\tau}_1 = 0{\cdot}2705, \quad \bar{\tau}_2 = 0{\cdot}3347, \quad \bar{\tau}_3 = 0{\cdot}3947.$$

Consequently, $S_u = 25{\cdot}93$. Now, S_u is distributed as χ_5^2 for large n and $P(\chi_5^2 > 25{\cdot}93) = 9{\cdot}2 \times 10^{-5}$. Hence the hypothesis of uniformity is rejected.

Tyror used "an equal area" investigation involving counting the number of points per area and the points in the four immediately adjacent areas. His conclusion was the same. However, the above analysis is more appropriate since we have the alternative of a Bingham distribution in mind. The Rayleigh test leads to the same conclusion. Following Mardia (1972a, pp. 225, 277–278) it is found that the distribution is girdle. (Here $\hat{\kappa}_1 = -0{\cdot}9545$, $\hat{\kappa}_2 = -0{\cdot}3978$, $S_g = 5{\cdot}74$, $S_b = 8{\cdot}01$.) Further, the normal to the preferred great circle has the direction cosines,

$$l = 0{\cdot}8522, \quad m = 0{\cdot}0397, \quad n = -0{\cdot}5217 \tag{7.2}$$

which is the direction of the eigenvector corresponding to $\bar{\tau}_1$.

If we now assume the distribution is girdle with p.d.f.

$$c \exp\{-\kappa(\boldsymbol{l}^{\mathrm{T}} \boldsymbol{\mu})^2\}, \quad \kappa > 0,$$

it remains to test

$$H_0\colon \boldsymbol{\mu}^{\mathrm{T}} = (0{\cdot}8772, -0{\cdot}0536, -0{\cdot}4772) \tag{7.3}$$

which defines the direction of the galactic pole. Let λ be the likelihood ratio criterion for this problem. We have

$$-2\log_e \lambda = -2n \log b(\hat{\kappa}_0) + 2\hat{\kappa}_0 \boldsymbol{\mu}^{\mathrm{T}} \mathbf{T} \boldsymbol{\mu} + 2n \log_e b(\hat{\kappa}_1) - 2\hat{\kappa}_1 \tau_1,$$

where

$$1/b(\kappa) = 2\int_0^1 \exp(-\kappa t^2)\, dt, \quad \hat{\kappa}_1 = D^{-1}(\bar{\tau}_1), \quad \hat{\kappa}_0 = D^{-1}(\boldsymbol{\mu}^{\mathrm{T}} \mathbf{T} \boldsymbol{\mu}/n)$$

with $D(\cdot)$ defined by Mardia (1972, p. 253) and $\boldsymbol{\mu}$ given by (7.3). We find that $-2\log_e \lambda = 0{\cdot}95$. Since n is large, $-2\log_e \lambda$ can be assumed to be distributed as χ_2^2. Hence there is strong evidence that the perihelion points exhibit a preference for lying near the galactic pole. This is not surprising since the angle between the directions (7.2) and (7.3) is only $6{\cdot}1°$. Hence, there is strong evidence that the Lyttleton theory is true. It may be noted that the angle between the mean direction and the apex of the sun's motion is $16{\cdot}3°$. Fig. 2 shows a Schmidt net (Mardia, 1972, p. 215) for the first 432 of these comets with its centre as the galactic pole.

The projection is an equal area projection and therefore we would expect more points towards the edges of the net under a girdle distribution. The figure reflects this feature. (The net is given only for the upper hemisphere. For the lower hemisphere the net is very similar.)

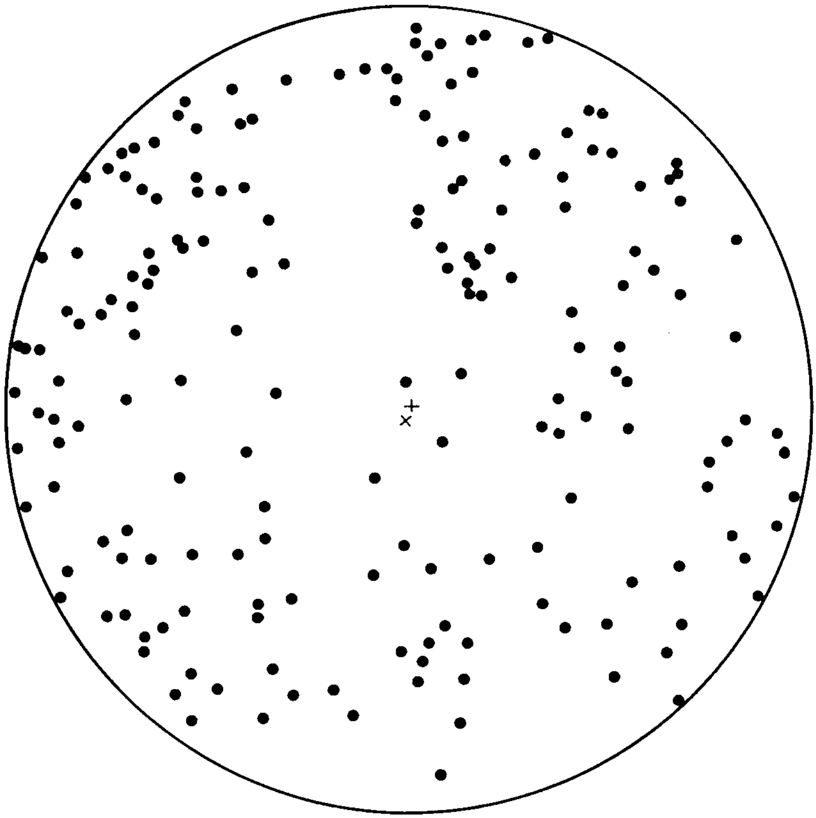

FIG. 2. Schmidt net for Tyror's data. +, Centre of the circle at the galactic pole; ×, estimated direction of the normal to the girdle plane.

7.3. *Cyclic Variation in Poisson Processes and Patient Arrival Data*

In some point processes, there are reasons to believe that there exists a cyclic effect, e.g. the arrival times of patients may have a time-of-day effect. Naturally, this leads to the question of testing a pure Poisson process against a non-homogeneous Poisson process with a cyclic trend. There exist various tests for the Poisson model (Cox and Lewis, 1966, pp. 152–164, 172–173). A simple test based on directional data analysis is as follows. Let 2π be the period suspected and suppose that we observe a pure Poisson process for a predetermined time t_0 which is an integral multiple of 2π. Further, let us suppose that n events are observed. We know (see, for example, Cox and Lewis, 1966, pp. 27–28) that the positions $U_1, \ldots, U_n$ of the events (measured as the distances from the origin) in a Poisson process are independently uniformly distributed over $(0, t_0)$ given that n events have occurred in all.

Since t_0 is an integral multiple of 2π, the random variables

$$\theta_i \equiv U_i \bmod 2\pi, \quad i = 1, \ldots, n \tag{7.4}$$

are independently uniformly distributed over the unit circle. Hence, conditionally on n, under the hypothesis of a pure Poisson process $\theta_1, \ldots, \theta_n$ is a random sample from the distribution with p.d.f.

$$f(\theta) = 1/(2\pi), \quad 0 < \theta \leqslant 2\pi.$$

Under the alternative hypothesis of a non-homogeneous Poisson process, we may assume that θ is distributed as $M(\mu_0, \kappa)$ which implies a cyclic effect. Hence, we are led to the Rayleigh test of uniformity when $\theta_1, \ldots, \theta_n$ are the observations (see Mardia, 1972b).

Example. Cox and Lewis (1966, pp. 254–255) give data regarding arrival times of patients at an intensive care unit. One expects arrivals to be influenced by a time-of-day effect. The number of arrivals is 254, and it is found that $\bar{R} = 0{\cdot}320$. As n is large we may use the chi-squared approximation. Since $2n\bar{R}^2 = 52{\cdot}02$, we reject the hypothesis of uniformity at the 0·1 per cent level. That is, the process is not pure Poisson. We can go one step further to enquire whether the von Mises distribution provides a good fit. It is found that $\hat{\mu}_0 = 96{\cdot}91°$ (= 1728 hours), $\hat{\kappa} = 0{\cdot}676$. The expected and observed frequencies are shown in Table 5. The value of the goodness

TABLE 5

Observed and expected frequencies when a von Mises distribution is fitted to the arrival times

Arrival times	*Observed frequency*	*Expected frequency*	*Arrival times*	*Observed frequency*	*Expected frequency*
$0 \leqslant t < 1$	5	7·9	12–	19	11·4
1–	9	6·7	13–	12	13·4
2–	11	5·9	14–	14	15·3
3–	6	5·3	15–	16	17·0
4–	4	4·9	16–	17	18·2
5–	1	4·8	17–	19	18·6
6–	1	4·9	18–	15	18·1
7–	7	5·3	19–	14	16·9
8–	3	5·9	20–	15	15·2
9–	2	6·8	21–	16	13·2
10–	12	8·0	22–	11	11·2
11–	13	9·5	23–	12	9·4

of fit statistic χ^2 is found to be 29·31. Now, the 10 per cent value of χ^2_{21} is 29·62, and hence we accept the null hypothesis.

7.4. *A Mixture of von Mises Distributions and some Turtle Data*

The bimodal distribution of von Mises type assumes that the two modes are of the same strength and are situated 180° apart. A more general bimodal distribution can be obtained by considering a mixture of two von Mises distributions. That is, the p.d.f. of θ is given by

$$f(\theta;\, \mu_{0,1}, \mu_{0,2}, \kappa_1, \kappa_2, \lambda) = \lambda g(\theta;\, \mu_{0,1}, \kappa_1) + (1-\lambda) g(\theta;\, \mu_{0,2}, \kappa_2), \tag{7.5}$$

where $g(\theta;\, \mu_0, \kappa)$ denotes the p.d.f. of $M(\mu_0, \kappa)$, $0 \leqslant \lambda \leqslant 1$ and $0 < \theta \leqslant 2\pi$. Sufficient conditions for (7.5) to be a bimodal distribution rather than unimodal are given in Mardia and Sutton (1975). The m.l.e. of the five parameters can be obtained numerically with the help of a program of Jones and James (1969). We use this method on some turtle data which have been previously investigated (Stephens, 1969; Boneva *et al.*, 1971, Mardia, 1972a, pp. 11, 129). Boneva *et al.* obtained a circular histospline; the raw data are given in Stephens (1969). Of course histospline analysis does not give a functional form of the density. For the grouped data used in Fig. 7 of Boneva *et al.*, it is found that

$$\lambda = 0{\cdot}85, \quad \mu_{0,1} = 65{\cdot}3^\circ, \quad \mu_{0,2} = 239{\cdot}0^\circ, \quad \kappa_1 = 1{\cdot}94, \quad \kappa_2 = 7{\cdot}76. \tag{7.6}$$

This uses a very wide cell length, but the estimators are comparable to the estimators for the raw data given by

$$\lambda = 0{\cdot}84, \quad \mu_{0,1} = 63{\cdot}5^\circ, \quad \mu_{0,2} = 241{\cdot}2^\circ, \quad \kappa_1 = 2{\cdot}62, \quad \kappa_2 = 8{\cdot}45.$$

Fig. 3 compares the fitted curve from the mixture with estimators given by (7.6) and the histospline. The two curves are similar and the modes are situated at the same

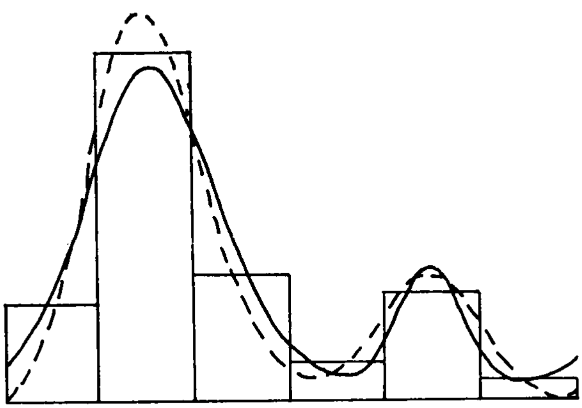

FIG. 3. Histospline (Boneva *et al.*, 1971) and a mixture fitted to turtle data. - - - -, Histospline; ——, mixture.

points. However, the strengths of the two modes given by the two methods differ slightly.

7.5. *Unknown Period*

In directional data analysis, it is generally assumed that the true period is known. However, there are situations where the true period may be unknown. Sometimes, particularly when mapping the real line onto the circle, it is not always clear which is the best period to use. We consider two problems.

(i) We may be looking, mistakenly, not at the true period, but at an integer multiple of it, for example, we may map the week or the year onto the circle when we really ought to be looking at the day. To ascertain if there is a sub-cycle, let us assume the parental population to have a density of the form

$$f(\theta) = c\exp\{\kappa\cos l(\theta-\mu_0)\}, \quad 0<\theta\leqslant 2\pi, \quad \kappa>0,$$

where l is an integer. The m.l.e. of l is l_0 where $R_{l_0}^2 = \max_l (R_l^2)$. The m.l.e.'s of κ and μ_0 are straightforward. The hypothesis of unimodality ($l=1$) against the hypothesis of a sub-cycle ($l=l_0$) can be treated by using the likelihood ratio principle.

(ii) Suppose data emerge from a point process where there exists a cycle. Professor J. N. Mills, of Manchester University, provided us with a set of data concerning the times of potassium excretion of a subject living on a 21-hour period rather than the normal period of 24 hours. The main problem is described in Fort and Mills (1970); we perform here a *confirmatory analysis* to the times of potassium excretion only.

Let $t_1, \ldots, t_n$ be the times in hours. For the data, $n = 68$. Let

$$2\pi\theta_i \equiv t_i \bmod 21, \quad 2\pi\theta_i^* \equiv t_i \bmod 24, \quad i = 1, \ldots, n.$$

We wish to test

$$H_0\colon\ \theta_i, \quad i=1,\ldots,n \quad \text{from} \quad M(\mu_0,\kappa), \quad \kappa>0,$$

against

$$H_1\colon\ \theta_i^*, \quad i=1,\ldots,n \quad \text{from} \quad M(\mu_0',\kappa'), \quad \kappa'>0.$$

We can write the likelihood ratio but its distribution is unknown. However, the circular statistics for varying periods should prove helpful. A plot of

$$R_\lambda^* = (n\pi)^{-1}\left[\left\{\sum_{i=1}^{n}\cos(2\pi t_i/\lambda)\right\}^2 + \left\{\sum_{i=1}^{n}\sin(2\pi t_i/\lambda)\right\}^2\right] = (n\pi)^{-1}\bar{R}_\lambda^2$$

against the period λ is given in Fig. 4. As expected R_λ^* has a maximum around $\lambda = 21$ hours, but it also has a minimum around $\lambda = 24$ hours. For $\lambda = 24$ hours, $\bar{R} = 0{\cdot}038$ so we accept the hypothesis of uniformity at the 10 per cent level, whereas for $\lambda = 21$ hours $\bar{R} = 0{\cdot}348$, so we reject this hypothesis at the 1 per cent level. The

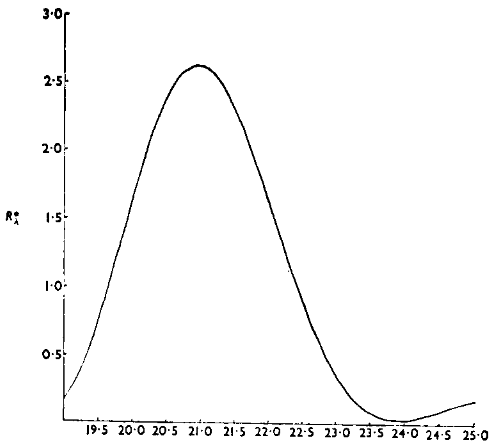

FIG. 4. R_λ^* for the excretion data.

expected and observed frequencies under H_0 were obtained. For a particular grouping, the value of the goodness-of-fit criterion χ^2 was found to be 8·87, distributed as χ_6^2. The 5 per cent value of χ_6^2 is 12·59.

8. CONCLUDING REMARKS

In this final section, we add some miscellaneous comments.

(i) There might exist a limiting process analogous to the central limit theorem leading to the von Mises–Fisher distribution. For the circular case, the Brownian motion distribution appears as a limiting distribution of the mean direction for the circular case (Rukhin, 1971). Hartman and Watson (1974) show that there exists a stopping time distribution for Brownian motion on S_p leading to the von Mises–Fisher distribution. Lewis (1975) has made a significant step forward by showing that von Mises is infinitely divisible for small κ.

(ii) There are various unsolved fundamental estimation problems For example, what does one mean by an "unbiased" estimator of an angular parameter such as the mean direction? When is an estimator best? Some attempts have been made by Mardia (1972, Chapter 5) and Rukhin (1971) The latter reference looks into these problems for a translation family. A Bayesian approach is discussed in Mardia and El-Atoum (1974).

(iii) The definition of circular rank correlation coefficient r_0 given by (6.4) can be modified to the range $(-1, 1)$ so that the positive value of r_0 corresponds to "positive" dependence, etc. However, such a distinction is artificial since we can have $\bar{R}_1^2 = \bar{R}_2^2$. For $n = 6$, and $r_1 = 4$, $r_2 = 1$, $r_3 = 6$, $r_4 = 5$, $r_5 = 2$, $r_6 = 3$, we find $\bar{R}_1^2 = \bar{R}_2^2 = 0{\cdot}25$. Here it is difficult to say whether there is positive or negative dependence.

If the true directions for θ and ϕ are known, say zero, a more appropriate coefficient would be

$$\max\left[\sum \cos\{2\pi(i-r_i)/n\}, \sum \cos\{2\pi(i+r_i)/n\}\right].$$

Its critical values were obtained simultaneously with the work for r_0 but such examples are not so common. Some other non-parametric circular correlation coefficients are given by Rothman (1971).

On following the argument of Section 6.2, we can define in general a circular correlation coefficient for (θ_i^*, ϕ_i^*), $i = 1, \ldots, n$, as

$$\max(D_+, D_-)/\{(1-\bar{R}_1^2)(1-\bar{R}_2^2)\},$$

where $\bar{R}_1$ and $\bar{R}_2$ are the mean resultant lengths of θ_i^* and ϕ_i^* and

$$n^2 D_\pm = \left\{\sum_{i=1}^{n} \cos(\theta_i^* \pm \phi_i^*) - n\bar{R}_1 \bar{R}_2\right\}^2 + \left\{\sum_{i=1}^{n} \sin(\theta_i^* \pm \phi_i^*)\right\}^2.$$

The mean directions for θ_i^* and ϕ_i^* are selected to be zero. A detailed investigation of this coefficient and related functions will appear elsewhere.

(iv) As in multivariate analysis, it is not surprising that there is a dearth of non-parametric tests of practical value for $p \geqslant 3$. In constructions of circular non-parametric tests, the circular uniform scores of Section 6 have been playing a key role (see Wheeler and Watson, 1964; Mardia, 1967, 1969, 1972b; Mardia and Spurr, 1973).

(v) There have been various other notable developments in the area within the past few years. Downs (1972) deals with orientations characterized by k directions (instead of a single direction) in p-dimensions. Matthews (1974) raises various important statistical problems in bird navigation and Kendall (1974a) provides a satisfactory solution to one of these through pole-seeking Brownian motion. Some practical problems of discriminant analysis are treated in Morris and Laycock (1974). Upton (1974) has given some new approximations.

Since the paper was submitted in 1973, it has been brought to the author's notice that Section 7.3 is related to Lewis P.A.W. (1972) and Section 7.5 to Kendall (1974b).

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor T. Lewis for his valuable help. I have benefited from discussions with various friends including Michael Bingham, Jim Thompson, Peter Zemroch, Harold Peers, Dick Gadsden, Barrie Spurr, Terry Sutton and Shafiq El-Atoum, to whom I am most grateful. My thanks are due to Mr H. Goldstein, Professor R. A. Lyttleton and Professor J. N. Mills for providing interesting case studies, and for their comments. The comet problem was brought to my attention by Professor D. G. Kendall. I am grateful to a referee for his helpful suggestions, and to Professor T. Downs for giving permission to use his data. I am also grateful to Mr P. J. Zemroch for his computing assistance.

REFERENCES

ANDERSON, T. W. and STEPHENS, M. A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika*, **59**, 613–621.

BINGHAM, C. (1964). *Distributions on the sphere and on the projective plane*. Ph.D. Thesis, Yale University.

—— (1972). An asymptotic expansion for the distribution of the eigenvalues of a 3 by 3 Wishart matrix. *Ann. Math. Statist.*, **43**, 1498–1506.

BINGHAM, M. S. and MARDIA, K. V. (1975). Maximum likelihood characterization of the von Mises distribution. In *Statistical Distributions in Scientific Work* (G. P. Patil, S. Kotz and J. K. Ord, eds), Vol. 3. Boston: D. Reidel.

BONEVA, L. I., KENDALL, D. G. and STEFANOV, I. (1971). Spline transformations: three new diagnostic aids for the statistical data-analyst. *J. R. Statist. Soc.* B, **33**, 1–70.

COX, D. R. and LEWIS, P. A. W. (1966). *Statistical Analysis of Series of Events*. London: Methuen.

CROUCH, P., WEISS, R. and GOLDSTEIN, H. (1975). Orientation of cancer cells on gridlines. (Unpublished.)

DOWNS, T. D. (1972). Orientation statistics. *Biometrika*, **59**, 665–675.

—— (1974). Rotational angular correlations. In *Biorhythms and Human Reproduction* (M. Ferin *et al.*, eds), pp. 97–104. New York: Wiley.

FISHER, R. A. (1953). Dispersion on a sphere. *Proc. Roy. Soc. Lond.* A, **217**, 295–305.

FORT, ANN and MILLS, J. N. (1970). Fitting sine curves to 24 h urinary data. *Nature*, **226**, 657–658.

HAJEK, J. and SIDAK, Z. (1967). *Theory of Rank Tests*. New York: Academic Press.

HARTMANN, P. and WATSON, G. S. (1974). "Normal" distributions functions on spheres and the modified Bessel functions. *Ann. Prob.*, **2**, 593–607.

JAMES, A. T. (1964). Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.*, **35**, 475–501.

JONES, T. A. and JAMES, W. R. (1969). Analysis of bimodal orientation data. *Math. Geol.*, **1**, 129–135.

KAGAN, A. M., LINNIK, YU. V. and RAO, C. R. (1973). *Characterization Problems in Mathematical Statistics*. New York: Wiley.

KENDALL, D. G. (1974a). Pole seeking Brownian motion and bird navigation. *J. R. Statist. Soc.* B, **36**, 365–417.

—— (1974b). Hunting quanta. *Phil. Trans. R. Soc. Lond.*, A, **276**, 231–266.

KLUYVER, J. C. (1906). A local probability theorem. *Ned. Akad. Wet. Proc.* A, **8**, 341–350.

KSHIRSAGAR, A. M. (1972). *Multivariate Analysis*. New York: Marcel Dekker.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. New York: Wiley.

LEWIS, P. A. W. (1972). Recent results in the statistical analysis of univariate point processes. In *Stochastic Point Processes*, pp. 1–54, (P.A.W. Lewis, ed.), New York: Wiley

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*, the M. S. Bartlett Volume (J. Gani, ed.), Academic Press, London.
LYTTLETON, R. A. (1953). *The Comets and their Origin.* Cambridge University Press.
—— (1961). On the statistical loss of long-period comets from the solar system, I. *Proc. 4th Berkeley Symp. Math. Statist. and Prob.*, 3, 229–244.
MARDIA, K. V. (1967). A non-parametric test for the bivariate two sample location problem. *J. R. Statist. Soc.* B, **29**, 320–342.
—— (1969). On Wheeler and Watson's test on a circle. *Sankhyā*, A, **31**, 177–190.
—— (1972a). *Statistics of Directional Data.* London: Academic Press.
—— (1972b). A multi-sample uniform scores test on a circle and its parametric competitor. *J. R. Statist. Soc.* B, **34**, 102–113.
—— (1975a). Characterizations of directional distributions. In *Statistical Distributions for Scientific Work* (G. P. Patil, S. Kotz and J. K. Ord, eds), Vol. 3. Boston: D. Reidel.
—— (1975b). Distribution theory for the von Mises–Fisher distribution. In *Statistical Distributions for Scientific Work* (G. P. Patil, S. Kotz and J. K. Ord, eds), Vol. 1. Boston: D. Reidel.
MARDIA, K. V. and EL-ATOUM, S. A. M. (1974). Bayesian inference for the von Mises–Fisher distribution. (Submitted to *Biometrika*.)
MARDIA, K. V. and SPURR, B. D. (1973). Multi-sample tests for multi-modal and axial data. *J. R. Statist. Soc.* B, **35**, 422–436.
MARDIA, K. V. and SUTTON, T. (1975). A mixture of two von Mises distributions. To appear in *Biometrika*.
MARDIA, K. V. and ZEMROCH, P. J. (1975a). Algorithm AS 81. Circular statistics. *Appl. Statist.*, **24**, 147–150.
—— (1975b). Algorithm AS 80. Spherical statistics. *Appl. Statist.* **24**, 144–146.
MATTHEWS, G. V. T. (1974). On bird navigation, with some statistical undertones. *J. R. Statist. Soc.* B, **36**, 349–364.
MORRIS, J. E. and LAYCOCK, P. J. (1974). Discriminant analysis of directional data. *Biometrika*, **61**, 335–342.
PEARSON, E. S. and HARTLEY, H. O. (1972). *Biometrika Tables for Statisticians*, Vol. 2. Cambridge University Press.
RAO, J. S. (1969). Some contributions to the analysis of circular data. Ph.D. Thesis, Indian Statistical Institute, Calcutta.
ROTHMAN, E. D. (1971). Tests of coordinate independence for a bivariate sample on a torus. *Ann. Math. Statist.*, **42**, 1962–1969.
RUKHIN, A. L. (1971). Some statistical decisions about distributions on a circle for large samples. *Sankhyā*, A, **34**, 243–250.
STEPHENS, M. A. (1969). Techniques for directional data. *Tech. Report No.* 150. Dept. of Statist., Stanford University.
TYROR, J. G. (1957). The distribution of the directions of perihelion of long-period comets. *Mon. Not. R. Astr. Soc.*, **117**, 369–379.
UPTON, G. J. G. (1974). New approximations to the distribution of certain angular statistics. *Biometrika*, **61**, 369–374.
WATSON, G. N. (1948). *A Treatise on the Theory of Bessel Functions.* 2nd ed. Cambridge Univ. Press.
WATSON, G. S. and WILLIAMS, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika*, **43**, 344–352.
WHEELER, S. and WATSON, G. S. (1964). A distribution-free two-sample test on a circle. *Biometrika*, **51**, 256–257.

DISCUSSION OF PROFESSOR MARDIA'S PAPER

Dr P. J. LAYCOCK (University of Manchester Institute of Science and Technology): I am delighted to propose the vote of thanks to Professor Mardia for the scholarly, yet practical, paper he has presented this evening. The paper represents an important step in the establishment of directional data analysis as a field of theoretical and practical importance in statistics.

There is little of a controversial nature in the paper, so I find myself somewhat at a loss for the traditional collection of "pointed" remarks with which to assail the speaker—despite the title of his paper!