Orientation Statistics Without Parametric Assumptions

Author(s): Michael J. Prentice

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, 1986, Vol. 48, No. 2 (1986), pp. 214-222

Published by: Wiley for the Royal Statistical Society

Stable URL: https://www.jstor.org/stable/2345716

J. R. Statist. Soc. B (1986)
48, *No.* 2, *pp.* 214–222

Orientation Statistics without Parametric Assumptions

By MICHAEL J. PRENTICE

University of Edinburgh, Scotland

[Received July 1985. Revision January 1986]

SUMMARY

Maximum likelihood estimation using the matrix von Mises–Fisher distribution in orientation statistics leads to unacceptably complicated likelihood equations, essentially because of the inconvenient form of the normalizing constant in the probability distribution. For the case of 3×2 or 3×3 orientations, the main cases of practical importance, an alternative approach is developed here by transforming the data into unsigned four-dimensional directions and using known results on the sampling properties of the spectral decomposition of the resulting sample moment of inertia matrix. It is demonstrated that the necessary computations are relatively simple by applying some of the techniques to a set of vectorcardiogram data.

Keywords: DIRECTIONAL DATA; MATRIX FISHER DISTRIBUTION; BINGHAM DISTRIBUTION; WATSON DISTRIBUTION; RANDOM ORTHOGONAL MATRIX; SAMPLING PROPERTIES OF PRINCIPAL COMPONENTS; VECTOCARDIOGRAPHY

1. INTRODUCTION

Parametric statistical methods for orientation statistics have been developed by for example Downs (1972), Khatri and Mardia (1977), Jupp and Mardia (1979) and Prentice (1982). Almost all of the methodology has been developed for the completely general case of m-frames in p dimensions, $m \leqslant p$, although for almost every practical application, p is at most 3. In this paper an alternative approach to statistical inference on the sample space $SO(3)$ of 3×3 rotation matrices is suggested. After a simple transformation not widely known among statisticians, the analysis is developed on H_4, the space of unsigned directions (axes) in four dimensions. Distribution-free procedures developed by Prentice (1984) are used to produce computationally convenient alternatives to the conventional parametric procedures.

A numerical example first analysed by Downs (1972) is treated by this alternative approach. His 98 matched pairs of vectocardiogram orientations obtained using two different lead systems on the same patients (Downs *et al.*, 1974) were transformed as in Section 2 into four-dimensional axial data. Significant differences were found between children aged 2–10 and those aged 11–19, but there were no significant differences between the sexes. Matched pair procedures for 3×3 orientations are also developed to demonstrate a significant difference between the orientations obtained from the two different lead systems.

2. UNIFORMLY DISTRIBUTED RANDOM ROTATIONS IN R^3

Miles (1965) has described a well-known parameterisation of random rotations in R^3. For present purposes, however, the following representation of $SO(3)$, the 3-dimensional rotation group, is more convenient. If $x = x_4 + ix_1 + jx_2 + kx_3$ is a unit quaternion ($x^{-1} = \bar{x}$), then it may be used to represent a rotation in R^3 in the sense that if z in R^3 is represented as a pure quaternion $z = iz_1 + jz_2 + kz_3 (\bar{z} = -z)$, then $z_x = xzx^{-1}$ is also a pure quaternion with the same modulus. Note that $z_x = z_{-x}$ and that every rotation in R^3 may be expressed uniquely,

Present address: Dept of Statistics, The University, King's Buildings, Mayfield Rd, Edinburgh EH9 3JZ, Scotland.

 0035/9246/86/48214 $2.00

up to the sign change, in this form (Porteous, 1969, p. 180). We may write z_x in the coordinate form Xz where now $z = (z_1, z_2, z_3)'$ and

$$X = \mu(x) = \begin{pmatrix} x_1^2 + x_4^2 - x_2^2 - x_3^2, & 2(x_1x_2 - x_3x_4), & 2(x_2x_4 + x_1x_3) \\ 2(x_3x_4 + x_1x_2), & x_2^2 + x_4^2 - x_1^2 - x_3^2, & 2(x_2x_3 - x_1x_4) \\ 2(x_1x_3 - x_2x_4), & 2(x_2x_3 + x_1x_4), & x_3^2 + x_4^2 - x_1^2 - x_2^2 \end{pmatrix}, \tag{2.1}$$

with trace $4x_4^2 - 1$, is a 3×3 rotation matrix with det $X = +1$ (Moran, 1976, p. 297, Prentice, 1978, p. 174). Also $\mu(x) = \mu(-x)$ and x is uniformly distributed on H_4 if and only if $X = \mu(x)$ is uniformly distributed on $SO(3)$. If $X = \mu(x)$ then $(\pm)x$ uniquely determines X and conversely.

Consider a random orientation of three mutually orthogonal directions in R^3, obeying the right-hand rule, such as a *QRS* loop described in Downs *et al.* (1974). A full description of the geometry of general orientation statistics is given in the Introductions to Khatri and Mardia (1977) and Downs (1972). Such a random orientation may be represented as a matrix X in $SO(3)$, where the rows of X specify the three mutually orthogonal directions. If H is any other rotation matrix, then $Y = XH'$ represents another orientation which is obtained from X by rotating all three rows of X through an angle $2\cos^{-1}(h_4)$ about the axis $(h_1, h_2, h_3)/(h_1^2 + h_2^2 + h_3^2)^{1/2}$, where $H = \mu(h)$. Every pair of orientations X, Y are related by such a transformation since choosing $H = Y'X$ transforms X into Y and $H = X'Y$ transforms Y into X. Note also that

$$\operatorname{tr}(XH') = 4(h'x)^2 - 1. \tag{2.2}$$

Also, if $w = w_4 + iw_1 + jw_2 + kw_3$ is any quaternion and u and v are unit quaternions then $w_{uv} = uwv$ has the same modulus as w, and (u, v) may be used to parameterize rotations in R^4 in the sense that every rotation in R^4 may be expressed uniquely in this form (Porteous, 1969, p. 182). In coordinate form we may write $w_{uv} = A(u, v)w$ where now $w = (w_1, w_2, w_3, w_4)'$ and $A(u, v)$ is a 4×4 rotation matrix, uniformly distributed on $SO(4)$ if and only if u and v are independently uniformly distributed on the unit sphere in R^4, but with (u, v) and $(-u, -v)$ identified (Moran, 1975, p. 300). Also if $X = \mu(x)$, $Y = \mu(y)$ and $y = A(u, v)x$ then

$$Y = UXV, \tag{2.3}$$

where $U = \mu(u)$ and $V = \mu(v)$. This link between parameterizations of $SO(3)$ and $SO(4)$ is useful for obtaining simple expressions for the relationships between the parameters of the distributions (3.1) and (3.2) below. I am grateful to a referee for reminding me of (2.3) and thus simplifying the presentation of Section 3.

3. THE 4-DIMENSIONAL BINGHAM DISTRIBUTION AND THE 3×3 MATRIX FISHER DISTRIBUTION

It will be shown in Section 4 that if X is a random 3×3 rotation matrix with an unspecified probability distribution, then any statistical analysis may be carried out by consideration of the relationship $X = \mu(x)$, as in (2.1), where x is a random 4-dimensional unsigned direction. To begin, consider the conventional parametric formulation of the one-sample inference problem for orientation statistics (Khatri and Mardia, 1977). It will be shown that X has a matrix Fisher distribution if and only if $x = \mu^{-1}(X)$ has a Bingham distribution, and the relationships between the various parameters will be clarified. In a slightly amended version of the notation of Khatri and Mardia (1977), suppose X has the matrix Fisher density

$$\{{}_0F_1(\tfrac{3}{2}; \tfrac{1}{4}D_\phi^2)\}^{-1} \exp\{\operatorname{tr}(FX')\} \, [dX] \tag{3.1}$$

where $[dX]$ represents the uniform distribution on $SO(3)$, ${}_0F_1(\tfrac{3}{2}; \tfrac{1}{4}D_\phi^2)$ is a normalising constant and F has singular value decomposition $F = \Delta D_\phi \Gamma$, where Δ and Γ are rotation matrices and $D_\phi = \operatorname{diag}(\phi_1, \phi_2, \phi_3)$ is definite with $|\phi_1| < |\phi_2| < |\phi_3|$. We impose the same uniqueness conditions as Khatri and Mardia (1977, p. 96) except that if then $\det(\Delta) = -1$ or $\det(\Gamma) = -1$ then Δ or Γ is replaced by $-\Delta$ or $-\Gamma$ respectively, and if exactly one of these sign changes is carried out then D_ϕ is negative definite. The difference between this notation and that of Khatri and Mardia is that they allow $\det(\Delta) = -1$ or $\det(\Gamma) = -1$, but require D_ϕ positive definite.

In all examples considered here, the maximum likelihood estimate of D_ϕ and its moment analogue D_h are positive, so that this apparently artificial distinction is actually irrelevant for the vectorcardiogram data analysed in Section 6. If $F = KM$, $M = \Delta\Gamma$, $K = \Delta D_\phi \Delta'$, where M is the polar component and K the elliptical component of F, then again it is assumed here that M is a rotation matrix so K is definite, not necessarily positive definite. Also, the density (3.1) has a maximum at M if D_ϕ is positive definite, but a minimum at M if D_ϕ is negative definite. Conventional maximum likelihood procedures (Khatri and Mardia, 1977) for a random sample $X_1, \ldots, X_n$ from the density (3.1) lead to consideration of the singular value decomposition of

$$\bar{X} = n^{-1} \sum_{i=1}^{n} X_i = \hat{\Delta} D_g \hat{\Gamma},$$

where here again D_g may in principle be negative definite if we require $\hat{\Delta}$ and $\hat{\Gamma}$ to be rotation matrices. It will be shown in Section 4 that D_g is a natural estimator of D_h, where $E(X) = \Delta D_h \Gamma$.

Let x be a random four-dimensional unsigned unit direction, distributed as $B_4(L, A)$, so x has the Bingham density

$$\{{}_1F_1(\tfrac{1}{2}; 2; L)\}^{-1} \exp\{\operatorname{tr}(LAxx'A')\}\ [dx] \tag{3.2}$$

where $[dx]$ represents the uniform distribution on H_4, ${}_1F_1(\frac{1}{2}; 2; L)$ is a normalizing constant, A is a 4×4 rotation matrix, so $\det(A) = +1$, and $L = \operatorname{diag}(l_1, l_2, l_3, l_4)$ with either $l_1 < l_2 < l_3 < l_4$ or $l_1 > l_2 > l_3 > l_4$. Note that $B_4(L + \alpha I_4, A)$ and $B_4(L, A)$ are indistinguishable, for arbitrary scalar α (Bingham, 1974, Lemma 2.1). For reasons soon to be made apparent, we choose to say that x has a *bipolar* Bingham distribution (or L is bipolar) if $l_1 + l_4 > l_2 + l_3$, and that x has an *equatorial* Bingham distribution if $l_1 + l_4 < l_2 + l_3$. The intermediate case with $l_1 + l_4 = l_2 + l_3$ corresponds to the case rank $(D_\phi) = 2$ and is of no immediate interest here. The motivation for this nomenclature is the (degenerate) special case of rotational symmetry (Prentice, 1984, Section 5); if $l_1 = l_2 = l_3 < l_4$ then x has a bipolar Watson density (Watson, 1965), and if $l_1 = l_2 = l_3 > l_4$ then x has an equatorial Watson density. In general, it is convenient to label the eigenvalues so that $l_1 < l_2 < l_3 < l_4$ in the bipolar case, and $l_1 > l_2 > l_3 > l_4$ in the equatorial case.

If $X = \mu(x)$ is a 3×3 rotation matrix, define $Q(X)$ by

$$4Q(X) = 4xx' - I_4 \tag{3.3}$$

so $\operatorname{tr}\{Q(X)\} = 0$, and denote the columns of I_3 by $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, and those of I_4 by (e_1, e_2, e_3, e_4). If $E_t = \mu(e_t)$, $t = 1, 2, 3, 4$, then $E_4 = I_3$, $E_t = 2\varepsilon_t\varepsilon_t' - I_3$, $t = 1, 2, 3$, and $E_1 + E_2 + E_3 + E_4 = 0$. If F is any real 3×3 matrix with singular value decomposition $F = \Delta D_\phi \Gamma$ where $D_\phi = \operatorname{diag}(\phi_1,\ \phi_2,\ \phi_3)$ then we may write $4F = \Sigma l_t \Delta E_t \Gamma$ $(1 \leqslant t \leqslant 4)$ uniquely, where $l_1 = \phi_1 - \phi_2 - \phi_3$, $l_2 = \phi_2 - \phi_1 - \phi_3$, $l_3 = \phi_3 - \phi_1 - \phi_2$ and $l_1 + l_2 + l_3 + l_4 = 0$. Also, since $4E_t = 3E_t - \Sigma E_s$, $s \neq t$, it follows that if we define

$$4Q(F) = \sum_{t=1}^{4} l_t Q(\Delta E_t \Gamma) \tag{3.4}$$

then (3.4) agrees with (3.3) on $SO(3)$. Lastly, if $\Delta E_t \Gamma = \mu(a_t)$, $t = 1, 2, 3, 4$ then a_1, a_2, a_3, a_4 are mutually orthogonal from (2.2), since $\operatorname{tr}\{\mu(a_s)\mu'(a_t)\} = -1$ if $s \neq t$. Hence we may write $4Q(F) = ALA'$ where $L = \operatorname{diag}(l_1, l_2, l_3, l_4)$ has zero trace, and $A = (a_1, a_2, a_3, a_4)$ is in $SO(4)$.

Note that $4Q(D_\phi) = L$ and

$$4\,\mathrm{tr}(FX') = \sum_{t=1}^{4} l_t\,\mathrm{tr}(\Delta E_t \Gamma X'),$$

or from (2.2),

$$\mathrm{tr}(FX') = x'ALA'x, \tag{3.5}$$

if $X = \mu(x)$, $F = \Delta D_\phi \Gamma$, $4Q(D_\phi) = L$, and $A = (a_1, a_2, a_3, a_4)$ with $\mu(a_t) = \Delta E_t \Gamma$, $t = 1, 2, 3, 4$. It follows that $X = \mu(x)$ has the matrix Fisher distribution (3.1) if and only if x has the Bingham distribution (3.2), where Δ, D_ϕ, Γ, L and A are related as in (3.5). Note that $\mu(a_t)\mu(a_4)' = 2\delta_t\delta_t' - I_3, t = 1, 2, 3$ where $\Delta = \pm(\delta_1, \delta_2, \delta_3)$ with the sign chosen to ensure that Δ is a rotation, that $4D_\phi = \mathrm{diag}(l_4 + l_1 - l_2 - l_3,\ l_4 + l_2 - l_1 - l_3,\ l_4 + l_3 - l_1 - l_2)$, that $\Gamma = \Delta'\mu(a_4)$, and that $\phi_1 < \phi_2 < \phi_3$ if and only if $l_1 < l_2 < l_3 < l_4$ whereas $\phi_1 > \phi_2 > \phi_3$ if and only if $l_1 > l_2 > l_3 > l_4$. Also D_ϕ is positive (negative) definite if and only if L is bipolar (equatorial).

From known results (Khatri and Mardia, 1977, Prentice, 1984) on maximum likelihood estimation for one random sample of data distributed as (3.1) or (3.2), some closely related equivalences obtain. Given a random sample $X_1, \ldots, X_n$ from (3.1) with sample mean $\bar{X} = n^{-1}\Sigma X_i$ $(1 \leqslant i \leqslant n)$ having singular value decomposition $\bar{X} = \hat{\Delta} D_g \hat{\Gamma}$, where $D_g = \mathrm{diag}(g_1, g_2, g_3)$, it is known that $\hat{\Delta}$ and $\hat{\Gamma}$ are maximum likelihood estimates of Δ and Γ respectively. Also, if $X_i = \mu(x_i)$, $1 \leqslant i \leqslant n$, and $T = n^{-1}\Sigma x_i x_i'$ has spectral decomposition $T = \hat{A}\hat{\Lambda}\hat{A}'$, then $\hat{A}$ is the maximum likelihood estimate of A. Hence the estimated location parameters $\hat{\Delta}$, $\hat{\Gamma}$, $\hat{A}$ are related in the same way as the true location parameters Δ, Γ, A in (3.5), and this is so *whether or not* the underlying distribution is Bingham in the axial domain, or equivalently matrix Fisher on $SO(3)$. Also if the Bingham assumption is true the maximum likelihood estimates of the elements of D_ϕ are monotonic functions of the elements of D_g (Khatri and Mardia, 1977, (4.3)); and, equivalently, the same is true of the maximum likelihood estimates of the elements of L in relation to the elements of $\hat{\Lambda}$.

4. GENERAL PROBABILITY DISTRIBUTIONS FOR 3 × 3 ORIENTATIONS AND 4-DIMENSIONAL AXES

Suppose that x is a random variable on H_4, or equivalently a random 4-dimensional unsigned direction with some unspecified antipodally symmetric distribution on S_4, the surface of the unit sphere in R^4. Clearly, all odd-order moments of x about the origin 0 are zero, and all even-order moments are finite. Where necessary it will be assumed that x has almost everywhere a probability density and that $E(xx') = \Omega$ is positive definite. Unless otherwise stated the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ of Ω are assumed distinct so that the spectral decomposition $\Omega = A\Lambda\Omega'$ is unique, with $\mathrm{tr}(\Omega) = \mathrm{tr}(\Lambda) = 1$.

Also of interest for present purposes is the matrix of fourth moments of x, which has a relatively simple form after rotation to principal axes. The random variable $x^* = A'x$ has variance matrix $E(x^*x^{*\prime}) = \Lambda$ and the 16×16 matrix $C^* = E(x^*x^{*\prime} \times x^*x^{*\prime})$ of fourth order moments of the principal components has elements $c_{ij,kl}$ which are assumed to take the value zero unless i, j, k, l take at most two distinct values with even multiplicities. This is certainly the case if the probability distribution of $(\pm x_1^*, \pm x_2^*, \pm x_3^*, \pm x_4^*)$ is the same for all 16 choices of signs, for example the Bingham distribution (3.2). Hence consideration of fourth moments is restricted to $C = E(x^{(2)}x^{(2)\prime})$ where $x^{(2)}$ is the 4-vector of diagonal entries in $x^*x^{*\prime}$, that is the squares of the elements of x^*. Clearly $V = \mathrm{var}(x^{(2)}) = C - \lambda\lambda'$, where $\lambda = \mathrm{vec}(\Lambda)$, the diagonal elements of Λ written as a 4-vector. It is assumed throughout that all elements of C are strictly positive and that V is positive semi-definite with rank 3; note that all row sums of V are zero.

Suppose now that X is a random 3×3 rotation matrix, an orientation statistic, and that $X = \mu(x)$, as in (2.1). Then it is straightforward to demonstrate certain relationships between the moments of X and x, and the sample summary statistics $\bar{X} = n^{-1}\Sigma X_i$ and $T =$

$n^{-1}\Sigma x_i x_i'$, $1 \leqslant i \leqslant n$, with X_i related to x_i as in (2.1). Again, where possible the notation of Khatri and Mardia (1977) and Prentice (1984) will be used. If $E(xx') = \Omega = A\Lambda A'$ with $\mathrm{tr}(\Lambda) = 1$, then for the same reasons advanced when considering the relationship between the densities (3.1) and (3.2), we choose to say that Ω (or Λ) is bipolar if $\lambda_1 + \lambda_4 > \frac{1}{2}$, in which case we adopt the convention $\lambda_1 < \lambda_2 < \lambda_3 < \lambda_4$, and that Ω (or Λ) is equatorial if $\lambda_1 + \lambda_4 < \frac{1}{2}$, when we assume $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$. The motivation for this nomenclature is as before, although it should be noted that if the distribution of x is actually Bingham (3.2), there is not a 1-1 correspondence between bipolar densities and bipolar covariance matrices, because the elements of L are monotonic but non-linear functions of the elements of Λ. Similarly, given a random sample $x_1, \ldots, x_n$ with $T = n^{-1}\Sigma x_i x_i'$, $1 \leqslant i \leqslant n$, having spectral decomposition $T = \hat{A}\hat{\Lambda}\hat{A}'$, we may choose to say that T (or $\hat{\Lambda}$) is bipolar or equatorial according as $\hat{\lambda}_1 + \hat{\lambda}_4 > \frac{1}{2}$ or $\hat{\lambda}_1 + \hat{\lambda}_4 < \frac{1}{2}$; again we take $\hat{\lambda}_1 < \hat{\lambda}_2 < \hat{\lambda}_3 < \hat{\lambda}_4$ in the bipolar case and $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 > \hat{\lambda}_4$ in the equatorial case.

It has already been remarked that $\hat{\Delta}$, $\hat{\Gamma}$, A bear the same relationship to each other as Δ, Γ, A in (3.6). It remains to explicate the relationship between D_g and $\hat{\Lambda}$, and between $D_h = \Delta' E(X)\Gamma'$ and Λ. If $\bar{X} = \hat{\Delta} D_g \hat{\Gamma}$ then $\mathrm{tr}(\bar{X}Z') = 4(z'Tz) - 1$, where $Z = \mu(z)$ is any rotation matrix, from (2.2), so $\mathrm{tr}(\bar{X}Z') = 4z'\hat{A}'\hat{\Lambda}\hat{A}'z - 1$ and $4Q(D_g) = 4\hat{\Lambda} - I_4$, for the same reason that $4Q(D_\phi) = L$ in the discussion after (3.4). Hence $4(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4) = (1 + g_1 - g_2 - g_3, 1 + g_2 - g_1 - g_3, 1 + g_3 - g_1 - g_2, \quad 1 + g_1 + g_2 + g_3)$ and $(g_1, \quad g_2, \quad g_3) = (\hat{\lambda}_4 + \hat{\lambda}_1 - \hat{\lambda}_3 - \hat{\lambda}_2, \hat{\lambda}_4 + \hat{\lambda}_2 - \hat{\lambda}_1 - \hat{\lambda}_3, \hat{\lambda}_4 + \hat{\lambda}_3 - \hat{\lambda}_1 - \hat{\lambda}_2)$. Also since $E(\bar{X}) = \Delta D_h \Gamma$ and $E(T) = A\Lambda A'$, the relationship between $D_h = \mathrm{diag}(h_1, h_2, h_3)$ and Λ is the same as that between D_g and $\hat{\Lambda}$.

Thus if the samples $x_1, \ldots, x_n$ and $X_1, \ldots, X_n$ are functionally related by $X_i = \mu(x_i)$ then $\hat{M}$, the polar component of $\bar{X}$, is related by (2.1) to the eigenvector $\hat{m} = \hat{a}_4$ corresponding to an eigenvalue $\hat{\lambda}_4$ of T. If $\hat{\lambda}_1 + \hat{\lambda}_4 > \frac{1}{2}$ then $\hat{M} = \mu(\hat{m})$ where $\hat{m}$ corresponds to the largest root of T, and if $\hat{\lambda}_1 + \hat{\lambda}_4 < \frac{1}{2}$ then $\hat{m}$ corresponds to the smallest root.

5. ESTIMATION AND HYPOTHESIS TESTING WITHOUT PARAMETRIC ASSUMPTIONS

Consider a random sample $X_1, \ldots, X_n$ on $SO(3)$, or equivalently $x_1, \ldots, x_n$, where $X_i = \mu(x_i)$ on H_4 as in Section 4. Let $T = \hat{A}\hat{\Lambda}\hat{A}'$, so from Prentice (1984) with $p = 4$, in large samples

$$n^{1/2}(\hat{\lambda} - \lambda) \to N_4(0, V) \text{ in distribution}$$

and

$$n^{1/2}(\hat{a}_k - a_k) \to N_4(0, F_k) \text{ in distribution, } k = 1, 2, 3, 4,$$

where the exact form of F_k is given in Prentice (1984, (3.2)). If we assume that $\Omega = E(xx')$ is very bipolar ($\lambda_4 \gg \lambda_3 > \lambda_2 > \lambda_1$) then interest centres upon $m = a_4$, $\hat{m} = \hat{a}_4$ and $\hat{M} = \mu(\hat{m})$, the natural estimate of M, the polar component of $E(X)$. An approximate $(1 - \alpha)$ confidence region for m is given by

$$H = nm'\hat{F}_4^- m \leqslant \chi^2_{3,\alpha} \tag{5.1}$$

where $\hat{F}_4^- = \hat{A}_{(4)} \mathrm{diag}[(\hat{\lambda}_4 - \hat{\lambda}_k)^2/\hat{c}_{4k}]\hat{A}'_{(4)}$, and $\hat{A}_{(4)} = (\hat{a}_1, \hat{a}_2, \hat{a}_3)$. This is a much simpler expression than (5.8) of Khatri and Mardia (1977). Similarly, an approximate $(1 - \alpha)$ confidence region for Λ is given by

$$n(\hat{\lambda} - \lambda)'\hat{V}^-(\hat{\lambda} - \lambda) \leqslant \chi^2_{3,\alpha}$$

where $\hat{V}^-$ is the Moore-Penrose inverse of $\hat{V}$, and an approximate $(1 - \alpha)$ confidence interval for λ_4 is $\hat{\lambda}_4 \pm (V_{44}/n)^{1/2} Z_{\alpha/2}$ where $Z_{\alpha/2}$ is the upper $\frac{1}{2}\alpha$ point of $N_1(0, 1)$. Also, a simple approximate test of bipolarity against equatoriality (whether D_h is positive definite or negative definite) may be constructed from

$$(d'\hat{V}d)^{-1/2} d'(\hat{\lambda} - \lambda) \to N_1(0, 1) \text{ in distribution,} \tag{5.2}$$

where $d = (1, -1, -1, 1)$. In terms of the original matrix variables, (5.1) is expressible as

$$(n/8) \sum_{k=1}^{3} \{\mathrm{tr}(D_g] - g_k\}^2 \{1 + \mathrm{tr}(M'\hat{A}_k]\}/\hat{C}_{4k} \leqslant \chi^2_{3,\alpha}$$

which is rather less convenient computationally.

A joint $(1 - \alpha)$ confidence region for Γ and Δ together may be constructed similarly using Prentice (1984, (3.9)). Any particular (Γ, Δ) corresponds to an essentially unique A so the set of (Γ, Δ) such that $\frac{1}{2} n \Sigma a'_k \hat{F}_k^- a_k \leqslant \chi^2_{6,\alpha}$, $1 \leqslant k \leqslant 4$, is a $(1 - \alpha)$ joint confidence region for (Γ, Δ).

Consider now the special case of spherical symmetry in the matrix domain. The random 3×3 rotation matrix X in Section 3 has a spherically symmetric first moment matrix if the elliptical component of $E(X)$ is of the form $K = \beta I_3$. From the relationships between D_h and Λ it follows that X has a spherically symmetric expectation if and only if x has a rotationally symmetric variance matrix, as in Section 5 of Prentice (1984). Although the hypothesis of spherical symmetry is of no particular interest, we present here some test procedures because of the relationship with rotational symmetry on H_4, as problems of ill-conditioning of estimates may arise if the data are compatible with rotational symmetry on H_4 (see section 6). Suppose that it is known whether Ω is bipolar or equatorial, and that the axis of symmetry is m corresponding to the largest (or smallest) eigenvalue λ of Ω. Given their sample analogues $\hat{m}$, $\hat{\lambda}$, clearly the natural estimate of β is $\hat{\beta} = (4\hat{\lambda} - 1)/3 = (g_1 + g_2 + g_3)/3 = \bar{g}$, positive if bipolar, and negative if equatorial. Also from Prentice (1984, (5.9)), an approximate $(1 - \alpha)$ confidence region for the axis of symmetry m is given by

$$n\hat{\beta}^2\{1 - (m'\hat{m})^2\}/\hat{c}_{12} \leqslant \chi^2_{3,\alpha} \tag{5.3}$$

where $\hat{c}_{12} = (\hat{\lambda} - \hat{c}_{11})/3$, or equivalently,

$$n\hat{\beta}^2\{3 - \mathrm{tr}(M'\hat{M})\}/(4\hat{c}_{12}) \leqslant \chi^2_{3,\alpha}.$$

Also, from Prentice (1984, (5.10)),

$$(\hat{\lambda} - \lambda)\{(n-1)/(\hat{c}_{11} - \hat{\lambda}^2)\}^{1/2} \rightarrow N_1(0, 1) \text{ in distribution}$$

or equivalently, $0.75(\hat{\beta} - \beta)\{(n-1)/(\hat{c}_{11} - \hat{\lambda}^2)\}^{1/2} \rightarrow N_1(0, 1)$ in distribution so confidence intervals for λ or β are easy to compute.

Finally, the natural test statistic for assessing the hypothesis of spherical symmetry on $SO(3)$, or equivalently rotational symmetry on H_4 is, from Prentice (1984, (6.1))

$$R = 15n[\mathrm{tr}(T^2) - \hat{\lambda}^2 - (1 - \hat{\lambda})^2/3]/(2\hat{c}_{12}) \tag{5.4}$$

or equivalently $R = (1.875n/\hat{c}_{12})\Sigma(g_k - \bar{g})^2$, $1 \leqslant k \leqslant 3$, distributed in large samples as χ^2_5.

6. APPLICATIONS

A data set of 98 matched pairs of vectorcardiogram orientations, described at length in Downs *et al.* (1974) is by now a classic data set analysed in part by Downs (1972) and Khatri and Mardia (1977). The data consist of 98 pairs of 3×2 orientations obtained with two different lead systems, those of Frank and McFee, from normal young people aged 2–19. It is convenient to subclassify the data by age and sex; there are 17 girls aged 2–10, 25 girls aged 11–19, 28 boys aged 2–10 and 28 boys aged 11–19. Interest naturally centres on any systematic differences between the sexes and also between the young (aged 2–10) and the juveniles (11–19). It is also of interest to describe any systematic differences between the results obtained on the same individuals from the two different lead systems. Thus the development of paired comparisons procedures for orientations seems relevant here.

Like Downs (1972) and Khatri and Mardia (1977), we consider first the data obtained from the McFee lead system applied to the young males and young females. Khatri and Mardia (1977) performed some parametric analyses assuming that the data are matrix Fisher distributed, and Downs (1972, Section 5.2) essentially used trivariate normal approximations in the tangent 3-space to $SO(3)$ at the estimated pole $\hat{M}$. Here we extend each 3×2 data matrix to a uniquely determined 3×3 rotation matrix using the right hand rule, convert to 4-dimensional axial form, using (2.1), and apply the methods of Section 5, making no assumptions about the underlying probability distributions. All tests and confidence regions constructed are based on estimated second and fourth-order moments in the axial domain, or equivalently first and second-order moments in the matrix domain. The distributional results of Section 5 relate to large samples, and it is by no means clear that the sample sizes used here are large enough to validate the various χ^2 distributional assumptions claimed for the test statistics. However, many of the procedures used below give extremely large values of the statistics, so this question is unimportant for present purposes. The advantage of avoiding computation of confluent hypergeometric functions of matrix argument, and their derivatives, is obvious.

For the sample of 28 young males we obtain

$$\bar{X} = \begin{pmatrix} 0.687 & 0.551 & 0.122 \\ 0.575 & -0.737 & 0.142 \\ 0.183 & -0.05 & -0.863 \end{pmatrix}, \qquad \hat{\Delta} = \begin{pmatrix} 0.039 & 0.818 & 0.575 \\ 0.993 & 0.031 & -0.112 \\ -0.109 & 0.575 & -0.811 \end{pmatrix},$$

$$D_g = \text{diag}\,(0.873, 0.896, 0.947),$$

$$\hat{\Gamma} = \begin{pmatrix} 0.611 & -0.750 & 0.254 \\ 0.764 & 0.474 & -0.438) \\ 0.208 & 0.461 & 0.863 \end{pmatrix}, \quad \text{and} \quad \hat{M} = \hat{\Delta}\hat{\Gamma} = \begin{pmatrix} 0.768 & 0.623 & 0.148 \\ 0.607 & -0.782 & 0.142 \\ 0.204 & 0.019 & -0.979 \end{pmatrix}.$$

Equivalently, in the axial domain, the sample moment of inertia matrix is

$$T = \begin{pmatrix} 0.821 & 0.282 & 0.076 & -0.037 \\ & 0.100 & 0.034 & -0.015 \\ & & 0.047 & 0.006 \\ & & & -0.022 \end{pmatrix} = \hat{A}\hat{\Lambda}\hat{A}'$$

where $\hat{\Lambda} = \text{diag}\,(0.008, 0.019, 0.044, 0.929)$

$$\text{and} \quad \hat{A} = \begin{pmatrix} 0.230 & 0.220 & 0.127 & 0.939 \\ -0.810 & -0.463 & -0.148 & 0.328 \\ 0.328 & -0.213 & -0.915 & 0.094 \\ -0.427 & 0.831 & -0.352 & -0.004 \end{pmatrix}.$$

Note that $\hat{m} = \hat{a}_4$ is related to $\hat{M}$ by $\mu(\hat{m}) = \hat{M}$, as claimed. The data set is clearly bipolar ($\hat{\lambda}_4 \gg \hat{\lambda}_3$) and concentrated somewhere in the neighbourhood of $e_1 = (1, 0, 0, 0)'$. Equivalently, in the matrix domain the data are concentrated somewhere near $\mu(e_1) = \text{diag}(1, -1, -1)$. By tradition we first test the data for uniformity using the matrix Rayleigh statistic $3n\ \text{tr}(\bar{X}'\bar{X}) = 12n(\text{tr}(T^2) - 0.25)$, the Bingham statistic on H_4. We obtain 206.9, clearly incompatible with the χ^2_9 distribution. Also, since $\hat{g}_1 = 0.873$ has estimated standard error 0.0323 (from (5.2)), we conclude that Ω is significantly bipolar rather than equatorial and proceed to test $\hat{\Lambda}$ for rotational symmetry, or equivalently D_g for spherical symmetry. From (5.4), we obtain $R = 12.24$, significant at the 5% level when compared with the χ^2_5 distribution; we conclude that $K \neq \beta I_3$.

As an example of a one-sample test procedure we investigate the hypothesis that $M = \text{diag}(1, -1, -1)$. From (5.1) we obtain $H = 13.4$, very significant when compared with χ^2_3.

Similar analyses on the sample of 17 young females lead to broadly similar summary statistics. However, perhaps because the sample is small, these data appear compatible with the hypothesis of spherical symmetry, since using (5.4) gives $R = 3.59$. Thus $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$, and therefore $\hat{\Gamma}$ and $\hat{\Delta}$, are ill-determined, even though $\hat{m} = \hat{a}_4$ and thus $\hat{M} = \hat{\Delta}\hat{\Gamma} = \mu(\hat{m})$ are well-determined. For this sample, $m = (0.939, 0.332, 0.086, 0.001)'$ and

$$\hat{M} = \begin{pmatrix} 0.765 & 0.624 & 0.161 \\ 0.623 & -0.780 & -0.059 \\ 0.163 & 0.055 & -0.985 \end{pmatrix}.$$

It seems inappropriate to test for equality of a_1, a_2, a_3 in the two samples since the female sample exhibits rotational symmetry. We therefore test for equality of poles ($M_1 = M_2$) in the two samples using a modified version of (5.1), without assuming rotational symmetry in both samples, by calculating $\hat{F}_4^-$ from all 45 young individuals and referring

$$(n_1 + n_2)^{-1} n_1 n_2 (\hat{m}_1 - \hat{m}_2) \hat{F}_4^- (\hat{m}_1 - \hat{m}_2) \text{ to } \chi_3^2,$$

where $n_1 = 28$ and $n_2 = 17$. We obtain 1.07 and conclude that there are no systematic differences between the sexes in the age group 2–10. Similar calculations comparing all males ($n = 56$) with all females ($n = 42$) and all young children ($n = 45$) with all juveniles ($n = 53$) give 1.57 and 23.47. We conclude that there are significant differences between the age groups, but not between the sexes.

As an example of the use of (5.3) for constructing confidence regions for the pole M, consider all 42 females. The test statistic (5.4) is $R = 6.61$, compatible with χ_5^2, so the hypothesis of spherical symmetry seems reasonable. A 95% confidence region, based on (5.3), is the set of rotation matrices M such that $\operatorname{tr}(M'\hat{M}) \geqslant 2.989$, where

$$\hat{M} = \begin{pmatrix} 0.688 & 0.658 & 0.306 \\ 0.677 & -0.734 & 0.055 \\ 0.261 & 0.169 & -0.950 \end{pmatrix}.$$

Equivalently, M is in the 95% confidence region if $M'\hat{M}$ represents a rotation of at most 6° about any axis, since $2.989 = 1 + 2 \cos 6°$ approximately.

Finally, we construct paired comparisons procedures to clarify the systematic differences between the two lead systems. If $(X_1, Y_1), \ldots, (X_n, Y_n)$ are a random sample of matched pairs of rotation matrices, it seems appropriate to describe the differences between the lead systems in terms of $Z_1, \ldots, Z_n$ where $Z_i = X_i' Y_i$, since Z_i is the 3×3 rotation matrix which transforms X_i into Y_i. It is appropriate to investigate spherical symmetry and also whether $E(Z_i)$ has polar component $M_z = I_3$. Among the females the hypothesis of spherical symmetry was just about tenable, $R = 10.3$ significant at 10% but not at 5%, but among the males it was not ($R = 30.1$). The dominant sample eigenvalue $\hat{\lambda}_4$ was extremely large in all four subgroups and the corresponding eigenvectors were very similar. There were no significant differences between the vectors $\hat{m}_{(1)}$, $\hat{m}_{(2)}$, $\hat{m}_{(3)}$, $\hat{m}_{(4)}$ from the four subgroups ($H = 9.29$) comparing

$$H = \sum_{j=1}^{4} n_j \hat{m}'_{(j)} \hat{F}_4^- \hat{m}_{(j)}$$

with χ_9^2 where $\hat{F}_4^-$ was calculated from all 98 observations; for k samples the distribution would be $\chi^2_{3(k-1)}$. It seems appropriate to summarize the distribution of the Z_i by $(g_1, g_2, g_3) = (0.927, 0.937, 0.969)$ and $\hat{m}_z = (0.059, -0.135, 0.058, 0.987)'$, or equivalently,

$$\hat{M}_z = \mu(\hat{m}_z) = \begin{pmatrix} 0.957 & -0.132 & -0.261 \\ 0.100 & 0.986 & -0.133 \\ 0.275 & 0.101 & 0.956 \end{pmatrix}.$$

Hence the typical systematic discrepancy between the two lead systems is a rotation of about $\cos^{-1}\{\frac{1}{2}(\mathrm{tr}(\hat{M}_z) - 1)\} = 2\cos^{-1}(0.987) = 18°$ approximately, broadly similar to the description of the observed typical differences by Downs *et al.* (1974, p. 218). In remains to demonstrate that $\hat{M}_z$ is significantly different from I_3, or equivalently that $m_z \neq (0, 0, 0, 1)'$. Using (5.1) we obtain $H = 146.1$, not at all compatible with χ_3^2. We conclude that there are very significant differences between the results obtained from the two lead systems when used on the same patients.

REFERENCES

Bingham, C. (1974) An antipodally symmetric distribution on the sphere. *Ann. Statist.*, **2**, 1201–25.
Downs, T. D. (1972) Orientation statistics. *Biometrika*, **59**, 665–676.
Downs, T. D., Liebman, J. and Mackay, W. (1974) Statistical methods for vectorcardiogram orientations. *Proc. XIth International Vectorcardiogram Symposium*, pp. 216–222. Amsterdam: North-Holland.
Jupp, P. E. and Mardia, K. V. (1979) Maximum likelihood estimation for the matrix von Mises-Fisher and Bingham distributions. *Ann. Statist.*, **7**, 599–606.
Khatri, C. G. and Mardia, K. V. (1977) The von Mises-Fisher matrix distribution in Orientation Statistics. *J. R. Statist. Soc.* B, **39**, 95–106.
Miles, R. E. (1965) On random rotations in R^3. *Biometrika*, **52**, 636–639.
Moran, P. A. P. (1975) Quaternions, Haar measure and the estimation of a palaeomagnetic rotation. *Perspectives in Probability and Statistics* (J. Gani, ed.), pp. 295–301. Applied Probability Trust.
Porteous, I. R. (1969). *Topological Geometry*. Van Nostrand Reinhold.
Prentice, M. J. (1978). On invariant tests of uniformity for directions and orientations. *Ann. Statist.*, **6**, 169–176.
——(1982) Antipodally symmetric distributions for orientation statistics. *Stat. Planning and Inference*, **6**, 205–214.
——(1984) A distribution-free method of interval estimation for unsigned directional data. *Biometrika*, **71**, 147–154.
Watson, G. S. (1965) Equatorial distributions on a sphere *Biometrika*, **52**, 193–201.