The Annals of Probability
1978, Vol. 6, No. 4, 688–694

DENSITY-PRESERVING STATISTICS AND DENSITIES FOR SAMPLE MEANS

BY P. E. JUPP[1] AND K. V. MARDIA

University of Leeds

Among functions between manifolds those which transform probability distributions with densities to probability distributions with densities are characterised as almost submersions. As an application, conditions are given on sample size for sample means of distributions with density on Stiefel manifolds and Grassmannians to have densities.

1. Introduction and summary. Consider a probability distribution with a density on an n-dimensional submanifold M of $\mathbb{R}^p$ and suppose that M is not contained in any proper affine subspace. A natural conjecture is that for samples of size r the sample mean has a density if r is large enough. This is false and a counterexample is given in Section 3. One might next conjecture instead that if M is analytic then the sample mean has a density if $r \geqq p/n$. This too is false and counterexamples are given in Section 3.

An obvious question now is: which functions between manifolds transform probability distributions with densities to probability distributions with densities? These functions are characterised as almost submersions in Section 2. In Section 3 this result is applied to the case of sample means of probability distributions on Stiefel manifolds and on Grassmannians. In fact this work was motivated by particular distributions of practical importance on these manifolds in Khatri and Mardia (1977).

2. Almost submersions. Which functions from one manifold to another transform probability distributions with densities to probability distributions with densities? To make this question more precise, first recall that although a C^∞ manifold does not in general possess a distinguished probability measure, coordinate charts transfer Lebesgue measure to each coordinate neighbourhood. Further, as the coordinate transformations are C^∞, the concept of null set (of measure zero) is well defined.

DEFINITION. A probability distribution μ on a manifold M *has a density* (resp. *a C^r density*) if its restriction to each coordinate neighbourhood has a density (resp. a C^r density) with respect to Lebesgue measure. *μ has a density which is* C^r a.e. if there is a closed null set Z in M with $\mu(Z) = 0$ and such that the restriction of μ to the manifold $M \backslash Z$ has a C^r density.

Received May 18, 1977; revised August 23, 1977.

[1] Research carried out while this author was on a Research Fellowship funded by the Science Research Council.

AMS 1970 *subject classifications.* Primary 62H10, 62E15; Secondary 58A10.

Key words and phrases. Density-preserving statistic, almost submersion, sample mean, Stiefel manifold, Grassmannian.

DEFINITION. Let $t: M \to N$ be a function between manifolds. t is *density-preserving* (resp. C^r *density-preserving*, resp. a.e. C^r *density preserving*) if for every probability distribution on M which has a density (resp. a C^r density, resp. a density which is C^r a.e.), the transformed probability distribution on N has a density (resp. a C^r density, resp. a density which is C^r a.e.).

The simplest example of a density-preserving function is a projection $\mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p$ sending a probability distribution to its marginal distribution, so we may reasonably expect the most complicated example to be "locally a projection," i.e., a submersion. Recall that $t: M \to N$ is a submersion if the derivative of t has rank everywhere equal to the dimension of N. Brickell and Clark (1970) give the basic facts about submersions which we shall need.

DEFINITION. Let $t: M \to N$ be a C^r function, $r \geqq 1$. A point of M at which the derivative of t has rank less than the dimension of N is a *critical point* of t. C will denote the set of critical points of t. t is an *almost submersion* if C is null.

Now we can characterise the density-preserving functions. First recall that t is *proper* if the inverse image of every compact set is compact.

THEOREM 1. *Let $t: M \to N$ be a C^∞ function. Then t is density-preserving if and only if t is an almost submersion. If t is proper (so, in particular, if M is compact) then t is* a.e. *C^r density-preserving if and only if t is an almost submersion.*

PROOF. By Sard's theorem (Sternberg (1964), page 47), $t(C)$ is null. If t is not an almost submersion, C is not null, so there is a probability distribution on M which has a C^∞ density giving C positive probability. The transformed probability distribution then gives $t(C)$ positive probability, so cannot have a density. Thus t is not density-preserving.

Conversely, suppose that t is an almost submersion. We shall proceed by "integration along the fibre." An outline of this procedure in the setting of differential topology is given by Bott (1972), pages 14, 15. Consider four cases of increasing complexity.

CASE 1. t is the projection $\mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$. t is obviously density-preserving—indeed t is C^r density-preserving for densities with compact support (by a standard argument using uniform continuity and differentiation under the integral sign).

CASE 2. t is a submersion.

By a variant of the inverse function theorem, for each point x in M, there are coordinate charts for M round x and N round $t(x)$ with respect to which t has the local form of projection on the first $n = \dim N$ coordinates, as in Case 1 (Brickell and Clark (1970), Lemma 6.1.1). Intuitively, we now sum over coordinate neighbourhoods. More precisely, we cover M by such charts and take a subordinate partition of unity $\{\phi_\alpha\}$ (Brickell and Clark (1970), Section 3.4). Given a probability measure μ on M with a density each $\phi_\alpha \mu$ has a density and is contained in one of the coordinate neighbourhoods. By Case 1, the transformed measure $t_*(\phi_\alpha \mu) = (\phi_\alpha \mu)_0 t^{-1}$ has a density. Thus $t_* \mu = t_*(\sum_\alpha \phi_\alpha \mu) = \sum_\alpha t_*(\phi_\alpha \mu)$ has a density. As there may be infinitely many terms in the sum, nothing can be said about differentiability of this density.

CASE 3. t is an almost submersion.

Put $U = M \backslash C$ and let s be the restriction of t to U. Then s is a submersion as in Case 2. Also, given a probability measure with density on M, the transformed measure by t is the transformed measure by s of the restriction to U.

CASE 4. t is a proper almost submersion. Let μ be a probability measure on M having a density which is C^r a.e. Let Z be a closed null set with $\mu(Z) = 0$ such that the restriction of μ to $M \backslash Z$ has a C^r density. In the notation of Case 3, $t^{-1}(t(C \cup Z)) = C \cup s^{-1}(t(C \cup Z))$. As s is a submersion, the inverse image under s of a null set is null (because this is true of projections). By hypothesis C and Z are null, so $t^{-1}(t(C \cup Z))$ is null. Put $V = M \backslash t^{-1}(t(C \cup Z))$. As t is proper and N is locally compact, V is open and the restriction of t to V is a proper submersion. Now we can proceed as in Case 2. However, as the submersion is proper, every point of $t(V)$ has a compact neighbourhood K with $t^{-1}(K)$ compact. Thus in the restriction of $\sum_\alpha t_*(\phi_\alpha \mu)$ to K only finitely many terms are nonzero. It follows that the density of the transformed probability measure is C^r on $t(V)$ and so C^r a.e. on N.

REMARKS. (i) It is usually reasonably easy to test whether or not a given function is an almost submersion.

(ii) The proof of the theorem provides (at least in principle) a method of calculating the density of the transformed measure.

(iii) The proof also provides restrictions on the possible singularities and discontinuities of the density of the transformed measure. They must be in $t(C)$.

(iv) With a little more work one can prove the result also for t proper and analytic and for densities analytic a.e.

EXAMPLE. Sample means for distributions on the circle $S^1 = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}$. Take $M = S^1 \times S^1 \times S^1$, $N = \mathbb{R}^2$, $t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{3}(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3)$. The derivative of t at $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) \to \frac{1}{3}(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + \boldsymbol{\xi}_3)$ where $\boldsymbol{\xi}_i \mathbf{x}_i' = 0$, $i = 1, 2, 3$. This has rank less than 2 precisely when the tangent lines to S^1 at $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are parallel. $C = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) : \mathbf{x}_1 = \pm\mathbf{x}_2 = \pm\mathbf{x}_3\}$. Thus if $\bar{R} = \|t(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\|$ and we consider a probability distribution with C^1 density on the circle, the density of the sample mean has its singularities, if any, where $\bar{R} = 1$ or $\bar{R} = \frac{1}{3}$ (cf. Mardia (1972), page 95).

As usual, if t is analytic we get an "all or nothing" result.

PROPOSITION. *If $t: M \to N$ is analytic and M is connected, either t is an almost submersion (and so is density preserving) or every point of M is a critical point (and no probability distribution with a density on M is transformed by t to a probability distribution with a density on N).*

PROOF. C is the set of points of M at which the derivative of t has rank less than $n = \dim N$. Thus C is the set of common zeros of the determinants of the $n \times n$ minors of the derivative of t. These determinants are analytic functions as t is analytic. Therefore either C is null or $C = M$. If $C = M$, $t(M)$ is null by Sard's theorem, so no probability distribution on M is transformed by t to a probability distribution with a density on N.

This proposition will greatly simplify the calculations in Section 3.

3. Sample means. Let M be a submanifold of $\mathbb{R}^n$ not contained in any proper affine subspace, and consider a probability distribution with a density on M. It is very tempting to assume that, for large enough samples, the sample mean has a density. This is too optimistic. Let M be a smooth 1-dimensional submanifold of $\mathbb{R}^2$ containing a straight line segment—rather like a letter D—and let the density be positive on this line segment. Then for all sample sizes, the sample mean is on the line segment with positive probability and so cannot have a density.

One would hope to forbid such nasty behavior by restricting attention to analytic submanifolds. Certainly all is well for most Stiefel manifolds. The Stiefel manifold $V_n(\mathbb{R}^p)$ $[= O(n, p)]$ is the set of orthonormal n-frames in $\mathbb{R}^p$. Considering this as the set of $n \times p$ matrices $\mathbf{X}$ satisfying $\mathbf{X}\mathbf{X}' = \mathbf{I}_n$ gives an analytic embedding of $V_n(\mathbb{R}^p)$ as a compact submanifold of $\mathbb{R}(n, p)$, the vector space of $n \times p$ matrices with real entries.

THEOREM 2. *The sample mean function* $\alpha : V_n(\mathbb{R}^p)^r \to \mathbb{R}(n, p)$ *defined by*

$$\alpha(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_r) = \frac{1}{r}(\mathbf{X}_1 + \cdots + \mathbf{X}_r)$$

is a.e. C^s *density-preserving if and only if*

$$r \geqq 2 \quad \text{for} \quad n < p\,,$$
$$r \geqq 3 \quad \text{for} \quad n = p \geqq 3\,.$$

If $n = p = 2$, α *is not density-preserving.*

PROOF. An immediate consequence of the definition of critical point is that a necessary condition for α to be an almost submersion is that $\dim V_n(\mathbb{R}^p)^r \geqq \dim \mathbb{R}(n, p)$. As $\dim V_n(\mathbb{R}^p) = np - n(n+1)/2$ and $\dim \mathbb{R}(n, p) = np$, a necessary condition for α to be a.e. C^s density-preserving is $r \geqq 2$ for $n < p$, $r \geqq 3$ for $n = p \geqq 3$.

To prove the condition sufficient we shall use the function $\beta : SO(p) \times SO(p) \to \mathbb{R}(p-1, p)$ defined by $\beta(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{2}\pi(\mathbf{X}_1 + \mathbf{X}_2)$, where $\pi : \mathbb{R}(p, p) \to \mathbb{R}(p-1, p)$ deletes the last row of a matrix. We shall show that β is an almost submersion. The tangent space of $O(p)$ at $\mathbf{X}_i$ is the set of $\mathbf{Y}_i$ $[= d\mathbf{X}_i]$ satisfying $\mathbf{X}_i\mathbf{Y}_i' + \mathbf{Y}_i\mathbf{X}_i' = \mathbf{0}_n$, and the derivative of β sends $(\mathbf{Y}_1, \mathbf{Y}_2)$ to $\frac{1}{2}\pi(\mathbf{Y}_1 + \mathbf{Y}_2)$. Let $\mathbf{G} \in O(p)$ be the permutation matrix with entries $g_{ij} = 1$ if $i = j + 1 \bmod p$, $g_{ij} = 0$ otherwise. Put $\mathbf{J} = \operatorname{diag}(1, \cdots, 1, -1)$ and define $\mathbf{\Gamma} \in SO(p)$ by $\mathbf{\Gamma} = \mathbf{G}$ for p odd, $\mathbf{\Gamma} = \mathbf{J}\mathbf{G}$ for p even. The derivative of β at $(\mathbf{I}_p, \mathbf{\Gamma})$ sends $(\mathbf{A}, \mathbf{B}\mathbf{\Gamma})$ to $\frac{1}{2}\pi(\mathbf{A}, \mathbf{B}\mathbf{\Gamma})$

where $\mathbf{A}$ and $\mathbf{B}$ are antisymmetric $p \times p$ matrices. Consider the associated linear function γ from antisymmetric $p \times p$ matrices to symmetric $(p-1) \times (p-1)$ matrices defined by $\gamma(\mathbf{B}) = \rho(\mathbf{B\Gamma} + (\mathbf{B\Gamma})')$, where ρ deletes the last row and column. If $\mathbf{B}$ has entries b_{ij}, $\gamma(\mathbf{B})$ has entries $b_{i,j+1} - b_{i+1,j}$. Take bases $\{\mathbf{E}_{ij} - \mathbf{E}_{ji}: 1 \leqq i < j \leqq p\}$ of the antisymmetric $p \times p$ matrices and $\{\mathbf{E}_{ij} + \mathbf{E}_{ji}: 1 \leqq i \leqq j \leqq p-1\}$ of the symmetric $(p-1) \times (p-1)$ matrices, where $\mathbf{E}_{ij}$ is the matrix with (i, j)th entry 1 and other entries 0. If these bases are ordered lexicographically γ is represented by a triangular matrix with nonzero diagonal entries. Thus γ is invertible. It follows that the derivative of β at $(\mathbf{I}_p, \mathbf{\Gamma})$ is onto, so $(\mathbf{I}_p, \mathbf{\Gamma})$ is not a critical point of β. As β is analytic and $SO(p)$ is connected, β is an almost submersion by the proposition.

Now for $n < p$ there are functions $\pi_n: SO(p) \to V_n(\mathbb{R}^p)$, $\bar{\pi}_n: \mathbb{R}(p-1, p) \to \mathbb{R}(n, p)$ which delete all but the first n rows of a matrix. π_n is onto and $\bar{\pi}_n$ is a submersion. Also $\alpha(\pi_n(\mathbf{X}_1), \pi_n(\mathbf{X}_2)) = \bar{\pi}_n(\beta(\mathbf{X}_1, \mathbf{X}_2))$. It follows that α is an almost submersion for $r = 2$ and hence for $r \geqq 2$. As $V_n(\mathbb{R}^p)$ is compact, Theorem 1 shows that α is a.e. C^s density preserving.

If $n = p \geqq 3$, put

$$\mathbf{H} = \begin{pmatrix} \mathbf{I}_{p-3} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}$$

with

$$\mathbf{R} = \begin{pmatrix} \kappa & \kappa & 0 \\ 0 & 0 & 1 \\ \kappa & -\kappa & 0 \end{pmatrix}$$

where $2\kappa^2 = 1$. Then $\mathbf{H} \in SO(p)$.

We claim that $(\mathbf{I}_p, \mathbf{G}, \mathbf{H})$ is not a critical point of α for $p \geqq 4$. To see this, first note that if $\mathbf{X} = (x_{ij})$ is of the form $\mathbf{A} + \mathbf{BG}$ with $\mathbf{A}$, $\mathbf{B}$ antisymmetric $p \times p$ matrices, then for $1 \leqq r \leqq p$, $\phi_r(\mathbf{X}) = \sum_{i+j=r} x_{ij} = 0$ where all subscripts are taken mod p. As γ (in the calculation for $n = p - 1$) is invertible, $\mathbf{X}$ is of the form $\mathbf{A} + \mathbf{BG}$ if and only if $\phi_r(\mathbf{X}) = 0$, $1 \leqq r \leqq p$. A brief calculation now shows that the function taking the antisymmetric matrix $\mathbf{C}$ to $(\phi_1(\mathbf{CH}), \cdots, \phi_p(\mathbf{CH}))$ maps onto $\mathbb{R}^p$. Thus the function $(\mathbf{A}, \mathbf{B}, \mathbf{C}) \to \mathbf{A} + \mathbf{BG} + \mathbf{CH}$ is onto, and so $(\mathbf{I}_p, \mathbf{G}, \mathbf{H})$ is not a critical point of α. If $p = 3$, $(\mathbf{I}_3, \mathbf{G}, \mathbf{H}^2)$ is not a critical point of α. Similarly $(\mathbf{I}_p, \mathbf{\Gamma}, \mathbf{H})$ is not a critical point of α. (Use $\psi_r(\mathbf{X}) = \sum_{i+j=r} x_{ij} - \sum_{i+j=p+r} x_{ij}$.) Thus $(\pm\mathbf{I}_p, \pm\mathbf{G}, \pm\mathbf{H})$ for p odd $(p \geqq 5)$, $(\pm\mathbf{I}_3, \pm\mathbf{G}, \pm\mathbf{H}^2)$ for $p = 3$, $(\mathbf{I}, \mathbf{G}, \mathbf{H})$, $(\mathbf{I}, \mathbf{\Gamma}, \mathbf{H})$, $(\mathbf{J}, \mathbf{\Gamma}, \mathbf{HJ})$ etc. for p even are noncritical points of α in each component of $O(p) \times O(p) \times O(p)$. Also α is analytic, so by the proposition α is an almost submersion for $r = 3$ and so for $r \geqq 3$. Thus α is a.e. C^s density-preserving.

Finally, recall that $SO(p) \subset O(p) = V_p(\mathbb{R}^p)$ and that normalised Haar measure on $V_p(\mathbb{R}^p)$ assigns to $SO(p)$ measure $\frac{1}{2}$. Now dim $\mathbb{R}(2, 2) = 4$, but $SO(2)$ is contained in the two-dimensional subspace of matrices of the form $\begin{pmatrix} c & s \\ -s & c \end{pmatrix}$. Thus if $n = p = 2$, for any r the image under α of normalised Haar measure gives positive measure to a proper subspace, so α is not density-preserving.

The above results lead to the conjecture that for a connected compact analytic submanifold M of $\mathbb{R}^p$ which is not contained in a proper affine subspace the sample mean function for sample size r is a.e. C^s density-preserving if $r \cdot \dim M \geqq p$. However, Theorem 3 below shows that for projective spaces this is too optimistic. The condition in the conjecture reduces to $r \geqq (p + 2)/2$. This is not a sufficient condition. The true necessary and sufficient condition is $r \geqq p$.

Consider the Grassmann manifold $G_n(\mathbb{R}^p)$ of n-dimensional subspaces of $\mathbb{R}^p$. Identifying each subspace with the orthogonal projection onto it we get an analytic embedding of $G_n(\mathbb{R}^p)$ onto $\mathbb{R}(p, p)$ with image the set of matrices $\mathbf{X}$ satisfying $\mathbf{X} = \mathbf{X}' = \mathbf{X}^2$ and $\operatorname{tr} \mathbf{X} = n$. Thus if $S_n(p)$ denotes the affine space of symmetric $p \times p$ matrices with trace n, $G_n(\mathbb{R}^p)$ is a compact connected analytic submanifold of $S_n(p)$.

THEOREM 3. *Define the sample mean function* $\alpha\colon G_n(\mathbb{R}^p)^r \to S_n(p)$ *by*

$$\alpha(\mathbf{X}_1, \cdots, \mathbf{X}_r) = \frac{1}{r}(\mathbf{X}_1 + \cdots + \mathbf{X}_r)\,.$$

Then for $r \geqq p$, $1 < n < p$, α *is* a.e. C^s *density-preserving. If* $n = 1$ (*so* $G_n(\mathbb{R}^p)$ *is real projective p-space*), α *is density-preserving if and only if* $r \geqq p$.

PROOF. Consider the function $\pi\colon V_n(\mathbb{R}^p) \to G_n(\mathbb{R}^p)$ defined by $\pi(\mathbf{X}) = \mathbf{X}'\mathbf{X}$. (The geometrical interpretation of π is that it maps each frame to the orthogonal projection onto the subspace it spans.) The tangent space to $V_n(\mathbb{R}^p)^r$ at $(\mathbf{X}_1, \cdots, \mathbf{X}_r)$ is the set of $(\mathbf{Y}_1, \cdots, \mathbf{Y}_r)$ satisfying $\mathbf{X}_i\mathbf{Y}_i' + \mathbf{Y}_i\mathbf{X}_i' = \mathbf{0}_n$, $1 \leqq i \leqq r$; the tangent space to $S_n(p)$ at any point is $S_0(p)$; and the derivative of $\alpha_0\pi^r$ at $(\mathbf{X}_1, \cdots, \mathbf{X}_r)$ sends $(\mathbf{Y}_1, \cdots, \mathbf{Y}_r)$ to $1/r \sum_{i=1}^r (\mathbf{Y}_i'\mathbf{X}_i + \mathbf{X}_i'\mathbf{Y}_i)$. Let $\{\mathbf{e}_1, \cdots, \mathbf{e}_p\}$ be an orthonormal base of $\mathbb{R}^p$. Then $\{\mathbf{e}_j'\mathbf{e}_j - \mathbf{e}_1'\mathbf{e}_1, j \geqq 2; \mathbf{e}_i'\mathbf{e}_j + \mathbf{e}_j'\mathbf{e}_i, i < j\}$ is a base of $S_0(p)$.

Define $\mathbf{X}_1, \cdots, \mathbf{X}_p \in V_n(\mathbb{R}^p)$ by

$$\mathbf{X}_i' = (\mathbf{e}_1', \cdots, \mathbf{e}_n')\,, \qquad \mathbf{X}_j' = ((\mathbf{e}_{j-1}' + \mathbf{e}_j')/2^{\frac{1}{2}}, \mathbf{e}_{j+1}', \cdots, \mathbf{e}_{j+n-1}') \quad j \geqq 2\,.$$

Then a brief calculation shows that $(\mathbf{X}_1, \cdots, \mathbf{X}_p)$ is not a critical point of $\alpha_0\pi^p$, so $(\pi(\mathbf{X}_1), \cdots, \pi(\mathbf{X}_p))$ is not a critical point of α. By the proposition α is an almost submersion for $r = p$ and so for $r \geqq p$. Thus by Theorem 1 α is a.e. C^s density-preserving.

If $n = 1$, $V_1(\mathbb{R}^p) = S^{p-1}$. Given $(\mathbf{X}_1, \cdots, \mathbf{X}_r) \in (S^{p-1})^r$, let $\{\mathbf{f}_1, \cdots, \mathbf{f}_s\}$ be an orthonormal base of the space spanned by $\mathbf{X}_1, \cdots, \mathbf{X}_r$ and extend it to an orthonormal base $\{\mathbf{f}_1, \cdots, \mathbf{f}_p\}$ of $\mathbb{R}^p$. The image under the derivative of $\alpha_0\pi^r$ of the tangent space at $(\mathbf{X}_1, \cdots, \mathbf{X}_r)$ is contained in the span of the $\mathbf{f}_i'\mathbf{f}_j + \mathbf{f}_j'\mathbf{f}_i$ with $i \leqq s$ or $j \leqq s$. As $s < p$, $\mathbf{f}_p'\mathbf{f}_p - \mathbf{f}_1'\mathbf{f}_1$ is not in this image. Thus $(\mathbf{X}_1, \cdots, \mathbf{X}_r)$ is a critical point of $\alpha_0\pi^r$, so $\alpha_0\pi^r$ is not an almost submersion. As π is a submersion it follows that α is not an almost submersion, so is not density-preserving.

REFERENCES

[1] BOTT, R., GITLER, S. and JAMES, I. M. (1972). *Lectures on Algebraic and Differential Topology.* Springer Lecture Notes **279**. Springer-Verlag, Berlin.

[2] Brickell, F. and Clark, R. S. (1970). *Differentiable Manifolds.* Van Nostrand-Reinhold, London.
[3] Khatri, C. G. and Mardia, K. V. (1977). The von Mises-Fisher matrix distribution in orientation statistics. *J. Roy. Statist. Soc. Ser.B* **39** 95–106.
[4] Mardia, K. V. (1972). *Statistics of Directional Data.* Academic Press, London.
[5] Sternberg, S. (1964). *Lectures on Differential Geometry.* Prentice-Hall, Englewood Cliffs, N.J.

Department of Statistics
School of Mathematics
University of Leeds
Leeds LS2 9JT, England