generalization is an important one, although the terminology tends to separate the subject from an area that is well established in the statistical literature: the theory of maximal invariance. Thus a shape in this generalized sense is also a maximal invariant under the action of the group $\mathscr{G}$. Such maximal invariants for data on riemannian manifolds are not uncommon in statistics. For example, data in directional statistics live on a one- or two-dimensional sphere for which the most useful group to generate transformation models is the rotation group. For models in which the rotation group generates a nuisance parameter, questions involving the testing of a concentration parameter in the absence of knowledge of the nuisance parameter require the reduction to the maximal rotation invariant. Fraser (1968) has emphasized the importance of transformation models and the fibers of data sets equivalent under the action of a group. Some relationships with the statistics of shape are developed in Small (1983).

I would like to close these comments with some remarks of a more specialized technical nature. The elegance of D. G. Kendall's theory of shape is especially clear for data sets in dimensions 1 and 2. In higher dimensions, singularities start to emerge. Although these singularities are not obviously detrimental to a theory of shape, they do detract from the elegance of the representation. Even in dimensions 1 and 2, the shape spaces are more easily constructed and represented than the corresponding size and shape spaces. The reason for the elegance of the shape space for the cases where $m = 1, 2$ is that in these cases a shape preserving transformation can be uniquely decomposed into two transformations that correspond to multiplication and addition in the real line and complex plane for the respective dimensions. For both $m = 1$ and $m = 2$ the group of shape preserving transformations is a solvable group with a dimension (i.e., number of degrees of freedom), which is an integral multiple of $m$. But in dimension $m = 3$ this fails to be the case. The group of shape preserving transformations is of dimension 7, which is not a multiple of $m$.

Let me conclude my remarks by congratulating David Kendall on some very interesting work. The new directions that are sketched in this paper seem to be promising for the analysis of geometric data of various kinds and from various sources. I hope that much more is forthcoming.

## ADDITIONAL REFERENCES


AMARI, S. (1985). *Differential-Geometric Methods in Statistics*. Springer, New York.

BROWN, B. M. (1983). Statistical uses of the spatial median. *J. Roy. Statist. Soc. Ser. B* **45** 25–30.

BUJA, A. and ASIMOV, D. (1985). Grand tour methods: An outline. *Computer Science and Statistics: Proceedings of the 17th Symposium on the Interface* (D. M. Allen, ed.) 63–67. North-Holland, Amsterdam.

FRASER, D. A. S. (1968). *The Structure of Inference*. Wiley, New York.

HURLEY, C. (1987). The data viewer: A program for graphical data analysis. Ph.D. dissertation and Technical Report, Dept. Statistics, Univ. Washington.

SMALL, C. G. (1983). Characterization of distributions from maximal invariant statistics. *Z. Wahrsch. verw. Gebiete* **63** 517–527.

SMALL, C. G. (1984). A classification theorem for planar distributions based on the shape statistics of independent tetrads. *Math. Proc. Cambridge Philos. Soc.* **96** 543–547.

SMALL, C. G. (1985). Decomposition of models whose marginal distributions are mixtures. *Canad. J. Statist.* **13** 131–136.


# Comment: Some Contributions to Shape Analysis

**Kanti V. Mardia**

There are no words to express the profound depth of Kendall's work. I have been working in this area intermittently since 1976 and I believe his fundamental work (as well as Bookstein, 1986) has opened up the field.

*Kanti V. Mardia is Professor and Head, Department of Statistics, University of Leeds, Leeds LS2 9JT, England.*

Bookstein (1986) has used the model for shape analysis assuming that the points are distributed independently as $N_2(\mu_i, \sigma^2\mathbf{I})$, $i = 1, 2, \cdots, p$. Consider $p = 3$. Let $\mathbf{x}$ be the point in Kendall's spherical shape space from these three points with $\ell$ representing the corresponding point in Kendall's space from $\mu_i$'s. Let $\bar{\mu}$ be their mean vector. Then using Mardia and Dryden (1989), it can be shown that the probability element of $\mathbf{x}$ is given by

$$\{1 + \varkappa(\ell'\mathbf{x} + 1)\}e^{\varkappa(\ell'\mathbf{x}-1)}dS, \quad \mathbf{x} \in S_2,$$

where $\varkappa \geq 0$, $\ell \in S_2$ and $dS$ is the uniform measure on $S_2$. In fact, $\varkappa = \sum_{i=1}^{3} |\mu_i - \bar{\mu}|^2/(4\sigma^2)$ where $|\cdot|$ denotes the Euclidean norm. This distribution on Kendall's space will be written as $K(\ell, \varkappa)$. It has a mode at $\ell$ and for $\varkappa = 0$, it is uniform on $S_2$. Further, as $\varkappa \to \infty$, we have bivariate normality, and as $\varkappa \to 0$, we have uniformity. The distribution $K(\ell, \varkappa)$ is, of course, not the Fisher distribution $F(\ell^*, \varkappa^*)$, but it belongs to the class of rotationally symmetric distributions as one would have expected. On equating the first order moments, this distribution (see Figure 1), is found to be very similar to the Fisher distribution with $\ell^* = \ell$. As we would mostly expect large $\varkappa$ for biological shapes, it seems we can carry out inference for the triangle case using the Fisher distribution. Note that because $K(\ell, \varkappa)$ is not in the exponential family, inference is somewhat more complicated for this distribution than for the Fisher distribution. One possible advantage of this approach for $p = 3$ may be as follows. It can be shown that for the variables in Bookstein's space, the second moments are infinite (Mardia and Dryden, 1989), although here all the "moments" are finite. The statistical implication of this point needs closer examination but these two approaches, directional and multivariate, should prove complementary.

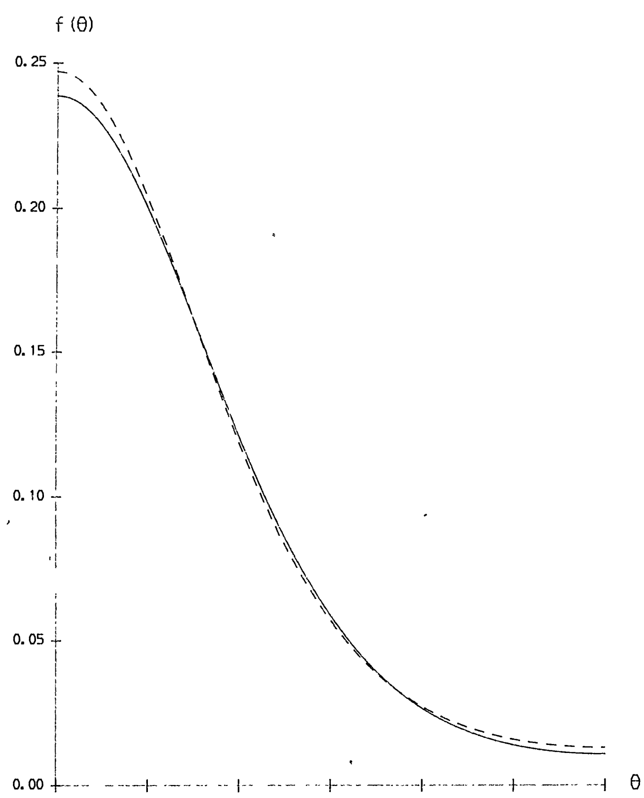



FIG. 1. *The profiles $f(\vartheta)$ of the two spherical distributions with $\ell'\mathbf{x} = \cos\vartheta$, $0° < \vartheta < 180°$. ——, exact distribution ($\varkappa = 1$); – – –, Fisher distribution ($\varkappa^* = 1.47$).*

Let $\alpha_1$, $\alpha_2$ and $\alpha_3$ be the three angles of the triangle and suppose the "handedness" of the triangle is ignored. In passing, we note that Mardia (1980) obtained the p.d.f. of two angles $\alpha_1$ and $\alpha_2$ for $\sigma \to \infty$ (or $\mu_i$'s equal),

$$6S/\{\pi(3 - C)^2\}$$

with $S = \Sigma \sin 2\alpha_i$, $C = \Sigma \cos 2\alpha_i$ saying "it is uniform in a certain sense" but could not see its implication. Of course, it now dawns that this is uniform in Kendall's spherical space.

It has been assumed that the landmarks are known but in practice they may not be. The determination of biological landmarks relies on expert opinion in creating homologous points. However, we emphasize mathematical landmarks that are some extremal points on the outlines (e.g., points of maximum curvature). These points can also be obtained by fitting poly-lines recursively. For example, for the palm shape in Figure 2, using any standard algorithm, we get the landmarks $P_1, \cdots, P_8$ (not necessarily in that order) with base $P_1P_2$. How many landmarks one should take for a given shape depends on the problem. For discrimination, we can carry out tests of dimensionality on shape variables recursively to select the number of landmarks (see Mardia, 1986). However, there are many problems in making sure that there is correspondence of the landmarks within and between the groups. If the fingers are not abducted, only the point $P_8$ will be very unstable and it may be necessary to select the threshold to be large in the poly-lines fitting algorithm. In such cases, a suitable model for the points must be independent $N_2(\mu_i, \sigma_i^2\mathbf{I})$ where the $\sigma_i$'s are not necessarily equal (see Mardia and Dryden,

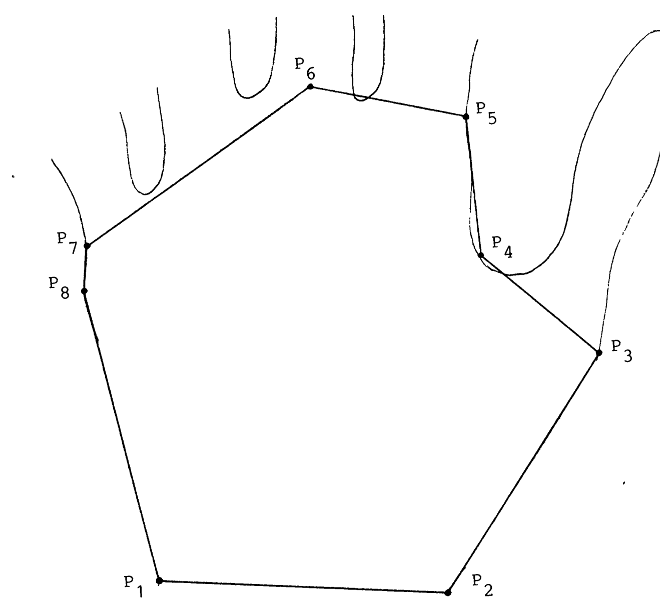



FIG. 2. *Landmarks estimation for a palm shape.*

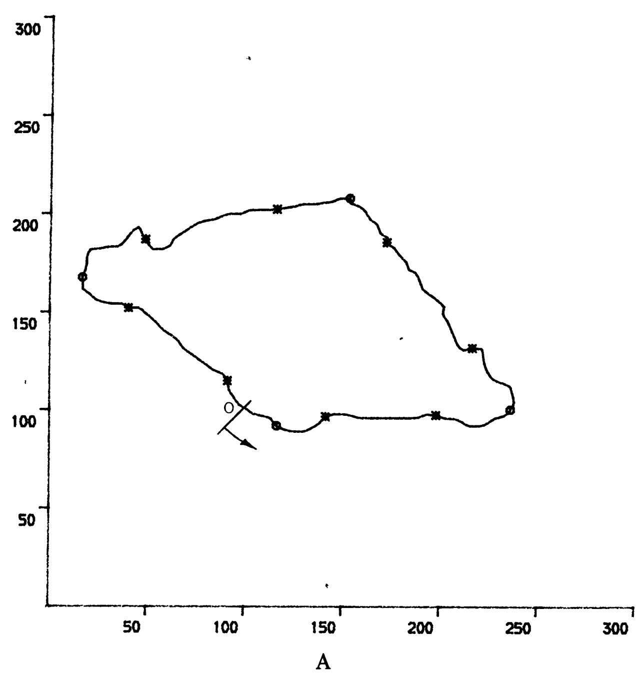
1989). The problem of shape change seems to be more difficult.

Nowadays, one can work with video-digitized images and the outline in Figure 3A gives a digital reconstruction of a $T1$ mouse vertebra bone (see Johnson, O'Higgins, McAndrew, Adams and Flinn, 1985). An algorithm to estimate landmarks, developed jointly with Ian Dryden, works as follows. The digitized outline is represented by a set of points (here over 300 points) and the local absolute curvature maxima are found (about 30 points here). Then a poly-line fitting algorithm is applied to this subset of points and the vertices of the resulting polygon are taken as landmarks (which are the points ∘ of maximum curvature in Figure 3B). Some further landmarks are located at either side of each landmark, at points of low absolute curvature (* in Figure 3B). We will call them pseudo-landmarks and they provide useful local shape information. Thus for the $T1$ vertebrae we obtain four landmarks and eight pseudo-landmarks as finally mapped onto the outline in Figure 3A. Consistency of the landmarks is verified by applying the algorithm to all the bones, and for the $T1$'s we have a reasonable correspondence both within and between the three groups. The algorithm is not straightforward for more complicated shapes like the $T2$ bone in Johnson, O'Higgins, McAndrew, Adams and Flinn (1985). It is expected that significant developments will take place in the estimation of landmarks, especially for shape change. Again, the work developed in the paper will play an important role.

Coming to the Central Place Theory application, Mardia (1977) also mentioned the use of the distribution of the circumradius $R$ in addition to the shape variables when the size is relevant but now I am convinced by the practical reasons put forward here. I agree, the method of eliminating the edge effect is effective. Indeed, it was used in simulations for Mardia (1977). "One method of eliminating boundary effects is to neglect those triangles whose circumcircles are not wholly within the sampling window." (See Edwards, 1980, pages 107 and 108.) In particular, the table below shows for $n = 44$ (mimicking the Iowa data) for 1000 simulations, the effect of boundary on the area $A$ of the shape of Delaunay triangles in a rectangle with sides in the ratio 1:2.

| | Simulated | Corrected | Miles' Dist. |
|---|---|---|---|
| $E(A)$ | 1.52 | 1.57 | 1.57 |
| $\text{var}(A)$ | 0.46 | 0.45 | 0.45 |

As expected, the correction process throws out the long thin triangles with low values of $A$ but in this case there is little effect. Even if we assume the independence of Delaunay's triangles, it has been pointed out that we will need a powerful test for the hypothesis of "equilateralness" under Delaunay's tesselations versus that under Central Place Theory. One new plausible approach is as follows. It can be shown that if we approximate Miles' density with $K(\ell, \varkappa)$ by equating the mode and the strength of the mode for the two distributions on the half-lune then $\ell' = (0, 0, 1)$, $\varkappa = 1.73 = \varkappa_0$, say. We can now test this null



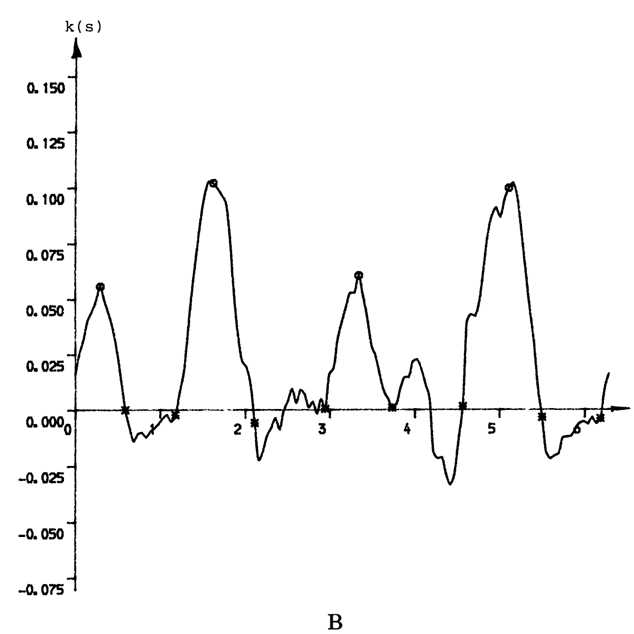




FIG. 3. *A, reconstructed $T1$ bone from video-digital image with estimated landmarks (∘), psuedo-landmarks (*). O indicates the start point together with its direction for B. B, curvature $k(s)$ versus arc length s for A, and the corresponding positions of estimated landmarks.*

hypothesis $\varkappa = \varkappa_0$ against $\varkappa > \varkappa_0$ because under Central Place Theory, $\varkappa$ will be very large under $H_1$. Note that under both hypotheses, we have $\ell' = (0, 0, 1)$.

Under the Fisher approximation to $K(\ell, \varkappa)$, we could use under $H_0$

$$2\gamma_0(n - \Sigma z_i) \sim \chi^2_{2n}, \gamma_0^{-1} = \varkappa_0^{-1} - \frac{1}{5}\varkappa_0^{-3},$$

where $(x_i, y_i, z_i)$, $i = 1, \cdots, n$, are the $n$ spherical coordinates for Delaunay's triangles specified on the half-lune as in Kendall (1983). The critical region is the lower tail of the distribution. Note that in terms of Bookstein's shape variables for the triangles $(Q_{1i}, Q_{2i})$, $i = 1, \cdots, n$, we have

$$z_i = \sqrt{3}Q_{2i}/(Q_{1i}^2 + Q_{2i}^2 - Q_{1i} + 1).$$

There is considerable room to improve the test. For example, we could estimate the percentage points of the test by simulating the Poisson process. Also, we could carry out a test for the non-nested hypothesis of the Miles' distribution versus $K(\ell, \varkappa)$ without any approximation. All these ideas require further investigation. Another approach when the size of the triangles is important is to use the mean area of triangles like Mardia (1977) but now without normalizing to $R = 1$. Its mean and variance are known under the Miles' distribution and thus we can test the null hypothesis. Of course, testing $H_0$ is only a small part of the main problem; the shape and size summary statistics themselves are revealing, e.g., in investigating comparative evidence of Central Place Theory for various different data. It would be interesting to see a detailed analysis of the Wisconsin data along the lines given in the paper.

Finally, let me say that I found the paper very stimulating and look forward to reading the forthcoming book.

## ADDITIONAL REFERENCES

EDWARDS, R. (1980). Some contributions to directional statistics. Ph.D. dissertation, Univ. Leeds.

JOHNSON, D. R., O'HIGGINS, P., MCANDREW, T. S., ADAMS, L.M. and FLINN, R. M. (1985). Measurements of biological shape: A general method applied to mouse vertebrae. *J. Embryol. Exp. Morphol.* **90** 363–377.

MARDIA, K. V. (1977). Discussion. *Bull. Internat. Statist. Inst.* **47** (2) 145–146.

MARDIA, K. V. (1980). Discussion of "Simulating the ley hunter," by S. R. Broadbent. *J. Roy. Statist. Soc. Ser. A* **143** 137.

MARDIA, K. V. (1986). Optimum design-based shape discriminators of an object and an analysis of palm shapes. Research Report, Dept. Statistics, Univ. Leeds, (presented to Princeton Symp. on Shape Analysis, May 1987).

MARDIA, K. V. and DRYDEN, I. L. (1989). Shape distributions for landmark data. *Adv. Probab.* To appear.

# Comment

**Wilfrid S. Kendall**

David Kendall has been my close collaborator from the very start of my scientific career, and so it gives me great pleasure to add to the discussion of this paper. I take as my theme the application of computer algebra in statistics and probability. As evidenced from the paper, some of the first instances of this have occurred in the statistical theory of shape. I shall make some remarks on the general application of computer algebra in statistical science, and then turn to the specific application (to the diffusion of shape) with which I have been involved recently.

### 1. COMPUTER ALGEBRA IN STATISTICS AND PROBABILITY

The reader will have noticed several references to the use of computer algebra (CA) in the investigations reported in the paper. To my knowledge this usage represents one of the first substantial applications of CA in the fields of statistics and probability. The others known to me are my own related work on shape diffusions (referred to in the paper as W. S. Kendall, 1988), which was encouraged by the success of CA in investigating the geometry of shape and is discussed further below; and the work on asymptotics in density estimation as described by Silverman and Young (1987). (I would be most grateful to hear of further instances.)

At present the use of CA in statistical science is in its infancy, although many exciting possibilities beckon. The emergence of readily available and powerful personal workstations gives reason to hope for rapid progress in the next few years. The wide screens, multiple tasking facilities and cut - and - paste editing of these workstations combine to yield a most productive environment for CA.

In what sort of areas might we anticipate CA's profitable employment? At the time of writing it seems

*Wilfrid S. Kendall is a Lecturer in Statistics, Department of Statistics, University of Warwick, Coventry, CV4 7AL, England. He is also David Kendall's son.*