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

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

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.

to me that maximum benefits will derive in areas that not only require complicated calculations but also exhibit considerable structure. The theory of statistics of shape is a good example: the calculations comprising Le Huiling's magnificent determination of shape densities are undergirded by a profound appreciation of the tessellation structure hidden within the density, and I believe it is this structure that facilitates the implementation of an effective algorithm in CA. In this case manual calculations preceded the implementation, and so CA played a supportive role. But one can envisage a future in which the CA implementation of the structure develops at the same pace as one's theoretical understanding of the structure (and this was indeed the case in the work on diffusion of shape). Other possible applications are to the theory of asymptotic distributional approximations, and to the related field of differential geometry of statistical inference.

Where the underlying structure is not properly appreciated, naive use of CA can very easily lead to a kind of algebraic overflow. In the words of Hearn (1985) "... we attempt to solve more and more complicated problems and succeed only in producing larger and larger expressions." Hearn's point here (as author of a CA system) is the need for structure-detecting algorithms in CA systems; I believe that as users we should draw the moral that increasing usage of CA will force us to acquire ever deeper theoretical understanding in our search for useful and practicable ways to express the underlying structure. (Consider the mathematical demands made on the reader of Davenport (1981), which expounds a CA algorithm for solving indefinite integrals.)

It is of course a legitimate concern whether one can accept as valid an argument which depends on a CA package, whose correctness has not been mathematically established. Although a proper treatment requires more philosophical expertise on the nature of a valid proof than I can muster, the following remarks may be helpful. Firstly, we learn from the history of famous conjectures that (even when CA is not employed) it is a nontrivial matter to establish whether an argument is valid. At the very least, potential bugs in a CA package are rather more public than possible bugs in my head! Secondly, and related to this, convincing mathematical arguments derive from good exposition, which builds up a coherent and checkable mathematical world in which the target result appears as a natural result. In a similar way a good application of CA should clearly implement a mathematical structure, susceptible to interactive checking by sceptics. Thus the final CA program should serve as a kind of "active text"; if an opponent can legitimately manipulate it to derive (for example) a negative variance then the argument fails. These points provide further motivation to use CA to implement mathematical structure rather than merely to crunch large formulas.

As remarked above, the next few years should see rapid development in the application of CA to statistical science. The work surveyed in the paper is pioneering in this as well as in other respects.

2. DIFFUSION OF SHAPE

The investigations reported in W. S. Kendall (1988) were motivated by a desire to develop CA in application to statistics and probability. About two years ago I realized that the basic idea of stochastic calculus ("replace the square dB^2 of the Brownian path increment by the time increment dt") lent itself naturally to implementation as a substitution role in a CA language such as REDUCE. To my surprise I found this implementation could actually be used to provide a stochastic calculus proof of the remarkable Clifford-Green result reported in Clifford, Green and Pilling (1987), and this led naturally to determination of the statistics of shape diffusion for triads of points.

Here is a summary of the argument concerning diffusion of shapes of triads.

(a) The shape of a triangle formed by $k = 3$ points in m-space (for $m \geq 3$) is parametrized by the ratios of the squared side-lengths of the triangle (the so-called homogeneous shape coordinates).

(b) If the vertices of the triangle diffuse as Ornstein-Uhlenbeck processes then their equilibrium distribution is rotationally-symmetric Gaussian.

(c) Computer algebra allows the ready determination of the statistics of the induced random process of homogeneous shape coordinates, using the implementation of stochastic calculus described above.

(d) Further computer algebraic manipulation reveals: the natural geometry of the shape space (as suggested by the shape process) is that of ("northern") hemisphere of radius ½; and moreover a time change of the shape process is Brownian motion on this hemisphere modified by a drift directed toward the north pole of the hemisphere. The drift can of course be found explicitly!

(e) The equilibrium distribution of the shape diffusion (and thus the shape density for a rotationally symmetric Gaussian triangle) can then be found using diffusion theory arguments.

(f) The same picture arises (with minor technical modifications) if the vertices diffuse by Brownian motion.

In particular these results make it clear that in high dimensions a Brownian triangle spends most of the time in shapes that are close to equilateral. This is the surprising conclusion referred to in the paper—

surprise however is a relative concept and readers of McKean (1973) would not be surprised at all!

No doubt readers will see other ways of addressing these problems using perhaps stochastic calculus without benefit of CA or the theory of Wishart distributions (indeed Mr. James of Leeds University has shown me how to use Wishart matrix theory to establish the Clifford-Green result mentioned above). The main purpose of this work has been to initiate the development of CA as an effective tool in the study of random processes, rather than to develop new results. More recently, and with the same motivation, I have been working on the use of CA to derive the statistics of shape diffusions for k-ads with $k > 3$. Here the technical problem is to find effective ways of dealing with sums involving k summands, when k is not fixed beforehand but must be treated as a symbolic quantity. Some progress has been made, but work is not yet complete.

ADDITIONAL REFERENCES

DAVENPORT, J. H. (1981). *On the Integration of Algebraic Functions.* Springer, Berlin.

HEARN, A. C. (1985). Structure: The key to improved algebraic computation. In *Symbolic and Algebraic Computations by Computers* (N. Inada and T. Soma, eds.). World Scientific, Singapore.

MCKEAN, H. P. (1973). Geometry of differential space. *Ann. Probab.* **1** 197–206.

SILVERMAN, B. W. and YOUNG, G. A. (1987). The bootstrap: To smooth or not to smooth? *Biometrika* **74** 469–479.

Comment

Geoffrey S. Watson

In stochastic geometry as in number theory, it is easy to ask questions that the layman can understand but that the specialist can only answer with difficulty or not at all. Under the older name, geometrical probability, the subject is old, e.g., Buffon's famous problem was invented around the time Buffon was preparing a French version of Newton's "fluxions." I don't know of any ancient and unresolved conjectures like Fermat's but it is easy to give simple-sounding problems that are hard to solve, e.g., the motivating problem of Kendall's theory of shape. How do the shapes of triangles vary when their vertices are independently and uniformly distributed in a fixed rectangle? This problem arises from questions about whether there is too much "collinearity" in sets of points (see Figures 1 and 2). A recent and very readable survey of Kendall's theory has been given by Small (1988).

All but the most mathematically gifted readers will find this paper difficult. Rather more basic details are given in Kendall (1984), but this too is written for mathematicians. I hope the promised book (now in preparation) by Carne, Kendall and Le will make it clear to statisticians, because I'm sure that this is a fascinating area for research and applications. To support this belief I will give a brief summary of my own related efforts, sticking mainly to triangles. This is reasonable because most of the suggested applications use them and they are the simplest case.

The shape of Δ, a triangle P_1, P_2, P_3, with vertex angles α_1, α_2, α_3, could be defined as the pair (α_1, α_2). But for most problems this is not easy to work with, or to generalize to k labeled points in $\mathscr{R}^m$. There are lots of other ways to define the shape of a triangle. We may think of Δ as a 2×3 matrix $[z_1, z_2, z_3]$, where the column z_i has elements x_i, y_i, and denotes the position of the vertex P_i in the plane. Because we are only interested in the shape of Δ we may translate, dilate and rotate Δ without changing the shape of Δ, so we seek a "canonical" triangle. Kendall's approach is a variant of the following. Change the origin to the centroid of the triangle and consider the singular value decomposition of the new 2×3 matrix, $R\Lambda L'$, where R is a 2×2 rotation and so irrelevant. By scaling we could make $\lambda_1^2 + \lambda_2^2 = 1$. The remaining object defines the shape. See Mannion (1988) for a simple description—it is very similar to the next suggestion—and Small (1988).

I found Kendall's reduction hard to understand and considered (in Watson, 1986) two alternatives, which worked well in the simple planar problem I had posed. Move P_1 to the origin (0, 0), move P_2 to (0, 1), which uses up the available transformations, and denote P_3 by z, which then serves to define the shape of Δ. It is natural to take it as a point in the complex plane. The other alternative came from taking z_1, z_2, z_3 as

Geoffrey S. Watson is Professor, Fine Hall, Washington Road, Princeton University, Princeton, New Jersey 08544.