Vector Correlation

M. A. Stephens

Biometrika, Volume 66, Issue 1 (Apr., 1979), 41-48.

Biometrika (1979), **66**, 1, *pp.* 41–8
Printed in Great Britain

Vector correlation

BY M. A. STEPHENS

Mathematics Department, Simon Fraser University, Burnaby, British Columbia

SUMMARY

This paper discusses the measurement of correlation between two sets of vectors. The vectors may be thought of as denoting directions in p dimensions. Two main measures of correlation are proposed, based on the premise that the two sets would be perfectly correlated if an orthogonal transformation, or less generally a rotation transformation, makes the second set coincide with the first. Natural extensions exist to cover correlation without rotation, or serial correlation. For testing for correlation, distributional results are given for $p = 2$ and 3, and especially for uniform parent populations.

Some key words: Correlation between directions; Correlation between vectors; Correlation on the circle; Correlation on the sphere.

1. INTRODUCTION

In some statistical work, the measurements of interest can be recorded as unit vectors. These are drawn from the centre O of a hypersphere, of unit radius, to points on the circumference or surface. In two or three dimensions the vectors may denote directional data, or, if the circle is regarded as a clock, the endpoints can represent events. Suppose that the data consist of paired vectors $u_i = OP_i$ and $v_i = OQ_i$, for $i = 1, \ldots, n$; for example, u_i and v_i might be directions of magnetization of the ith rock sample before and after laboratory treatment. Then P_i and Q_i are points on the hypersphere of unit radius. It is of interest to develop a theory of correlation for such paired vectors.

In this paper we propose several vector-correlation statistics, depending on the nature of the relation expected. The main definition of correlation will be based on measuring how close the u vectors can be brought to the v vectors by an orthogonal transformation H. Suppose the two sets u_i and v_i, in p dimensions, are the rows of $n \times p$ matrices U and V; we shall use U and V to refer both to the matrices and to the sets of vectors. Let H' denote the transpose of H. If H can be found so that $H'u_i$ equals v_i for all i, we shall describe U and V as perfectly correlated. The definition will be tightened to insist that the orthogonal transformation be a rotation. Simple adaptations lead to measures of correlation between U and V *in situ*, with no transformation allowed, for example to measure the directional correlation between the prevailing wind at airports and the main runways, and of serial correlation in one vector set U. These uses relate to work of Epp, Tukey & Watson (1971) and of Watson & Beran (1967).

The problem of finding H to bring one set of vectors close to another can arise in many configurational problems; for instance, Mackenzie (1957) discusses it in connexion with crystallography and Downs, Liebman & McKay (1967) in connexion with vector cardiography. Downs (1972) makes a general study of orientation problems, and Downs, Liebman & McKay also propose a different definition of vector correlation.

In the next section we define the correlation statistics and give the calculations to find H and the statistics. The correlations might be used as measures of the U and V relation, with different measures being appropriate for different problems; or they might be used as test statistics for testing independence of U and V. For this purpose, null distributions will be required; these depend on the parent populations of U and V and are in general difficult to find. In §§ 3 and 4, we give some results for uniform parent populations, where, in p dimensions, the point P moves uniformly over the surface of the p-sphere, together with tests and tables.

2. Vector correlation coefficients

2·1. *Definition of correlation*

An orthogonal transformation H' is applied to u_i to give $w_i = H'u_i$. The maximum, r, of $r^* = \Sigma v_i' w_i/n$ as H is varied, will be called the vector correlation coefficient. Clearly r is positive and is 1 for perfect correlation. For some applications in two and three dimensions, one may have reason to restrict H to the class of orthogonal matrices which describe rotations. The orthogonal matrix must then have determinant $+1$, and such matrices will be called H_+ matrices. When r^* is maximized over all H_+ matrices we define the maximum, say r_+, to be the sample vector rotation-correlation coefficient. Clearly when the H which maximizes r^* overall is an H_+ matrix, $r = r_+$; otherwise $r > r_+$. Another measure of close fit would be the minimum, with respect to H or H_+, of $s^* = n^{-1}\Sigma(v_i - w_i)'(v_i - w_i)$. But $s^* = 2(1-r^*)$, so that this leads immediately to maximizing r^* as before.

Let $\operatorname{tr}(Z)$ stand for the trace of a matrix Z. Suppose that the transformed vectors w_i' are rows of a matrix W, with $W = UH$. Then $nr^* = \operatorname{tr}(V'W)$, and for the maximum of r^* we must find $\max \operatorname{tr}(H'A)$, where $A = U'V/n$ and the maximum is taken over orthogonal matrices H. Similarly $r_+ = \max \operatorname{tr}(P'A)$, where P is an H_+ matrix and the maximum is over all matrices P.

2·2. *Calculation of r and of r_+, and of the associated maximizing matrices*

The determination of H, to maximize $\operatorname{tr}(H'A)$ for given A, is an old problem, though the restriction to H_+ has attracted less interest. We quote a solution essentially given by Mackenzie (1957); it uses the singular value decomposition. Other proofs are possible. Suppose that $A'A$, positive-definite with probability one, has eigenvalues $\lambda_1 \geqslant \ldots \geqslant \lambda_p > 0$, and let Λ and Z be the diagonal matrices with $\Lambda_{ii} = \lambda_i$, and $Z_{ii} = \sqrt{\lambda_i}$ $(i = 1, \ldots, p)$. Then $A'A = C\Lambda C'$, where C is orthogonal. Write $A = GZC'$, where $G = ACZ^{-1}$; G is orthogonal, since $G'G = Z^{-1}C'A'ACZ^{-1} = Z^{-1}\Lambda Z^{-1} = I$. Then

$$\operatorname{tr}(H'A) = \operatorname{tr}(H'GZC') = \operatorname{tr}(C'H'GZ) = \operatorname{tr}(RZ),$$

where $R = C'H'G$ is orthogonal, and where Z is diagonal with positive diagonal elements. Clearly $\operatorname{tr}(RZ)$ is maximized when $R = I$, and then $H' = CG' = CZC'A^{-1} = TA^{-1}$, where we define

$$T = CZC'. \tag{1}$$

Thus we have:

(a) the matrix H_m which maximizes r^* is given by

$$H_m' = TA^{-1}; \tag{2}$$

(b) the maximum r^* is given by

$$r = \max(r^* \mid H_m) = \operatorname{tr}(H_m' A) = \operatorname{tr}(T) = \Sigma\sqrt{\lambda_i}, \tag{3}$$

where the λ_i are the eigenvalues of $A'A$.

Since $\det(T)$ is positive, $\det(H_m)$ has the same sign as $\det(A)$. The matrix A equals $U'V/n$, and comes from the data. When the data give negative $\det(A)$, we may still wish to maximize r^*, to obtain r_+, by allowing only rotation for the vectors V, as discussed above. Then suppose that the H_+ matrix which maximizes r^* is H_{+m}. Define matrix Y to be the same as matrix Z, but with a sign change for Y_{pp}, that is $Y_{ii} = \surd\lambda_i$ $(i = 1, \ldots, p-1)$, but $Y_{pp} = -\surd\lambda_p$, and define

$$r_1 = \surd\lambda_1 + \ldots + \surd\lambda_{p-1} - \surd\lambda_p.$$

Applying arguments similar to those above, we have:

(a) the H_+ matrix H_{+m} which maximizes r^* when $\det(A)$ is negative is given by $H'_{+m} = T^* A^{-1}$, where $T^* = CYC'$, and

(b) the maximum r^* is then r_1 given above. Thus in general, $r_+ = r_1$ when $\det(A)$ is negative, and $r_+ = r$ when $\det(A)$ is positive.

2·3. *Comments*

The correlation coefficients are unaltered if either U or V undergoes an orthogonal transformation before the correlation is measured. Let U be transformed to $U_1 = UH$; then $A = U'V$ becomes $A_1 = U_1'V$, different from A, but $A'A = A_1'A_1$, and the correlation coefficients, which depend on $A'A$, are unaltered. This is an important property for consistency in the definitions of r and r_+.

In certain cases, a geometric interpretation can be made of the correlation coefficients. For example, suppose that U and V are perfectly correlated, so that an H exists for which $V = UH$. Then λ_i, the eigenvalues of $A'A$, are the same as those of AA' which is now $U'U$. Suppose vector u_i makes angle θ_i with an arbitrary unit vector T; then λ_1 and λ_p are respectively the maximum and minimum values of $S = \Sigma_i \cos^2\theta_i$ as T is varied (Watson, 1965; Anderson & Stephens, 1972). If the components of the vector u_i are thought of as separate random variables, the eigenvector associated with λ_i is the ith principal component of these variables. When λ_p is very small, vectors u_i must be close to the plane perpendicular to the eigenvector corresponding to λ_p; because of the perfect correlation, vectors v_i must also be close to a plane, and in this situation r_1 and r are almost equal, that is r_+, for $\det(A)$ negative, almost equals r. It is then possible to find a rotation in two or three dimensions to produce nearly as good correlation as allowing any orthogonal transformation. For the case of two dimensions, an interesting geometrical construction exists to connect H_m and H_{+m}. More details are given in an unpublished report obtainable from the author.

Coefficients r and r_+ cannot be negative, in contrast to the situation for bivariate scalar variables, say x and y, where the usual correlation coefficient r is an estimate of a population parameter ρ which is zero if x and y are independent. With vectors, independent sets U and V can give large values of r; for example, if each set is tightly clustered around an axis.

Equation (2) may be written $H_m = AT^{-1}$, and this solution for the general case was given without proof by Downs, Liebman & McKay (1967); these authors also did not consider the rotation case as a separate problem. They defined a correlation coefficient different from those above; its distribution would seem to be more difficult than those of r and r_+, and no results were given along these lines. The authors gave their coefficient the sign of $\det(H)$. However, a value of r with a negative $\det(H)$ differs only very little from its corresponding r_+, with positive $\det(H)$, if λ_p is small, as we have seen above; this would be true also for the coefficient proposed by Downs, Liebman & McKay, so that it seems difficult to give a clear interpretation to the sign of a correlation coefficient.

2·4. *Statistics for correlation without transformation and for serial correlation*

The natural extension for correlation without moving U or V, that is correlation *in situ*, would be $r_0 = \text{tr}(U'V/n)$. The statistic nr_0 was proposed by Epp, Tukey & Watson (1971), and they discuss its permutation distribution. For serial correlation, of a single set U, Watson & Beran (1967) suggested the statistic $L = \Sigma u_i' u_{i+1}$, with $m = n-1$ terms in the sum; the natural modification to fit into the r class is $r_s = L/(n-1)$. Watson & Beran, and Epp, Tukey & Watson discuss the permutation test based on L. For both statistics we give some sampling results in § 3.

2·5. *Example*

The data below concern directions of magnetization of rock samples in three dimensions, before, U, and after, V, heat treatment in a laboratory. The sample is rather small, $n = 6$, but will serve for illustration. The basic data, omitted to save space, were the spherical polar coordinates (θ, ϕ); the three components given in U or V are $\sin\theta\cos\phi$, $\sin\theta\sin\phi$ and $\cos\theta$. The U and V data matrices, to 3 decimal places, are given in Table 1, together with matrices A, H_m and W. For this example, det(A) is positive, but if it were necessary to use r_1 as the correlation, the value, 0·8985, differs negligibly from r; this is because the smallest value of $\sqrt{\lambda}$ for $A'A$ is $0{\cdot}454 \times 10^{-3}$, and the difference between r and r_1 is therefore 0·0009. The matrix W is printed for interest, although it is not always easy from inspection to see how close W is to V. The values of $w_i' v_i$ are successively: 0·9898, 0·9841, 0·9607, 0·5928, 0·9690, 0·9002, so that the vectors match closely except for the fourth pair. Although it has no practical significance in this context, the correlation *in situ* is 0·7958.

Table 1. *Data, U and V, and matrix calculations for rock magnetism data*

	U			V	
−0·321	0·580	0·749	−0·237	0·268	0·934
−0·387	0·505	0·772	−0·293	0·283	0·914
−0·074	0·849	0·522	−0·109	0·156	0·982
0·645	0·716	0·267	−0·186	0·214	0·959
−0·188	0·455	0·870	−0·179	0·252	0·951
−0·313	0·329	0·891	−0·149	0·132	0·980

	$A = U'V/n$			H_m	
0·026	−0·026	−0·099	0·491	−0·854	0·172
−0·107	0·124	0·546	0·569	0·463	0·679
−0·133	0·149	0·646	−0·660	−0·235	0·714

	$W = UH_m$	
−0·322	0·367	0·873
−0·412	0·383	0·827
0·103	0·334	0·937
0·548	−0·282	0·788
−0·407	0·167	0·898
−0·553	0·210	0·806

Eigenvalues of $A'A$: 0·793, $0{\cdot}692 \times 10^{-4}$, $0{\cdot}206 \times 10^{-6}$;
det$(A) = 0{\cdot}336 \times 10^{-5}$;
$r = 0{\cdot}8994$, $r_1 = 0{\cdot}8985$, $r_0 = 0{\cdot}7958$.

3. Distribution theory of vector-correlation statistics

3·1. *Distribution theory for uniform parent populations*

If measures of correlation are to be used as test statistics, usually for independence of U and V, the null distributions will be needed. These will depend on the parent populations of U and V. For scalar variables x and y, the distribution of the correlation coefficient r also depends on parent populations, but is dependent only on ρ when this population is bivariate normal, and is further much simplified when $\rho = 0$. There does not appear to be such a convenient general population for vectors. However, results can be found when U and V have uniform populations for the two extremes when U and V are independent or perfectly correlated, i.e. there exists an H for which $V = UH$.

3·2. *Distribution theory for independence*

When U and V are independent, the entries in A, for large n, become normally distributed, with zero means and variances $1/(np^2)$, and with asymptotically zero covariances. Suppose that the column vectors of A are $Z_1, \ldots, Z_p$; then $A'A = \Sigma Z_i Z_i'$. Asymptotically, $A'A$ has a Wishart distribution $W(V, p)$, where $V = I_p/(np^2)$ is the covariance of each Z_i. The joint density of the λ_i can then be found, but for $p > 2$ it does not seem easy to derive the distribution of $r = \Sigma\sqrt{\lambda_i}$, and still less that of r_1. However, the asymptotic distribution of $np^2\Sigma\lambda_i$, from the diagonal terms of the Wishart matrix, is χ_p^2. This implies that $r\sqrt{n}$ and $r_1\sqrt{n}$ have asymptotic distributions, and so therefore does $r_+\sqrt{n}$; all three statistics r, r_1 and r_+ approach zero in probability as $n \to \infty$.

3·3. *Distribution theory with complete dependence*

At the other extreme, suppose now that U is a sample from a uniform population, and that V is perfectly correlated with U, that is $V = UH$. The roots of matrix $A'A$ now depend on those of UU'/n; this matrix has been considered in another connexion by Anderson & Stephens (1972). Let $t_i = \sqrt{n}(\sqrt{\lambda_i} - 1/p)$; since $r = 1$, $\Sigma\sqrt{\lambda_i} = 1$ and $\Sigma t_i = 0$. Anderson & Stephens have given the joint density of the t_i, and have shown that asymptotically $\frac{1}{2}p(p+2)\Sigma t_1^2$ has the χ_k^2 distribution, where $k = \frac{1}{2}p(p+1) - 1$. In principle, the density of t_p could be found, and hence percentage points for r_1, since $\sqrt{n}(r_1 - 1 + 2/p) = 2t_p$, but this will be difficult for $p > 3$; results for $p = 2$ and 3 are given below. However, it follows from these results that, for U uniform and U, V perfectly correlated, $\Sigma\lambda_i$ converges in probability to $1/p$ and r_1 to $1 - 2/p$. More details are given in the unpublished report referred to above.

3·4. *Further results for two and three dimensions*

When $p = 3$, with perfect correlation, we have $r_1 = 1 - 2\sqrt{\lambda_3}$ and $\sqrt{n}(r_1 - \frac{1}{3}) = -2t_3$; Anderson & Stephens (1972) have given the asymptotic density of $-t_3$, and have tabulated percentage points of $\sqrt{\lambda_3} = t_3/\sqrt{n} + \frac{1}{3}$ (Anderson & Stephens, Table 1); $\sqrt{\lambda_3}$ is their $S_{\min}$. For example, 95% and 99% points for r_1, when $n = 10$, are 0·826 and 0·884; when $n = 20$, they are 0·698 and 0·752. The corresponding asymptotic points for $\sqrt{n}(r_1 - \frac{1}{3})$ are 1·746 and 2·076.

For the circle we have more results for r and r_+. Suppose that the four entries of A are in usual notation $a_{11}, a_{12}, a_{21}, a_{22}$; it can be shown that

$$r^2 = (\sqrt{\lambda_1} + \sqrt{\lambda_2})^2 = \lambda_1 + \lambda_2 + 2\{\det(A'A)\}^{\frac{1}{2}} = \Sigma_{i,j} a_{ij}^2 + 2|\det(A)|,$$
$$r_1^2 = \Sigma_{i,j} a_{ij}^2 - 2|\det(A)|.$$

Let $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ be positive, and let $b_1 = a_{11} + a_{22}$, $b_2 = a_{12} - a_{21}$, $b_3 = a_{11} - a_{22}$, $b_4 = a_{12} + a_{21}$; then $r^2 = b_1^2 + b_2^2$ and $r_1^2 = b_3^2 + b_4^2$. If $\det(A)$ is negative, $r^2 = b_3^2 + b_4^2$ and $r_1^2 = b_1^2 + b_2^2$. When U and V are independent, $\det(A) > 0$ with probability $\frac{1}{2}$, and r^2 and r_1^2 are respectively the larger and smaller of two quantities with the same marginal distribution, that of $b_1^2 + b_2^2$; when in addition the parent populations are uniform, $2n(b_1^2 + b_2^2)$ is asymptotically χ_2^2. Then nr^2 and nr_1^2 are asymptotically distributed as the order statistics of a sample of two from the density $f(x) = e^{-x}$ $(x > 0)$. The asymptotic densities of $P = nr^2$ and $Q = nr_1^2$ are then for $x > 0$

$$f_P(x) = 2(e^{-x} - e^{-2x}), \quad f_Q(x) = 2e^{-2x}. \tag{4}$$

Finally, $r_+ = r$ or $r_+ = r_1$ each with probability $\frac{1}{2}$ in this situation; thus if $S = nr_+^2$, the asymptotic density of S must be

$$f_S(x) = \tfrac{1}{2}f_P(x) + \tfrac{1}{2}f_Q(x) = 2e^{-x}. \tag{5}$$

Equations (4) and (5) imply that asymptotically $4nr_1^2 = \chi_2^2$ and $2nr_+^2 = \chi_2^2$.

When V is perfectly correlated with U, we have $r = \sqrt{\lambda_1}+\sqrt{\lambda_2} = 1$ and $r_1 = \sqrt{\lambda_1}-\sqrt{\lambda_2}$. Let R be the length of the resultant or vector sum of the set u_i. Then $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$ are respectively $\frac{1}{2}(1+R/n)$ and $\frac{1}{2}(1-R/n)$ (Anderson & Stephens, 1972, § 4), and r_1 has the same distribution as R/n, for which exact percentage points for all n are given by Durand & Greenwood (1957) and Stephens (1969). The known asymptotic result $2R^2/n = \chi_2^2$ gives $4n(\Sigma\lambda_i - 0{\cdot}5) = \chi_2^2$ in agreement with the result for Σt_i^2 in § 3·3 above, and also $2nr_1^2 = \chi_2^2$; thus r_1^2 is stochastically half as big for independence as for perfect correlation.

3·5. *Correlation without rotation and serial correlation*

For independent uniform populations, it is straightforward to find the asymptotic distribution of r_0 defined in § 2·4. For dimension p, a term in the diagonal of A will be asymptotically normal, mean zero, variance $(np^2)^{-1}$; thus r_0 is asymptotically normal, mean zero, variance $(np)^{-1}$. Similarly, for r_s measuring serial correlation, the asymptotic distribution will be normal, mean zero, variance $\{(n-1)\,p\}^{-1}$.

For r_0 we can also find the moments for finite n, for uniform independent parents. For $p = 2$, cumulants of $r_0\sqrt{n}$ become $\kappa_2 = \frac{1}{2}$ and $\kappa_4 = -9/(16n)$; the β_1 value is 0, and $\beta_2 = 3-9/(4n)$. Similar calculations for $p = 3$ give, for $r_0\sqrt{n}$, $\kappa_2 = \frac{1}{3}$, $\beta_1 = 0$ and $\beta_2 = 3-6/(5n)$. If we approximate the distribution using the first four cumulants to fit Pearson curves we can expect excellent accuracy, and percentage points given by these approximations are given in Table 2. On the hypothesis of no serial correlation, points for r_s, for a sample of size $n+1$, will be the same as those for r_0 for a sample of size n.

4. TESTS AND TABLES OF PERCENTAGE POINTS

The above distributional results can be used for tests of the hypothesis H_0 that u and v are independent, assuming the parent populations are uniform. The procedure is as follows.

(i) Calculate $A = U'V/n$ and find eigenvalues λ_i of $A'A$; let λ_p be the smallest eigenvalue.

(ii) Then $r = \Sigma_i \lambda_i$ and $r_1 = r - 2\sqrt{\lambda_p}$. From these, find r_+ as follows. If $\det(A)$ is positive, $r_+ = r$. If $\det(A)$ is negative, $r_+ = r_1$.

(iii) The *in situ* coefficient r_0 is $\mathrm{tr}(A)$.

In making a test for independence, use r if any orthogonal transformation is permitted to obtain correlation. If a rotation only is allowed, use r_+; when $\det(A)$ is positive, this then equals r, but the rotation constraint means that a different table is used for the test. Similarly, when the data are such that $\det(A)$ is negative, one might use r_1 and the r_1 table, making the test conditional on the sign of $\det(A)$, but such an application would seem to be rare. If the test required is for correlation *in situ*, use r_0.

The statistics should be used with the appropriate part of Table 2; reject H_0 at level α if the statistic exceeds the percentage point at this level. As an illustration, suppose, in three dimensions, n is large enough to enable asymptotic points to be used, and the test is to be made at the 5% level. Then $r_0\sqrt{n}$ can go as high as 0·95 before rejection; if a rotation is allowed to bring the U nearer the V, the critical value, for $r_+\sqrt{n}$, goes up to 1·70, while if any orthogonal transformation, i.e. rotation accompanied by reflexion, is permitted, $r\sqrt{n}$ can go to 2·06.

Table 2 has been constructed from the asymptotic results given in § 3, together with Monte Carlo studies for n finite. It can be seen that the percentage points converge fairly quickly to the asymptotic values. The asymptotic points, for $r\sqrt{n}$, $p = 3$ were found from a plot of percentage points against $1/n$ for fixed values of α.

It will be useful to have a preliminary test for uniformity of U and V. A good test against clustering around one axis uses the length R of the resultant, or vector sum, of the appropriate set; exact percentage points are given by Watson (1956) and Stephens (1964, 1969). For large samples, pR^2/n is to be compared with the upper tail of χ_p^2, in p dimensions. Tests for randomness against other alternatives are given by Anderson & Stephens (1972) and Watson (1965). The tests for correlation will not apply to the example above since the R test clearly indicates that both U and V are strongly clustered. For U, $R^2 = 28{\cdot}8$, and for V, $R^2 = 35{\cdot}7$. For a more complete treatment of vector correlation, it will be necessary to deal with samples exhibiting clustering, and it is hoped to return to this subject in another paper.

There are other uses to which the basic ideas may be adapted. For instance, the vectors to be matched may not be of unit length; 12 vectors around a circle might denote the monthly incidence of a disease, for example, and one wants to compare Britain with Australia. Also, the two sets may not be paired, but it may be desired to rotate one to match the other. These points too will be discussed on another occasion.

Table 2. *Upper tail percentage points of correlation statistics for independent, uniform vectors in p dimensions*

		$p = 2$ Percentage level				$p = 3$ Percentage level			
Statistic	n	10	5	2·5	1	10	5	2·5	1
$r\sqrt{n}$	5	1·62	1·78	1·96	2·10	1·76	1·86	1·91	2·02
	10	1·67	1·85	2·03	2·20	1·81	1·95	2·05	2·18
	20	1·70	1·89	2·06	2·25	1·84	2·00	2·12	2·27
	50	1·71	1·91	2·08	2·28	1·88	2·04	2·18	2·33
	∞	1·72	1·92	2·10	2·30	1·89	2·06	2·22	2·38
$r_+\sqrt{n}$	5	1·44	1·61	1·78	1·98	1·67	1·76	1·87	1·93
	10	1·48	1·67	1·85	2·06	1·71	1·84	1·96	2·07
	20	1·50	1·70	1·89	2·10	1·73	1·88	2·00	2·15
	50	1·51	1·72	1·91	2·13	1·75	1·90	2·03	2·19
	∞	1·52	1·73	1·92	2·15	1·76	1·92	2·05	2·21
$r_1\sqrt{n}$	5	1·03	1·19	1·36	1·49	1·50	1·59	1·69	1·81
	10	1·05	1·21	1·36	1·50	1·53	1·64	1·74	1·86
	20	1·06	1·21	1·36	1·51	1·55	1·67	1·77	1·89
	50	1·07	1·22	1·36	1·52	1·56	1·69	1·78	1·91
	∞	1·07	1·22	1·36	1·52	1·57	1·70	1·79	1·92
$r_0\sqrt{n}$	5	0·933	1·166	1·352	1·547	0·751	0·952	1·120	1·306
	10	0·919	1·166	1·373	1·603	0·745	0·951	1·126	1·326
	20	0·912	1·165	1·380	1·625	0·743	0·950	1·129	1·335
	50	0·909	1·164	1·384	1·637	0·741	0·950	1·130	1·340
	∞	0·906	1·163	1·386	1·645	0·740	0·950	1·132	1·343

This work was partly supported by the National Research Council of Canada, and by the U.S. Office of Naval Research, and the author thanks both these agencies for their support. The assistance of Mr C. S. Davis with computations is also gratefully acknowledged.

References

Anderson, T. W. & Stephens, M. A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika* **59**, 613–22.

Downs, T. D. (1972). Orientation statistics. *Biometrika* **59**, 665–76.

Downs, T., Liebman, J. & McKay, W. (1967). Statistical methods for vectorcardiogram orientations. In *Proc. XI Int. Vectorcardiography Symp.*, pp. 216–22. New York: North Holland.

Durand, D. & Greenwood, J. A. (1957). Random unit vectors II: usefulness of Gram–Charlier and related series in approximating distributions. *Ann. Math. Statist.* **28**, 978–86.

EPP, R. J., TUKEY, J. W. & WATSON, G. S. (1971). Testing unit vectors for correlation. *J. Geophys. Res.* **76**, 8480–3.

MACKENZIE, J. K. (1957). The estimation of an orientation relationship. *Acta Cryst.* **10**, 61–2.

STEPHENS, M. A. (1964). The testing of unit vectors for randomness. *J. Am. Statist. Assoc.* **59**, 160–7.

STEPHENS, M. A. (1969). Tests for randomness of directions against two circular alternatives. *J. Am. Statist. Assoc.* **64**, 280–9.

WATSON, G. S. (1956). A test for randomness of directions. *Mon. Notices R. Astr. Soc., Geophys. Supp.* **7**, 160–71.

WATSON, G. S. (1965). Equatorial distributions on a sphere. *Biometrika* **52**, 193–201.

WATSON, G. S. & BERAN, R. J. (1967). Testing a sequence of unit vectors for serial correlation. *J. Geophys. Res.* **72**, 5655–9.

[*Received January* 1977. *Revised June* 1978]