STATISTICAL THEORY AND DATA ANALYSIS II

Proceedings of the Second Pacific Area Statistical Conference

edited by

K. MATUSITA
Teikyo University
Tokyo

1988

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

STATISTICAL THEORY AND DATA ANALYSIS II
K. Matusita (Editor)

A CORRELATION COEFFICIENT FOR ANGULAR VARIABLES

S. Rao JAMMALAMADAKA and Y. Ramakrishna SARMA

1. INTRODUCTION AND REVIEW

Observations in many biological and physical sciences are made in the form of directions in two or three dimensional spaces. For instance, a geologist observes the azimuths of lee faces of sand ripples, foreset planes of cross-beddings, orientation of elongate pebbles, etc. to study the currents responsible for transportation and deposition of sediments in river beds. The navigational direction of a migratory bird or a homing bird is observed by a biologist. A meteorologist observes the directions of wind, rainfall or the movement of clouds. In medical studies, observations in the form of directions are made on vector cardiographs and x-rays. Studies on any periodic phenomenon of known period such as biological rhythms can be represented in the form of directions. In all such cases, a convenient sample frame for two dimensional directions can be the circumference of a unit circle centered at the origin. Similarly the surface of a unit sphere in three dimensions centered at the origin, can be the appropriate sample space for directions in space, each point on the surface representing an observable direction. The analysis of directional data does not fit into the classical methods of linear statistical analysis. A comprehensive treatment of the analysis of directional data can be found in Mardia [11], Batschelet [1] and Watson [19]. For a survey of the nonparametric methods in this area, see S. Rao Jammalamadaka [7].

In many cases, observations are made on two or more directional variables on the same object - as in all bivariate studies. For example, a biologist may observe the wind direction and flight direction of homing birds. A geologist may wish to examine if the orientations of pebbles lying on foresets can be correlated with the foreset azimuths. In such cases one has to make precise the concept of association or correlation between two directional random variables. In what follows, the concept of correlation between two circular random variables is discussed. This aspect of statistical analysis has received the serious attention of statisticians during the last ten years or so, although tests for independence date back to MacKenzie [10]. The concept of a measure of dependence for random variables on a torus was considered by Downs [2] and more formally by Mardia in [12]. Thompson in the discussion on the paper by Mardia [12], considered the problem from the point of view of prediction and proposed a measure of correlation. However this measure was not studied further except for some comments made by Jupp and Mardia [9]. Johnson and Wehrly [8] have proposed a measure of angular correlation, based on the method of canonical correlation. The asymptotic distribution of sample angular correlation was also discussed there. Mardia and Puri [13] have proposed a scale invariant measure of correlation and some simulation studies were made to compare the sample analogue of this measure with the one proposed by Mardia [12]. Stephens [18] developed a measure of correlation based on measuring how close the unit vectors corresponding to the two angles can be brought by an orthogonal transformation. This idea was also used by MacKenzie [10] in connection with problems in crystallography and by Downs et al [3] in studying statistical methods for vector cardiogram orientations. Jupp and Mardia [9] introduced a general correlation coefficient for random variables taking values in general Riemannian manifolds and specialized it to the bivariate circular data. They also compare it with other measures proposed earlier. Rivest [17] proposed a functioand β by measuring separately the degree to which α can be predicted from β using one of the relationships $\alpha = (\pm \beta + \theta) \bmod 2\pi$ for some arbitrary fixed direction θ and constructing a measure by taking the difference of these. This leand β by measuring separately the degree to which α can be predicted from β using one of the relationships $\alpha = (\pm \beta + \theta) \bmod 2\pi$ for some arbitrary fixed direction θ and constructing a measure byand β by measuring separately the degree to which α can be predicted from β using one of the relationships $\alpha = (\pm \beta + \theta) \bmod 2\pi$ for some arbitrary fixed direction θ and constructing a measure by taking the difference of these. This leads to the definition

$$\rho_T = \frac{E \sin(\alpha_1-\alpha_2) \sin(\beta_1-\beta_2)}{[E \sin^2(\alpha_1-\alpha_2) \cdot E \sin^2(\beta_1-\beta_2)]^{1/2}} \tag{1.1}$$

where (α_1,β_1) and (α_2,β_2) are independently distributed as (α,β).

In all these studies, measures have been proposed mainly for testing the hypothesis of independence or no association between the two random variables. Another feature of these measures of correlation is that, except for those proposed by Rivest [17] and Fisher and Lee [5], they take non-negative values and then a sign is added as a secondary step in some cases. There are several practical situations where there can be negative association between the two variables and a measure which can exhibit such a negative association as well as the positive association in a natural way (much like the Pearson's product moment correlation in the linear case) is desirable. In what follows, a circular correlation coefficient which has all the desirable properties of a correlation coefficient is defined. The paper amplifies and provides some further details on a definition that was given in Rao and Sarma [16]. A sample analogue of the measure is introduced and its asymptotic distribution is obtained. The third section deals with some parametric models and illustrates how the definition given in section 2 works for these models. In these special cases, estimates of the proposed measure and their asymptotic distributions are derived. These estimates are then used to test hypotheses concerning angular correlations. Finally a nonparametric correlation coefficient is introduced in section 4.

2. A MEASURE OF CORRELATION

Let (α,β) denote a pair of random variables which are directions, both measured as angles with reference to the same zero direction and the same sense of rotation, i.e., both measured in the clockwise or in the counter clockwise direction. Let $f(\alpha,\beta)$ denote the joint pdf on the torus $0 \le \alpha,\beta < 2\pi$. Let μ and ν denote the circular mean directions of the marginal distributions of α and β respectively (cf Mardia [11] p. 45). Then we define

$$\rho_c(\alpha,\beta) = \frac{E \sin(\alpha-\mu) \sin(\beta-\nu)}{[E \sin^2(\alpha-\mu) \cdot E \sin^2(\beta-\nu)]^{1/2}} \tag{2.1}$$

as a measure of correlation between α and β. The motivation for such a definition comes from the observation that $E \sin(\alpha-\mu) = E \sin(\beta-\nu) = 0$ which corresponds to the fact in linear statistics, that the first central moment is zero. Thus $\sin(\alpha-\mu)$ and $\sin(\beta-\nu)$ can be taken to represent the deviations of α and β from their mean directions μ and ν respectively and thus the circular correlation $\rho_c(\alpha,\beta)$ is simply the product moment correlation between the sine components. Further motivation is given by the following remark 2.1.

<u>Remark 2.1:</u> One may rewrite (2.1) as

$$\rho_c(\alpha,\beta) = \frac{E\{\cos(\alpha-\beta-\mu+\nu) - \cos(\alpha+\beta-\mu-\nu)\}}{2[E \sin^2(\alpha-\mu) \cdot E \sin^2(\beta-\nu)]^{1/2}} \tag{2.2}$$

Observe that the first term in the numerator of (2.2), namely $E \cos(\alpha-\beta-\mu+\nu)$ represents how strongly the distribution of $(\alpha-\mu)-(\beta-\nu)$ is concentrated and this is a measure of the positive part of the correlation. Similarly the second term in the numerator, corresponding to the concentration of $(\alpha-\mu)+(\beta-\nu)$ distribution, measures the negative part of the correlation.

Remark 2.2: If μ or (and) ν is (are) not well defined, or arbitrary because either or both α and β have uniform marginals, then μ and ν in (2.2) are chosen in such a way that they maximize the two terms in the numerator of (2.2) individually, corresponding to the positive and negative parts of the correlation, i.e., to maximize $|E\ \cos(\alpha-\beta-\mu+\nu)|$ and $|E\ \cos(\alpha+\beta-\mu-\nu)|$. Thus, whenever there is ambiguity in the choice of the mean direction, they are chosen to yield the largest possible association in both positive and negative directions. This leads to the choice of μ and ν such that $(\mu-\nu)$ is the mean direction of $(\alpha-\beta)$ and $(\mu+\nu)$ is the mean direction of $(\alpha+\beta)$. The numerator then becomes the difference in the lengths of the mean vectors of $(\alpha-\beta)$ and $(\alpha+\beta)$. With $R_{\alpha\pm\beta} = |E(e^{i(\alpha\pm\beta)}|$, (2.2) can now be written as

$$\rho_c = (R_{\alpha-\beta} - R_{\alpha+\beta})/2\sqrt{[E\ \sin^2(\alpha-\mu)\ E\ \sin^2(\beta-\nu)]} \quad . \tag{2.3}$$

It may be noted here that in this case (of uniform marginals) our coefficient is equivalent to the one proposed by Rivest [17] in the sense that they differ only in the denominators which are like the norming constants. It should, however, be noted that the measure in (2.1) is different from that proposed by Rivest [17] in the general case and the two are arrived at from entirely different motivations.

Remark 2.3: The definition of the angular correlation ρ_c as given in (2.1) or (2.2) still makes sense even if (α,β) have support which is less than the full circle, say a half circle.

Remark 2.4: The primary motivation for the definition (2.1) is to have a suitable circular version of Pearson's product moment correlation defined by

$$\frac{E[(X-E(X))\ (Y-E(Y))]}{[E((X-E(X))^2)\ E((Y-E(Y))^2)]^{1/2}} \tag{2.4}$$

Notice however that (2.4) can be written equivalently as

$$\frac{E((X_1-X_2)(Y_1-Y_2))}{[E((X_1-X_2)^2)\ E((Y_1-Y_2)^2)]^{1/2}} \tag{2.5}$$

where (X_1,Y_1) and (X_2,Y_2) are independently distributed as (X,Y). While our definition (2.1) is the circular analogue of (2.4), Fisher and Lee's [5] definition (1.1) is the circular analogue of (2.5). In the circular case, the definitions (1.1) and (2.1) are algebraically different and yield different values. The example in Johnson and Wehrly [8] gives a value of 0.1914 for the sample analogue of (1.1) while the measure we proposed in (2.1) has a value of 0.2620.

Remark 2.5: In section 3 estimation and testing problems concerning ρ_c are dealt with, when some underlying population models are assumed. In these cases, we consider methods of estimation such as MLE and exact tests for hypotheses about ρ_c. It is shown below that when no further assumptions about the population model are made, one can use $r_{c,n}$ for estimating ρ_c as well as for testing hypotheses concerning ρ_c.

The following result summarizes the properties of ρ_c.

Theorem 2.1: The circular correlation coefficient $\rho_c(\alpha,\beta)$ in (2.1) or (2.2) satisfies the following properties.

(a) $\rho_c(\alpha,\beta)$ does not depend on the zero direction used for either variable.

(b) $\rho_c(\alpha,\beta) = \rho_c(\beta,\alpha)$

(c) $-1 \le \rho_c(\alpha,\beta) \le 1$

(d) $\rho_c(\alpha,\beta) = 0$ if α and β are independent, although the converse need not be true.

(e) If α and β have full support, $\rho_c(\alpha,\beta) = 1$ iff $\alpha = \beta + \text{const}(\text{mod } 2\pi)$ and $\rho_c(\alpha,\beta) = -1$ iff $\alpha+\beta = \text{const}(\text{mod } 2\pi)$.

(f) $\rho_c(\alpha,\beta) = \rho(\alpha,\beta)$ the product moment correlation if α and β are unimodal and are highly concentrated.

Proof: All the properties are easily verified except for (e). To prove (e), suppose $\alpha = \beta+\delta$ (a.e). Then $\mu = \nu+\delta$ and consequently $(\alpha-\mu) = (\beta-\nu)$ and thus $\sin(\alpha-\mu) = \sin(\beta-\nu)$ (a.e) which implies $\rho_c(\alpha,\beta) = 1$.

To prove the converse, one has by Cauchy-Schwarz inequality that $\rho_c(\alpha,\beta) = \pm 1$ iff $\sin(\alpha-\mu) = b\sin(\beta-\nu)$ a.e for some $b \neq 0$. Now suppose $\rho_c(\alpha,\beta) = 1$. Then clearly $b > 0$. Moreover b has to be 1 in view of the fact that the ranges of $(\alpha-\mu)$ and $(\beta-\nu)$ are the same. Thus it follows that $\alpha-\mu = \beta-\nu$ a.e or $\alpha = \beta+\delta$ (a.e) where $\delta = (\mu-\nu)$.

The other part can be proved in a similar way. □

To define a sample correlation coefficient, let (α_i, β_i) $i = 1,\ldots,n$ be a sample of size n from some bivariate circular distribution. Then one may define the sample analogue of ρ_c as follows:

$$r_{c,n} = \frac{\sum_1^n \sin(\alpha_i - \bar{\alpha}_n)\sin(\beta_i - \bar{\beta}_n)}{\left(\sum_1^n \sin^2(\alpha_i - \bar{\alpha}_n) \cdot \sum_1^n \sin^2(\beta_i - \bar{\beta}_n)\right)^{1/2}} \tag{2.6}$$

where $\bar{\alpha}_n$ and $\bar{\beta}_n$ are the resultant mean directions of $\alpha_1,\ldots,\alpha_n$ and $\beta_1,\ldots,\beta_n$ respectively.

Let $\lambda_{ij} = E(\sin^i(\alpha-\mu)\sin^j(\beta-\nu))$ for $i,j = 0,1,2,3,4$.

Define

$$T_{1,n} = \frac{1}{\sqrt{n}} \sum_1^n \sin(\alpha_i - \bar{\alpha}_n)\sin(\beta_i - \bar{\beta}_n)$$

$$T_{2,n} = \frac{1}{\sqrt{n}} \sum_1^n \sin^2(\alpha_i - \bar{\alpha}_n) \quad \text{and} \quad T_{3,n} = \frac{1}{\sqrt{n}} \sum_1^n \sin^2(\beta_i - \bar{\beta}_n) .$$

Then by the multivariate central limit theorem one has that

$T_n' = (T_{1,n}, T_{2,n}, T_{3,n})$ is asymptotically $N_3(\underset{\sim}{\lambda}, \Sigma)$ where

$\underset{\sim}{\lambda}' = (\lambda_{11}, \lambda_{20}, \lambda_{02})$ and

$$\Sigma = \begin{bmatrix} \lambda_{22}-\lambda_{11}^2 & \lambda_{31}-\lambda_{20}\lambda_{11} & \lambda_{13}-\lambda_{02}\lambda_{11} \\ \cdot & \lambda_{40}-\lambda_{20}^2 & \lambda_{22}-\lambda_{20}\lambda_{02} \\ \cdot & \cdot & \lambda_{04}-\lambda_{02}^2 \end{bmatrix}$$

From this one immediately has the following

<u>Theorem 2.2</u>: If neither marginal is uniform, $\sqrt{n}(r_{c,n} - \rho_c(\alpha,\beta))$ converges in distribution to $N(0,\sigma^2)$ where

$$\sigma^2 = \frac{\lambda_{22}}{\lambda_{20}\lambda_{02}} - \rho_c(\alpha,\beta) \left[\frac{\lambda_{13}}{\lambda_{20}\sqrt{\lambda_{20}\lambda_{02}}} + \frac{\lambda_{31}}{\lambda_{02}\sqrt{\lambda_{20}\lambda_{02}}}\right]$$

$$+ \frac{\rho_c^2(\alpha,\beta)}{4} \left[1 + \frac{\lambda_{40}}{\lambda_{20}^2} + \frac{\lambda_{04}}{\lambda_{02}^2} + \frac{\lambda_{22}}{\lambda_{20}\lambda_{02}}\right] .$$

Proof: Writing $r_{c,n} = T_{1,n}/\sqrt{T_{2,n}T_{3,n}}$ the result follows by an application of the so called δ-method for functions of random variables (see eg. Rao [14] p. 387) after some algebraic simplifications.

Corollary 2.3: If α and β are uncorrelated, i.e., $\rho_c = 0$ then $\sqrt{n}\cdot r_{c,n}$ converges in distribution to $N(0,\lambda_{22}/\lambda_{20}\lambda_{02})$.

Remark: When n is sufficiently large, λ_{ij}'s can be replaced by their estimates

$$\hat{\lambda}_{ij} = \frac{1}{n} \sum_{\ell=1}^{n} \sin^i(\alpha_\ell - \bar{\alpha}_n) \sin^j(\beta_\ell - \bar{\beta}_n) .$$

Remark: This result suggests that even without any further assumptions being made about the model, one can use $r_{c,n}$ for estimating ρ_c as well as for testing hypotheses concerning ρ_c.

3. SOME PARAMETRIC MODELS

3.1 Wrapped bivariate normal distribution.

Let $\underset{\sim}{X}' = (X_1,X_2)$ be a bivariate normal vector with mean $\underset{\sim}{\mu}' = (\mu_1,\mu_2)$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \cdot & \sigma_2^2 \end{pmatrix}$. Let $\alpha = X_1 \pmod{2\pi}$ and $\beta = X_2 \pmod{2\pi}$. The distribution of (α,β) is the wrapped bivariate normal with characteristic function given by

$$\varphi(m,n) = \exp\{i(m\mu_1+n\mu_2) - \frac{1}{2}(m\sigma_1^2+2mn\sigma_{12}+n\sigma_2^2)\} \tag{3.1}$$

where m and n are integers. To evaluate $\rho_c(\alpha,\beta)$ in this case it is convenient to use (2.2).

Let $\delta^+ = (\alpha+\beta)(\text{mod } 2\pi) = (X_1+X_2)(\text{mod } 2\pi)$

$\delta^- = (\alpha-\beta)(\text{mod } 2\pi) = (X_1-X_2)(\text{mod } 2\pi)$.

Then δ^+ is $WN(\mu_1+\mu_2, \sigma_1^2 + \sigma_2^2 + 2\sigma_{12})$ and δ^- is $WN(\mu_1-\mu_2, \sigma_1^2 + \sigma_2^2 + 2\sigma_{12})$ where WN denotes a wrapped normal distribution. Thus the length of polar vector for δ^+ is

$$\zeta(\delta^+) = |E(e^{i\delta^+})| = \exp\left(-\frac{1}{2}(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})\right)$$

and that of δ^- is

$$\zeta(\delta^-) = \exp\left(-\frac{1}{2}(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})\right) .$$

Similarly the lengths of the polar vectors of 2α and 2β are

$\zeta(2\alpha) = \exp(-2\sigma_1^2)$ and $\zeta(2\beta) = \exp(-2\sigma_2^2)$ respectively.

Further the mean directions of α and β are $\mu = \mu_1 \pmod{2\pi}$ and $\nu = \mu_2 \pmod{2\pi}$ and the mean directions of 2α and 2β are $\bar{\alpha}_2 = 2\mu_1$ (mod 2π) and $\bar{\beta}_2 = 2\mu_2 \pmod{2\pi}$. Finally the mean directions of δ^+ and δ^- are $\bar{\delta}^+ = (\mu_1+\mu_2)\pmod{2\pi}$ and $\bar{\delta}^- = (\mu_1-\mu_2)\pmod{2\pi}$ respectively. Using these expressions, after some algebraic simplifications, one has

$$\rho_c(\alpha,\beta) = \frac{\zeta(\delta^-)\cos[\bar{\delta}^- -(\mu-\nu)] - \zeta(\delta^+)\cos[\bar{\delta}^+-(\mu+\nu)]}{[\{1-\zeta(2\alpha)\cos(\bar{\alpha}_2-2\mu)\}\{1-\zeta(2\beta)\cos(\bar{\beta}_2-2\nu)\}]^{1/2}} \tag{3.2}$$

Now substituting the values, one finally has

$$\rho_c(\alpha,\beta) = (\sinh \sigma_{12})/[\sinh \sigma_1^2 \sinh \sigma_2^2]^{1/2} \tag{3.3}$$

Remark 3.1: The expression $\rho_c(\alpha,\beta)$ in this case coincides with λ_2 obtained by Johnson and Wehrly [8] except that here the sign of association between α and β is naturally incorporated.

Remark 3.2: From the expression for $\rho_c(\alpha,\beta)$ it may be noted that $\rho_c(\alpha,\beta) = 0$ if X_1 and X_2 are independent. However even if X_1 and X_2 are perfectly correlated in the linear sense, it does not follow that $\rho_c = 1$ unless $\sigma_1 = \sigma_2$. In this case, of course $\alpha = \beta$+const. from (e) of Theorem 2.1. Similarly when the linear correlation between X_1 and X_2 is -1 and $\sigma_1 = \sigma_2$, then $\rho_c = -1$.

<u>Estimation of $\rho_c(\alpha,\beta)$</u>: The Maximum likelihood equations for estimation the parameters in this case are not tractable. So the parameters are estimated using the method of moments on the random variables $Z_1 = \exp(i\alpha)$ and $Z_2 = \exp(i\beta)$. Since ρ_c does not depend on μ_1 and μ_2, for simplicity, they are assumed to be zero. Then equating the theoretical moments with the corresponding sample moments, one has the estimating equations

$$E(Z_1) = E(\cos\alpha) = e^{-\frac{1}{2}\sigma_1^2} = \frac{1}{n}\sum_1^n \cos\alpha_i = \bar{X} \text{ (say)}$$

$$E(Z_2) = E(\cos\beta) = e^{-\frac{1}{2}\sigma_2^2} = \frac{1}{n}\sum_1^n \cos\beta_i = \bar{Y} \text{ (say)} \tag{3.4}$$

$$E(Z_1Z_2) = E\cos(\alpha+\beta) = \exp(-\tfrac{1}{2}(\sigma_1^2+\sigma_2^2+2\sigma_{12}))$$

$$= \frac{1}{n}\Sigma\cos(\alpha_i+\beta_i) = \bar{Z} \quad \text{(say)}$$

Solution of these equations leads to

$$\hat{\sigma}_1^2 = -2\log\bar{X}, \qquad \hat{\sigma}_2^2 = -2\log\bar{Y} \quad \text{and} \quad \hat{\sigma}_{12} = \log(\bar{X}\bar{Y}/\bar{Z}). \tag{3.5}$$

It is interesting to note that the minimum distance method which minimizes the sum of the distances of the observed $(\cos\alpha_i, \cos\beta_i, \cos(\alpha_i+\beta_i))$ from their expectations suggested in Rao, J.S. [15], leads to the same estimates, which are shown there to be consistent and asymptotically normal. In fact by an application of multivariate central limit theorem, it follows that $(\bar{X},\bar{Y},\bar{Z})$ is asymptotically normal with mean vector

$$\left(e^{-\frac{1}{2}\sigma_1^2},\ e^{-\frac{1}{2}\sigma_2^2},\ e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2+2\sigma_{12})}\right)$$

and covariance matrix

$$\Sigma = \begin{bmatrix} \dfrac{(1-e^{-\sigma_1^2})^2}{2} & e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}\left(\dfrac{e^{\sigma_{12}}+e^{-\sigma_{12}}}{2}-1\right) & \dfrac{e^{-\frac{1}{2}\sigma_2^2}}{2}\left[1-e^{-(\sigma_1^2+\sigma_{12})}\right]^2 \\ * & \dfrac{(1-e^{-\sigma_2^2})^2}{2} & \dfrac{e^{-\frac{1}{2}\sigma_1^2}}{2}\left[1-e^{-(\sigma_2^2+\sigma_{12})}\right]^2 \\ * & * & \dfrac{1+e^{-2(\sigma_1^2+\sigma_2^2+2\sigma_{12})}}{2} \end{bmatrix}$$

This in turn yields the asymptotic normality of the estimates in (3.5). Finally one has that the estimated correlation coefficient

$$\hat{\rho}_c(\alpha,\beta) = \sinh \hat{\sigma}_{12}/[\sinh \hat{\sigma}_1^2 \sinh \hat{\sigma}_2^2]^{1/2} \quad . \tag{3.6}$$

is a consistent estimate of ρ_c and is asymptotically normal with mean ρ_c and variance

$$1 + \tanh \frac{\sigma_1^2}{2} \tanh \frac{\sigma_2^2}{2} - \frac{2\rho_c \cosh \sigma_{12}}{(\sinh \sigma_1^2 \sinh \sigma_2^2)^{1/2}}$$

$$\cdot \left[\frac{1}{\tanh \sigma_1^2} (2 \sinh \tfrac{1}{2} \sigma_1^2 + \cosh \sigma_{12} - 1 - 2 \sinh^2 \tfrac{1}{2}(\sigma_1^2+\sigma_{12})) \right.$$

$$\left. + \frac{1}{\tanh \sigma_2^2} (2 \sinh \tfrac{1}{2} \sigma_2^2 + \cosh \sigma_{12} - 1 - 2 \sinh^2 \tfrac{1}{2}(\sigma_2^2+\sigma_{12})) \right]$$

$$+ 2\rho_c^2 \left[\frac{\sinh^2 \frac{1}{2} \sigma_1^2}{\tanh^2 \sigma_1^2} + \frac{\sinh^2 \frac{1}{2} \sigma_2^2}{\tanh^2 \sigma_2^2} \right] \tag{3.7}$$

Remark 3.3: The estimate obtained here can also be used for large sample testing of hypotheses concerning ρ_c.

3.2 A model with uniform marginals

Suppose (α,β) is a random vector such that α has the uniform distribution on the circle and the conditional distribution of β given α is circular normal with density function

$$f(\beta|\alpha : \theta,a,k) = \frac{1}{2\pi I_0(k)} e^{k \cos(\beta-a\alpha-\theta)} \tag{3.8}$$

where $k \geq 0$, $0 \leq \theta < 2\pi$ and $a = \pm 1$ are the unknown parameters and $I_p(k)$ is the modified Bessel function of first kind and p^{th} order.

Notice that the marginal distribution of β is also uniform on the circle and the joint distribution of α and β is given by the density function

$$f(\alpha,\beta : \theta,a,k) = \frac{1}{4\pi^2 I_0(k)} \exp \{k \cos(\beta-a\alpha-\theta)\} \; . \tag{3.9}$$

The model with $a = 1$ has been considered before by Johnson and Wehrly [8]. In this example, since α and β have uniform marginals, μ and ν are arbitrary and will be chosen later as stated in Remark 2.2. Suppose for definiteness that $a = 1$. In the numerator of ρ_c in the definition (2.2), the second term is zero since

$$E\cos(\alpha+\beta-\mu-\nu) = E_\alpha E_{\beta|\alpha}\cos[\beta-\alpha-\theta+(2\alpha+\theta-\mu-\nu)]$$

$$= \frac{I_1(k)}{I_0(k)} E_\alpha \cos(2\alpha+\theta-\mu-\nu) = 0$$

The penultimate step here follows from (3.8) and the last step because α is uniform. Since μ and ν are arbitrary, their choice is made so as to maximize $E\cos[\alpha-\beta-(\mu-\nu)]$.

This happens when $(\nu-\mu)$ is the mean direction of $(\beta-\alpha)$ which from (3.9), is θ. Also from (3.9)

$$E\cos(\beta-\alpha-\theta) = I_1(k)/I_0(k) .$$

Combining this with the fact that $E\sin^2(\alpha-\mu) = E\sin^2(\beta-\nu) = \frac{1}{2}$ (because of the uniform marginals), one obtains

$$\rho_c = I_1(k)/I_0(k). \tag{3.10}$$

Similarly when $a = -1$ it can be shown that

$$\rho_c = -I_1(k)/I_0(k). \tag{3.11}$$

<u>Remark 3.4</u>: In this model again, the correlation coefficient ρ_c coincides with the one given by Johnson and Wehrly [8] but only when $a = 1$. In the present discussion, the possibility of negative correlation is also permitted.

<u>Estimation of ρ_c</u>: In this case the maximum likelihood estimates of θ, k and $a = \pm 1$ may be obtained as follows. Fix $a = +1$. Then the ML estimates of the parameters say, k_+ and θ_+ denoted by $\hat{k}_+$ and $\hat{\theta}_+$, are obtained from the usual equations

$$\frac{\partial \log L}{\partial \theta} = \sum_1^n \sin(\beta_i-\alpha_i-\theta) = 0 \qquad \text{and}$$

$$\frac{\partial \log L}{\partial k} = n\frac{I_1(k_+)}{I_0(k_+)} - \Sigma\cos(\beta_i-\alpha_i-\theta_+) = 0 \tag{3.12}$$

where $L(a,k,\theta)$ denotes the likelihood of the sample. Denote $(\beta_i-\alpha_i) = \delta_i^-$ and $(\beta_i+\alpha_i) = \delta_i^+$. Then the equations (3.12) give $\hat{\theta}_+ = \bar{\delta}^-$ and $\hat{k}_+$ is obtained from $A(\hat{k}_+) = \frac{1}{n} R_{\delta^-}$, where $\bar{\delta}^-$ and R_{δ^-} are the mean direction and the resultant length of $\delta_1^-,\ldots,\delta_n^-$ respectively and $A(k) = I_1(k)/I_0(k)$. Similarly, fixing $a = -1$, the ML estimates of the parameters say, $\hat{k}_-$ and $\hat{\theta}_-$ are obtained as $\hat{\theta}_- = \bar{\delta}^+$ and $A(\hat{k}_-) = \frac{1}{n} R_{\delta^+}$ where $\bar{\delta}^+$ and R_{δ^+} are the mean direction and the resultant length of $\delta_1^+,\ldots,\delta_n^+$ respectively. Finally a is estimated to be that value of $\hat{a}$ which corresponds to the maximum of $L(1,k_+,\theta_+)$ and $L(-1,k_-,\theta_-)$. This, however, amounts to choosing

$$\hat{a} = \begin{cases} 1 & \text{if} \quad R_{\delta^-} > R_{\delta^+} \\ -1 & \text{if} \quad R_{\delta^-} < R_{\delta} \end{cases} \tag{3.13}$$

since
$$L(1,\hat{k}_+,\hat{\theta}_+) = \text{const.} I_0(\hat{k}_+)^{-n} \exp(\hat{k}_+ R_{\delta^-}) = g(R_{\delta^-}) \text{ , say}$$

and
$$L(-1,\hat{k}_+,\hat{\theta}_+) = \text{const.} I_0(\hat{k}_-)^{-n} \exp(\hat{k}_+ R_{\delta^+}) = g(R_{\delta^+})$$

where $g(\cdot)$ is a monotonic increasing function (Mardia, [11] p. 134). Thus the ML estimate of $\rho_c(\alpha,\beta)$ is

$$\hat{\rho}_c(\alpha,\beta) = \hat{a}\cdot A(\hat{k}) = \begin{cases} \frac{1}{n} R_{\delta^-} & \text{if} \quad R_{\delta^-} > R_{\delta^+} \\ \frac{-1}{n} R_{\delta^+} & \text{if} \quad R_{\delta^-} < R_{\delta^+} \end{cases} \tag{3.14}$$

<u>Remark 3.5</u>: Mardia [12] considered the case where α and β have uniform marginals and proposed the correlation coefficient

$$r = \max(\bar{R}_{\delta^-},\bar{R}_{\delta^+}) = |\rho_c(\alpha,\beta)|.$$

Thus the correlation coefficient proposed here coincides with that proposed by Mardia [12] in absolute value and moreover incorporates the sign of association in a natural way.

<u>Test for $H_0 : \rho_c(\alpha,\beta) = 0$</u>: This hypothesis is equivalent to $k = 0$. Under this hypothesis, α and β are independent and have uniform distributions on the circle (cf (3.9)). Hence both $\delta^+ = \beta+\alpha$ and $\delta^- = \beta-\alpha$ have uniform distribution on the circle and moreover they are independent. Hence R_{δ^+} and R_{δ^-} are i.i.d random variables.

(a) To test $\rho_c = 0$ vs $\rho_c > 0$ (or $\rho_c < 0$) the test is based on R_{δ^-} (or R_{δ^+}). In either case the table for Rayleigh's test (Table 2.5 Mardia [11], p.300) can be used without any modification.

(b) Suppose now that the alternative is two sided, $\rho_c \neq 0$. Then large values of $|\rho_c(\alpha,\beta)|$ lead to the rejection of H_0. This is equivalent to rejection of H_0 for large values of $\max(R_{\delta^+}, R_{\delta^-})$. Since, under H_0, the probability density function of R_{δ^+} (or R_{δ^-}) is that of the resultant length of n random unit vectors, the pdf and cdf respectively are given by,

$$f_n(r) = r \int_0^\infty tJ_0(tr)J_0^n(t)dt \qquad (3.15)$$

$$F_n(r) = r \int_0^\infty J_1(tr)J_0^n(t)dt \qquad \text{for} \quad 0 \leq r \leq n$$

where $J_k(t)$ are the standard Bessel functions of first kind. Hence the pdf of $\max(R_{\delta^+}, R_{\delta^-})$ is

$$g_n(r) = 2f_n(r)F_n(r) \qquad \text{for} \quad 0 \leq r \leq n.$$

Since the upper $100(1-\alpha)$ percentage point of $g_n(r)$ corresponds to the $100\sqrt{(1-\alpha)}$ percentage point of $f_n(r)$ one may adapt the table for Rayleigh's test mentioned above. On the other hand, if n is large, writing the approximation that $\frac{2}{n} R_{\delta^+}^2$ is $\chi^2(2)$, $g_n(r)$ may be approximated by

$$g_n^*(r) = \frac{4r}{n} e^{-r^2/n} [1-e^{-r^2/n}]. \qquad (3.16)$$

A table of percentage points for $n\rho^2 = \frac{1}{n}[\max(R_{\delta^+}, R_{\delta^-})]^2$ based on this approximation as well as on Monte Carlo results, is given by Stephens [18] and is reproduced partly in Batschelet ([1] p. 183).

(c) Similarly tests for the hypothesis $\rho_c = \rho_c^0$ can be converted into tests for k and critical values may be obtained using the known distributions of the resultant lengths for circular normal samples.

4. A NONPARAMETRIC OR RANK CORRELATION FOR ANGULAR VARIABLES.

The purpose of this section is to demonstrate how definition (2.1) can be used to obtain a rank correlation measure.

Given (α_i, β_i), $i = 1,\ldots,n$, a random sample of size of n from some bivariate circular distribution, measured with respect to some arbitrary zero direction and sense of rotation these are first converted into uniform scores (Batschelet [1] p.186). That is, one of these sets of variables, say $(\alpha_1,\ldots,\alpha_n)$ are linearly ranked and then let r_i denote the rank of the β_i which corresponds to the ith largest α_i for $i = 1,\ldots,n$. Thus the observations (α_i, β_i), $i = 1,\ldots,n$ are converted into ranks (i, r_i), $i = 1,\ldots,n$ and are replaced by the uniform scores defined by

$$\varphi_i = \frac{2\pi i}{n} \quad \text{and} \quad \psi_i = \frac{2\pi r_i}{n} \tag{4.1}$$

for $i = 1,\ldots,n$. The nonparametric (or rank) circular correlation η_c is defined now as the circular correlation ρ_c for the uniform scores (φ_i, ψ_i), $i = 1,\ldots,n$. Using definition (2.1)

$$\eta_c = \frac{\sum_1^n \sin(\varphi_i - \bar{\varphi})\sin(\psi_i - \bar{\psi})}{[\Sigma \sin^2(\varphi_i - \bar{\varphi})\Sigma \sin^2(\psi_i - \bar{\psi})]^{1/2}} . \tag{4.2}$$

Noting that $\sum_1^n \sin^2\left(\frac{2\pi i}{n} - \bar{\varphi}\right) = \frac{n}{2}$ for any choice of $\bar{\varphi}$, (4.2) reduces to

$$\eta_c = \frac{1}{n}\Sigma \cos[\varphi_i - \psi_i - (\bar{\varphi} - \bar{\psi})] - \frac{1}{n}\sum_1^n \cos[\varphi_i + \psi_i - (\bar{\varphi} + \bar{\psi})] . \tag{4.3}$$

Now since $\bar{\varphi}$ and $\bar{\psi}$ are arbitrary, they are chosen again (see Remark 2.2) to maximize the two terms on the RHS of (4.3) individually. This maximization occurs when $(\bar{\varphi} - \bar{\psi})$ is the resultant direction of $\delta_i^- = (\varphi_i - \psi_i)$, $i = 1,\ldots,n$

and $(\overline{\varphi+\psi})$ is the resultant direction of $\delta_i^+ = (\varphi_i+\psi_i)$, $i = 1,\ldots,n$. Thus

$$\eta_c = \frac{1}{n} R_{\delta^-} - \frac{1}{n} R_{\delta^+} \tag{4.4}$$

The coefficient η_c being the difference in the lengths of the resultants is clearly invariant under any choice of zero direction or sense of rotation. Mardia [12] suggests the coefficient which is max $\left(\frac{R_{\delta^-}}{n}, \frac{R_{\delta^+}}{n}\right)$ and adding the sign corresponding to the larger resultant. Here (4.4) on the other hand, incorporates the sign naturally depending on whether the positive or negative part of the correlation dominates the other.

One may compare this coefficient with that proposed by Fisher and Lee [4]. Our measure (4.4) differs from their definition in the same way that our ρ_c (cf eqn. (2.1)) differs from their circular correlation (1.1)(cf. Fisher and Lee [5]).

REFERENCES

[1] Batschelet (1981). Circular Statistics in Biology, Academic Press, New York.

[2] Downs, T.D. (1974). Rotational angular correlations (in) Biorthythms in Human Reproduction. Eds Ferin, M. et al, 91-104, John-Wiley, New York.

[3] Downs, T.D., Liebman, J. and Mckay, W. (1967). Statistical methods for vector cardiogram orientations. Proc. XI. Int. Vector cardiography Symp. 216-222, North Holland, New York.

[4] Fisher, N.I. and Lee, A.J. (1982). Nonparametric measures of angular-angular association. Biometrika, 69, 315-22.

[5] Fisher, N.I. and Lee, A.J. (1983). A correlation coefficient for circular data. Biometrika, 70, 327-332.

[6] Greenwood, J.A. and Durand, D. (1955). The distribution of the length and components of the sum of n random unit vectors. Ann. Math. Statist. 26, 233-246.

[7] Jammalamadaka, S. Rao (1984), Nonparametric methods in directional data analysis, Handbook of Statistics, Vol. 4 (Eds. P.R. Krishnaiah and P.K. Sen) 755-770, North Holland, New York.

[8] Johnson, R.A. and Wehrly, T.E. (1977). Measures and models for angular correlations and angular-linear correlation. J. Roy. Statist. Soc. B. 39, 222-229.

[9] Jupp, P.E. and Mardia, K.V. (1980). A general correlation coefficient for directional data and related regression problems. Biometrika, 67, 163-173.

[10] MacKenzie, J.K. (1957). The estimation of an orientational relationship. Acta Cryst. 10, 61-62.

[11] Mardia, K.V. (1972). Statistics of Directional Data. Acad. Press, New York.

[12] Mardia, K.V. (1975). Statistics of directional data. (with discussion) J. Roy. Statist. Soc. B, 37, 349-393.

[13] Mardia, K.V. and Puri, M.L. (1978) A spherical correlation coefficient robust against scale, Biometrika, 65, 391-395.

[14] Rao, C.R. (1973). Linear Statistical Inference and its Applications 2nd Ed. John Wiley, New York.

[15] Rao, J.S. (1969). Some Contributions to the Analysis of Circular Data, thesis submitted for Ph.D. of ISI, Calcutta.

[16] Rao, J.S. and Sarma, Y.R. (1983). A measure of association for circular data. Proc. ISI, 44th session, Vol. 2, 571-574.

[17] Rivest, P. (1982). Some statistical methods for bivariate circular data. J. Roy. Statist. Soc. B. 37, 81-90.

[18] Stephens, M.A. (1979). Vector correlation. Biometrika, 66, 41-48.

[19] Watson, G.S. (1984). Statistics on Spheres. Wiley-Interscience, John Wiley, New York.

University of California
Indian Statistical Institute