Statistica Sinica 4(1994), 729-747

TESTING INDEPENDENCE OF BIVARIATE CIRCULAR DATA AND WEIGHTED DEGENERATE U-STATISTICS

Grace S. Shieh, Richard A. Johnson and Edward W. Frees

University of Missouri and University of Wisconsin

Abstract: A class of new statistics for testing independence of bivariate circular data is obtained by averaging a "weighted Kendall's tau" over its marginals. The averaging is done by first fixing the two origins, calculating a weighted Kendall's tau rank statistic and then averaging over cyclic permutations of the two sets of ranks. These statistics are based on ranks, are distribution-free, and are invariant under different choices of origin and rotation. They could, for instance, be applied to testing independence of bird flight and prevailing wind direction. We obtain the asymptotic distribution of our rank statistic as a special case of a general class of statistics, called weighted degenerate U-statistics. Let $Z_1, \ldots, Z_n$ be i.i.d. r.v.'s with $E(Z_1) = 0$ and $\mathrm{Var}(Z_1) = 1$, and $E(Z_1^4) < \infty$. We define a weighted degenerate U-statistic as $WU_n = \sum_{i \neq j} d_{ijn} h(Z_i, Z_j)$. Here, $\{d_{ijn}\}$ are non-stochastic weights and h is degenerate in the sense that $\mathrm{Var}[E(h(Z_1, Z_2)|Z_2)] = 0$. Under regularity conditions, the limit distribution of WU_n is shown to be a linear combination of independent chi square random variables. It is interesting that a special case (using equal weights) of our general procedures, a Circular Kendall's tau, turns out to be equivalent to a statistic proposed by Fisher and Lee (1982). A power study and an application are presented.

Key words and phrases: Directional data, Kendall's tau, limit distribution, rank correlation.

1. Introduction

Researchers are sometimes confronted with bivariate circular data, for instance, the direction of bird flight and the prevailing wind direction occurring in biology. A problem of concern is whether or not these two directions are independent. Many test statistics have been proposed, but not many of them are distribution-free (Jupp and Mardia (1980) and Jupp and Spurr (1985)). When there is no natural joint distribution for the two orientations, the distribution-free property is desirable. In this paper, we introduce a class of distribution-free statistics for testing coordinate independence of bivariate circular data. In Section 2, these statistics are derived from a "weighted Kendall's tau". These statistics are invariant under different choices of origin and direction of rotation. When the origin of a circle is fixed, each circle is equivalent to a line. For any fixed origins of the two circles, these statistics are functions of ranks. Further, these statistics are invariant under different choices of origin and direction of rotation on both circles. Hence they are distribution-free on the torus which is the cross product of the two unit circles. To obtain their limit distributions, which are linear combinations of independent chi square variates with one degree of freedom, new results are derived concerning weighted degenerate U-statistics. Section 3 is devoted to developing these limit distributions. In Section 4, we investigate the asymptotic and finite sample properties of a particular test statistic which is derived from Kendall's tau. We term it a Circular Kendall's tau. A power study and an application of Circular Kendall's tau to estimate angular-angular association in a set of isotropic data are presented. We conclude with some remarks in Section 5.

1.1. The problem of testing independence

Let $\mathbf{X}_i = (\Theta_i, \Phi_i), 1 \leq i \leq n$, be a vector of angles, for instance of bird flight and wind direction, observed at time i. Further, let $\mathbf{R}_i = (R\Theta_i, R\Phi_i)$ be the vector of ranks of Θ_i and Φ_i with respect to some fixed but arbitrary origins on the two circles, respectively. We assume that Θ_i's and Φ_i's are i.i.d. r.v.'s with continuous distributions F and G, respectively. These marginal distributions may be uniform $(0, 2\pi)$, a wrapped normal distribution on the unit circle, or any number of other distributions. Further, denote the joint distribution of Θ_i's and Φ_i's by H. The problem of testing independence may be formulated as:

$$\begin{aligned} H_0 &: H(\theta_1, \phi_1) = F(\theta_1)G(\phi_1) \quad \text{for all } (\theta_1, \phi_1) \text{ versus} \\ H_1 &: H(\theta_1, \phi_1) \neq F(\theta_1)G(\phi_1) \quad \text{for some } (\theta_1, \phi_1). \end{aligned}$$

Under independence, the joint distribution will factor whatever the choice of the two origins.

1.2. Short literature review

Test statistics for coordinate independence of bivariate circular data or circular correlation coefficients may be classified into the following three types. The first type consists of functions of the corrected covariate matrix of p-dimensional directions $\mathbf{x}$ and $\mathbf{y}$. Statistics proposed by Downs (1974), Johnson and Wehrly (1977), Mardia and Puri (1978) and Jupp and Mardia (1980) fall into this class. The second type consists of functions of the uncorrected covariance matrix. Some examples include those introduced by Watson and Beran (1967), Epp et al. (1971), Stephens (1979), Rivest (1982), Fisher and Lee (1983, 1986) and Jupp and Spurr (1985). The third type consists of statistics based on the empirical distribution function (EDF) and "axial EDF". Test statistics introduced by Rothman (1971), Puri and Rao (1977) and Fisher and Lee (1982) belong to this class. For further background on testing independence of bivariate circular data, see Jupp and Mardia (1989) and Shieh (1990). The test statistic that we introduce below falls into the third category. Its limit distribution is derived via the more general limit theory for "weighted degenerate U-statistics". Weighted degenerate U-statistics are an extension of U-statistics in the sense that they are the weighted average of degenerate kernels, while U-statistics are just averages of kernels. Let $Z_1, Z_2, \ldots$ be i.i.d. r.v.'s. Assume that $E(Z_1) = 0$, $\mathrm{Var}(Z_1) = 1$, and $EZ_1^4 < \infty$. Let $h(x,y)$ be a real valued function with finite second moment. A weighted U-statistic has the form

$$WU_n = \sum_{i \neq j} d_{ijn} h(Z_i, Z_j). \tag{1}$$

Here, $\{d_{ijn}\}$ are non-stochastic weights. The statistic WU_n is called a weighted degenerate U-statistic if the kernel h is degenerate in the sense that $\mathrm{Var}[h_1(Z_1)] = 0$, where $h_1(z_1) = Eh(z_1, Z_2)$.

Gregory (1977) and Serfling (1980) independently established the limit distribution of degenerate (equal-weighted) U-statistics. Weber (1981) established the limit distribution of incomplete degenerate U-statistics which is a special case of (1) with weights equal to 0 or $\frac{1}{n(n-1)}$. Janson (1984) extended Weber's (1981) result to incomplete U-statistics with weights which may be random. For weighted non-degenerate U-statistics, the asymptotic normality result is due to Shapiro and Hubert (1979).

2. Main Results

Two well-known rank statistics for testing independence of linear data are Kendall's tau (τ) and Spearman's rho (ρ), where

$$\tau = \frac{1}{n(n-1)} \sum_{i \neq j} \mathrm{sign}(\Theta_i - \Theta_j)\mathrm{sign}(\Phi_i - \Phi_j)$$

and

$$\rho = \frac{12}{n(n^2-1)} \sum_{i=1}^{n} \Big(R\Theta_i - \frac{n+1}{2}\Big)\Big(R\Phi_i - \frac{n+1}{2}\Big).$$

However, for directional data, values of τ and ρ vary as the choice of origin varies, an important disadvantage when testing independence of bivariate circular data. Our approach is to modify a weighted Kendall's tau, thus forming a class of rank statistics by averaging it over cyclic permutations of the marginal ranks. We define a (linear) weighted Kendall's tau statistic as

$$t_w = \sum_{i \neq j} c_{ijn} \mathrm{sign}(\Theta_i - \Theta_j) \mathrm{sign}(\Phi_i - \Phi_j), \tag{2}$$

where $c_{ijn} = c_{jin}$. The following notations are useful in the derivation of a class of statistics, that are invariant under the choice of origins and separate continuous one-to-one transformations of each circle onto itself. These transformations must preserve the clockwise say, orientation between any two points on a single circle (all points transversed will have their images transversed).
First define the lth cyclic permutation of a rank vector $\mathbf{R}$ as

$$\mathbf{C}_l(\mathbf{R}) = (C_l(R_1), \ldots, C_l(R_n)), \text{ where}$$
$$C_l(R_i) = R_i + l \quad (\text{modulo } n), \text{ for } 1 \leq i \leq n.$$

Second, let AC_Θ be the operation that averages a function of ranks over all cyclic permutations of the ranks of Θ_i's, i.e.,

$$AC_\Theta[f(\mathbf{R\Theta}, \mathbf{R\Phi})] = \frac{1}{n} \sum_{l=1}^{n} f[\mathbf{C}_l(\mathbf{R\Theta}), \mathbf{R\Phi}], \text{ where}$$
$$\mathbf{R\Theta} = (R\Theta_1, \ldots, R\Theta_n) \text{ and } \mathbf{R\Phi} = (R\Phi_1, \ldots, R\Phi_n).$$

Similarly, AC_Φ is defined as

$$AC_\Phi[f(\mathbf{R\Theta}, \mathbf{R\Phi})] = \frac{1}{n} \sum_{l=1}^{n} f[\mathbf{R\Theta}, \mathbf{C}_l(\mathbf{R\Phi})].$$

Let AC be the operation of averaging over both sets of cyclic permutations. By the definitions of AC_Θ and AC_Φ,

$$AC = AC_\Theta AC_\Phi = AC_\Phi AC_\Theta.$$

For any fixed choices of origins, separate continuous monotone 1-1 transformations of the coordinates lead to the sets of ranks as invariants. Because orientation preserving transformations can change the origins, the sets of permuted ranks are appropriate invariants. Distribution-free test statistics must be constant on these sets of permuted ranks. We note that shifting the origin counter-clockwise across one datum corresponds to a cyclic permutation of the ranks. Thus, to modify a weighted Kendall's tau into a class of test statistics for circular data, we average the evaluations of t_w with respect to all possible origins, i.e., apply AC to t_w. This is in the spirit of "a permutation test conditional on the marginals" (Jupp

(1987)). The resulting statistic is distribution free under the null hypothesis with any of the $n!$ permutations being equally likely for $\mathbf{R\Theta}$ and for $\mathbf{R\Phi}$.

2.1. A class of invariant statistics

Consider the weighted statistic t_w with weight matrix in the class C_n as specified in Equation (3) below. Define

$$\mathbf{C}_{mn} = (cc_{ij(m)})_{n\times n},$$

where

$$cc_{ij(m)} = \begin{cases} 1, & |i-j| = m \text{ or } n-m, \\ 0, & \text{otherwise}, \end{cases}$$

for $m = 1, 2, \ldots, [n/2]$. Note that $\mathbf{C}_{mn}$ contains only elements which are m or $(n-m)$ apart from the diagonal. Let C_n be the collection of matrices which are linear combinations of $\mathbf{C}_{mn}$, i.e.,

$$C_n = \Big\{\sum_{m=1}^{M} k_m \mathbf{C}_{mn} : k_m \text{ depends on } m \text{ and } n \text{ and is of order } n^{-2}\Big\}, \qquad (3)$$

for $M = 1, 2, \ldots, [n/2]$. We note that taking pairwise angles for t_w as in (2) implies that $\mathbf{C}_{mn}$ assumes a special form. This special form yields a group, $(C_n, +, 0)$, under addition. That is, for any two matrices in C_n, their linear combination is still in C_n. If we believe that closer pairs of observations would contribute more to the correlation than those further apart, then we may take $k_1 \geq \ldots \geq k_M$. However, if we believe that any pair of observations should contribute equally, then we may take $k_1 = \cdots = k_M = O(n^{-2})$, $M = [n/2]$. Thus C_n in (3) is a fairly rich family for weight matrices when defining a weighted Kendall's tau. The following notation is needed for Property 1 below. Recall that $\mathbf{X}_i = (\Theta_i, \Phi_i)$. For any real valued function f, $\Sigma_{i_3=i_4}$ denotes summation of f over all distinct indices in (i_1, i_2, i_3, i_4) except that i_3 may equal i_4, i.e.,

$$\sum_{i_3=i_4} f(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) = \sum_{p(3)} f(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_3}) + \sum_{p(4)} f(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}), \qquad (4)$$

where $\Sigma_{p(m)}$ sums over all distinct permutations of $\{i_1, \ldots, i_m\}$ from $\{1, \ldots, n\}$ and m is the number of distinct i_j's in f. Property 1 shows that an invariant statistic may be derived from t_w with any weight matrix in C_n.

Property 1. *For any weighted statistic t_w with weight matrix in the class C_n (defined in (3)), let T_w be the statistic resulting from a cyclic permutation of t_w,* *i.e.*, $T_w = AC(t_w)$. *For any constant* $1 \le M \le [n/2]$ *and for any integer* $n \ge 2$,

(i)
$$T_w = n^{-2} \sum_{m=1}^{M} k_m \sum_{i_3=i_4} cc_{i_1 i_2(m)} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}),$$
$$= n^{-2} \sum_{i_3=i_4} c_{i_1 i_2} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}),$$

where $(c_{i_1 i_2}) \in C_n$, $(cc_{i_1 i_2(m)}) \in \mathbf{C}_{mn}$,

$$h_w(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \text{csign}(\Theta_1, \Theta_2, \Theta_3)\text{csign}(\Phi_1, \Phi_2, \Phi_4) \tag{5}$$

and

$$\text{csign}(x, y, z) = [\text{sign}(x-y) + \text{sign}(y-z) + \text{sign}(z-x)]. \tag{6}$$

(ii) T_w *is invariant under different choices of origin and direction of rotation.*

The proof is in Appendix 1.

Averaging the "weighted Kendall's tau" over cyclic permutations of each coordinate is in the spirit of a permutation test conditional on the marginals. Property 1 (ii) states that a permutation test conditional on the marginals is invariant, as suggested in Jupp (1987). In the following, we present two interesting test statistics derived from a weighted Kendall's tau after a special choice of weights. The first one, presented in Corollary 1 below, is derived from Kendall's tau (i.e., t_w in (2) with equal weights $c_{ij} \equiv \frac{1}{n(n-1)}$).

Corollary 1. *Let* $T_e = AC(\tau)$. *Then*

$$T_e = \tau - \frac{2(n+1)}{3n}\rho, \tag{7}$$

where τ *and* ρ, *defined in Section* 2, *are Kendall's tau and Spearman's rho, respectively.*

Applying cyclic permutations to one coordinate of Kendall's tau, say the $R\Theta_i$'s, we obtain $\tau - \frac{2(n+1)}{3n}\rho$. Note that ρ, after being averaged over all cyclic permutations with respect to any coordinate, equals 0. Due to this special property of ρ, $AC_\Phi(T_e)$ yields T_e again. The invariance of T_e is easily seen from (7), since $AC(\rho) = 0$ and thus $AC(T_e) = AC(\tau) = T_e$.

The following is the second invariant test statistic derived from the weighted Kendall's tau with weights in (8) below.

$$n^{-2} \begin{bmatrix} 0 & b & 0 & \cdots & b \\ b & 0 & b & & 0 \\ 0 & b & 0 & & \vdots \\ \vdots & 0 & b & \ddots & 0 \\ 0 & \vdots & 0 & & b \\ b & 0 & \cdots & b & 0 \end{bmatrix}_{n \times n}, \tag{8}$$

where $0 < b$. Suppose we now consider that observations are taken adjacent in time. Instead of using equal weight in Kendall's tau, we shall put more weight on adjacent observations than on non-adjacent observations. We note that the constant b in (8) accommodates the above intuition and the fact that the nth and first rank sit adjacently, which in turn accommodates the nature of cyclic time series data. The index 'a' in T_a indicates that T_a is derived from t_w with autocorrelation weight matrix (8).

Corollary 2. *Let t_a be t_w with weights in* (8) *and define $T_a = AC(t_a)$. Define $\mathbf{X}_0 = \mathbf{X}_n$ and $\mathbf{X}_{n+1} = \mathbf{X}_1$. Then,*

$$T_a = \frac{2b}{n} \sum_{i_1=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_1-1}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}).$$

3. Limit Distributions for Weighted Degenerate U-statistics

The motivation for this section is to establish a limit distribution for the test statistic T_w with weights satisfying (3). It turns out that the limit distribution is an immediate application of the result for weighted degenerate U-statistics. To see this connection, define

$$\hat{T}_w = \Big(\frac{n-2}{n}\Big)^2 \sum_{i_1 \neq i_2} c_{i_1 i_2} \hat{h}_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}) \tag{9}$$

with

$$\hat{h}_w(\mathbf{X}_1, \mathbf{X}_2) = \Big(\mathrm{sign}(\Theta_1 - \Theta_2) - 2[F(\Theta_1) - F(\Theta_2)]\Big) \cdot \Big(\mathrm{sign}(\Phi_1 - \Phi_2) - 2[G(\Phi_1) - G(\Phi_2)]\Big).$$

The statistic $\hat{T}_w$ is the projection of T_w. In Appendix 2, we show that $E[n(T_w - \hat{T}_w)]^2 = O(n^{-1})$, thus

$$n(T_w - \hat{T}_w) \to_D 0. \tag{10}$$

In Appendix 2, we also show that the kernel, $\hat{h}_w$, is degenerate. Hence $\hat{T}_w$ is a weighted degenerate U-statistic as defined in (1). We conjecture but have been unable to prove that the operation, averaging t_w over the cyclic permutations of its marginal ranks, results in the degeneracy of T_w's, and hence the degeneracy of $\hat{T}_w$'s. The degeneracy of $\hat{h}_w$ implies that the limit distribution of $n\hat{T}_w$ be, instead of a normal, a linear combination of independent chi square r.v.'s. We explore its limit distribution below.

We now consider a general weighted degenerate U-statistic (WU_n). Similar to Serfling (1980), we use the theory of linear operators (Dunford and Schwartz (1963)) and Fourier series (Kufner and Kadlec (1971)) to expand h into an infinite weighted sum of eigenfunctions. This infinite series is asymptotically equivalent to its finite version (a quadratic form) in L_2, since the kernel is assumed to have finite second moments, i.e.,

$$E[h(Z_1, Z_2)]^2 < \infty. \tag{11}$$

Thus, h may be used to define an operator A in the functional space $L_2(\mathbf{R}, F)$, by

$$A\phi(z_1) = \int_{-\infty}^{\infty} \widetilde{h}(z_1, z_2)\phi(z_2)\, dF(z_2), \quad \text{for any } z_1 \in \mathbf{R}, \phi \in L_2.$$

Let ϕ_i's be the distinct eigenfunctions of A and α_i's the corresponding eigenvalues. Assuming that h satisfies (11), we may expand the centered kernel, $\widetilde{h} = h - Eh$, as a weighted sum of product of eigenfunctions, i.e.,

$$\widetilde{h}(Z_i, Z_j) = \sum_{k=1}^{\infty} \alpha_k \phi_k(Z_i)\phi_k(Z_j).$$

See Serfling (1980, p.196) or Dunford and Schwartz (1963) for further details. Define $Z_{ki} = \phi_k(Z_i)$. Thus we may write

$$WU_n = \sum_{i \neq j} d_{ijn} \sum_{k=1}^{\infty} \alpha_k Z_{ki} Z_{kj}.$$

The following notations are used in the statement of Theorem 1 below. Let $\mathbf{B}_n = (b_{imn})$ and $\mathbf{D}_n = (d_{ijn})$, where $b_{imn} \in R$, for $i = 1, \ldots, n$; $m, n = 1, 2, \ldots$. Here, $\mathbf{B}_n$ is an orthogonal matrix such that $\mathbf{B}'_n \mathbf{D}_n \mathbf{B}_n = \Lambda_n$, where Λ_n is a diagonal matrix with λ_{mn} as the mth diagonal element. Assume $\lim_{n\to\infty} \lambda_{mn} = \lambda_m$ and use the notation $\delta_{km} = 1$, if $k = m$, and $\delta_{km} = 0$ otherwise.

Theorem 1. *Assume that $E[h(Z_1, Z_2)]^2 < \infty$ and the following conditions hold:*
(i) $\max_{1\le i\le n} |b_{imn}| \to 0$ *as* $n \to \infty$ *for each* m,
(ii) $\sum_{i=1}^n b_{imn} b_{ikn} \to \delta_{mk}$ *as* $n \to \infty$ *for all* m, k,
(iii) $\sum_{i=1}^n \sum_{j=1}^n d_{ijn}^2 \to \sum_{m=1}^{\infty} \lambda_m^2 < \infty$,
(iv) $\sum_{i=1}^n \sum_{j=1}^n d_{ijn} b_{imn} b_{jmn} \to \lambda_m$ *as* $n \to \infty$, *for all* m.
Then

$$WU_n \to_D G_0 = \Big(\sum_{k=1}^{\infty} \alpha_k \Big[\sum_{m=1}^{\infty} \lambda_m (Y_{km}^2 - 1)\Big]\Big),$$

where the Y_{km}'s are independent $N(0,1)$ variates. The proof is in Appendix 3.

By Theorem 1 of Verrill and Johnson (1988), (ii)-(iv) are conditions for weights of quadratic forms such that $\mathbf{D}_n - \{\lambda_m\} - \mathbf{B}_n$ have an approximate symmetric matrix-eigenvalues-eigenvectors relationship. Condition (i) is the central limit theorem negligibility condition that ensures the asymptotic normality of $\mathbf{B}_n'\mathbf{Z}_{kn}$, where $\mathbf{Z}_{kn}' = [Z_{k1}, \ldots, Z_{kn}]$, for each fixed k.

Since the kernel of $\hat{T}_w$ is degenerate and the weights satisfy the conditions in Theorem 1, the limit distribution of $n\hat{T}_w$ is immediate from Theorem 1. Hence,

$$n\hat{T}_w \to_D \sum_{k=1}^{\infty} \alpha_k \Big[\sum_{m=1}^{\infty} \lambda_m (Y_{km}^2 - 1) \Big]. \tag{12}$$

Equations (10) and (12) yield the limit distribution of nT_w stated in Corollary 3 below.

Corollary 3. *Let $(c_{i_1 i_2}) \in C_n$. Then, the limit distribution of T_w is given by*

$$nT_w \to_D \sum_{k=1}^{\infty} \alpha_k \Big[\sum_{m=1}^{\infty} \lambda_m (Y_{km}^2 - 1) \Big], \tag{13}$$

where α_k, λ_m are the kth *eigenvalue and the* mth *limiting eigenvalue of the kernel and the weight matrix, respectively.*

4. Circular Kendall's Tau (T_n)

We have obtained limit distributions of T_w with general weights in C_n. In this section, asymptotic and finite sample properties of Circular Kendall's tau (T_n) are investigated, where $T_n = \frac{3n}{n-2} T_e$ and T_e is derived from applying cyclic permutations to each set of ranks for Kendall's tau. We apply Theorem 1 to obtain the limit distribution of T_n. In Property 2 (i) below, using an expression of U-statistics, it is shown that T_n is equal to the test statistic ($\hat{\Delta}_n$) in Fisher and Lee (1982).

Property 2. *Let $T_n = \frac{3n}{n-2} T_e$. Then*

(i) $T_n = \hat{\Delta}_n = \binom{n}{3}^{-1} \sum_{i<j<k} \delta(\mathbf{X}_i, \mathbf{X}_j, \mathbf{X}_k)$,

where $\delta(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = \mathrm{sign}(\Theta_1 - \Theta_2)\mathrm{sign}(\Theta_2 - \Theta_3)\mathrm{sign}(\Theta_3 - \Theta_1) \times \mathrm{sign}(\Phi_1 - \Phi_2)\mathrm{sign}(\Phi_2 - \Phi_3)\mathrm{sign}(\Phi_3 - \Phi_1)$.

(ii) $-1 \le T_n \le 1$.

(iii) $T_n = 1$, *if both directions are of identical order, i.e.,* $R\Theta_i = R\Phi_i$, *for* $1 \le i \le n$.

(iv) $T_n = -1$, *if both directions are of reverse order, i.e.,* $R\Theta_i = R\Phi_{n+1-i}$, *for* $1 \le i \le n$.

(v) *The limit distribution of nT_n is given by*

$$nT_n \to_D V = \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \frac{3}{km\pi^2} (\xi_{km}^2 + \eta_{km}^2 - \zeta_{km}^2 - \omega_{km}^2), \text{ where } \xi_{km}, \eta_{km}, \zeta_{km}$$

and ω_{km} *are i.i.d.* $N(0,1)$ *variates.*

The proof of statement (i) is in Appendix 4. The proof of statements (ii)-(v) can be found in Fisher and Lee (1982) and Shieh (1990).

From Property 2 (v), the limit distribution of nT_n is the sum of independent variates with zero means and variances $\frac{9\times 8}{\pi^4 k^2 m^2}$. By straight-forward calculation, the asymptotic variance of T_n is equal to 2. For tables of critical values of T_n, see Fisher and Lee (1982).

Since $\hat{\Delta}_n = T_n$, from (7) and the definition of T_n, we have

$$\hat{\Delta}_n = \frac{1}{n-2}\{3n\tau - 2(n+1)\rho\},$$

i.e., $\hat{\Delta}_n$ is a linear combination of Kendall's tau and Spearman's rho. Note that for T_w derived from t_w with non-equal weights, those T_w can not be expressed in terms of $\hat{\Delta}_n$. This is easily seen from the following case. Take $n = 4$,

$$c_{ij(m)} = \begin{cases} 1, & |i-j| = 2, \\ 0, & \text{otherwise.} \end{cases}$$

4.1. Power study — comparison of T_n to Hillman's (1974) τ_{aa}

In Hillman (1974), a test statistic with a form related to T_n was proposed, namely

$$\tau_{aa} = \max_{l,m} \frac{1}{n(n-1)} \sum_{i\neq j} \text{sign}(\Theta_{i+l} - \Theta_{j+l})\text{sign}(\Phi_{i+m} - \Phi_{j+m}).$$

In this section, the power of the statistics T_n and τ_{aa} are compared via a Monte Carlo study for sample size $n = 10$. The model of dependence considered is

$$\Phi = \Theta + \text{von Mises } (0, 2\pi; k),$$

where Θ is uniform $(0, 2\pi)$, the probability density function g of the von Mises $(0, 2\pi; k)$ distribution is

$$g(\eta; \mu_0, k) = \frac{1}{2\pi I_0(k)} e^{k\cos(\eta-\mu_0)}, \quad 0 \le \eta \le 2\pi,\ k > 0,\ 0 \le \mu_0 < 2\pi,$$

$I_0(k) = \sum_{r=0}^{\infty} \frac{1}{r!^2}(\frac{k}{2})^{2r}$ is the modified Bessel function of the first kind and order zero, and $k = 2, 5$ and 10. Note that larger k indicates higher degree of dependence.

For $n = 10$, $k = 0$, 4000 values of τ_{aa} and T_n were generated to obtain 5% critical values. Under the one-sided alternative hypothesis that there exists positive correlation between Θ and Φ, for $n = 10$, T_n exhibits better power (0.58, 0.91 and 0.98 for $k = 2, 5$ and 10, respectively) than τ_{aa} (0.57, 0.80 and 0.87). We note that the Hillman's statistic is not invariant under the alternative hypothesis. Its power varies when the mode of the von Mises distribution varies. The powers of τ_{aa} shown above are the highest we have obtained thus far.

4.2. Example

In the following, we apply T_n to estimate angular-angular association in one set of isotropic data.

Example 1 (Fisher and Lee (1986)). Magnetic remanence at 680°C and 685°C in each of 52 rock specimens was measured. The estimated association between the 680°C and 685°C is $T_n = 0.0120$, and it is significant at the 5% level. An approximate 95% confidence interval for T_n is $(0.0105, 0.0135)$.

Readers interested in further applications may refer to a recent review paper on circular correlation by Hanson et al. (1992).

5. Concluding Remarks

We have derived a new class of statistics for testing independence with bivariate circular data. These statistics are derived via averaging over cyclic permutations of the weighted Kendall's tau in (2). These statistics are distribution-free. Further, they have the desirable property of being invariant under different choices of origin and direction of rotation. Among these statistics, we explore the asymptotic and finite sample properties of a Circular Kendall's tau. Its limit distribution is obtained via weighted degenerate U-statistics. Our asymptotic result concerning weighted degenerate U-statistics extends the theory of degenerate U-statistics.

Appendix: Proofs

Appendix 1: Proof of Property 1

To prove Property 1 (i), we first work with T_a, a special case of t_w in (2), which is derived from t_w with autocorrelation weights in (8). i.e., take $M = m = 1$ and $k_1 = \frac{b}{n^2}$. Applying result in the derivation of T_a in Shieh (1990), we have

$$T_a = \frac{2b}{n^2}\sum_{i=1}^{n}\Big\{\Big[\operatorname{sign}(R\Theta_i - R\Theta_{i+1}) - \frac{2(R\Theta_i - R\Theta_{i+1})}{n}\Big] \times\Big[\operatorname{sign}(R\Phi_i - R\Phi_{i+1}) - \frac{2(R\Phi_i - R\Phi_{i+1})}{n}\Big]\Big\}. \tag{A.1.1}$$

By (A.210) of Lehmann (1975),

$$R\Theta_i = \frac{1}{2}\sum_{k=1}^{n}[\text{sign}(\Theta_i - \Theta_k)] + \frac{n+1}{2}.$$

Thus

$$[\text{sign}(\Theta_i - \Theta_{i+1}) - \frac{2}{n}(R\Theta_i - R\Theta_{i+1})] = n^{-1}\sum_{k=1}^{n}\text{csign}(\Theta_i, \Theta_{i+1}, \Theta_k), \qquad \text{(A.1.2)}$$

where $\text{csign}(\Theta_i, \Theta_{i+1}, \Theta_k)$ is defined in (6). Putting this in (A.1.1), we have

$$T_a = \frac{2b}{n^4}\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\text{csign}(\Theta_i, \Theta_{i+1}, \Theta_k)\text{csign}(\Phi_i, \Phi_{i+1}, \Phi_l).$$

Define $\mathbf{X}_0 = \mathbf{X}_n$ and $\mathbf{X}_{n+1} = \mathbf{X}_1$. Now, for fixed k,

$$\sum_{i=1}^{n}\text{csign}(\Theta_i, \Theta_{i+1}, \Theta_k) = -\sum_{i=1}^{n}\text{csign}(\Theta_{i+1}, \Theta_i, \Theta_k) = -\sum_{i=1}^{n}\text{csign}(\Theta_i, \Theta_{i-1}, \Theta_k).$$

Likewise, the above equality holds for Φ_i's, for $1 \le i \le n$. Thus

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\text{csign}(\Theta_i, \Theta_{i+1}, \Theta_k)\text{csign}(\Phi_i, \Phi_{i+1}, \Phi_l)\\ =&\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\text{csign}(\Theta_i, \Theta_{i-1}, \Theta_k)\text{csign}(\Phi_i, \Phi_{i-1}, \Phi_l).\end{aligned}$$

So

$$T_a = n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}c_{ij}\text{csign}(\Theta_i, \Theta_j, \Theta_k)\text{csign}(\Phi_i, \Phi_j, \Phi_l),$$

where $c_{ij} = \frac{b}{n^2}$ for $|i-j| = 1$ or $n-1$, and $c_{ij} = 0$, otherwise. Substitute $h_w(\mathbf{X}_i, \mathbf{X}_j, \mathbf{X}_k, \mathbf{X}_l)$ in (5) for $\text{csign}(\Theta_i, \Theta_j, \Theta_k)\text{csign}(\Phi_i, \Phi_j, \Phi_l)$ and replace indices i, j, k, l by i_1, i_2, i_3, i_4. Further, note that $h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) = 0$, for any two identical $\mathbf{X}_{i_j}$'s except $\mathbf{X}_{i_3} = \mathbf{X}_{i_4}$. Thus

$$T_a = n^{-2}\sum_{i_3=i_4} c_{i_1 i_2} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}), \qquad \text{(A.1.3)}$$

where $\Sigma_{i_3=i_4}$ is defined in (4).

Note that the above result may be generalized to

$$\Big[\sum_{i=1}^{n}\text{sign}(\Theta_i - \Theta_{i+m})\text{sign}(\Phi_i - \Phi_{i+m})\Big],$$

for any $1 \le m \le M$. Thus by (A.1.3),

$$\begin{aligned}T_w = AC(t_w) =& n^{-2}\sum_{m=1}^{M} k_m \sum_{i_3=i_4} cc_{i_1i_2(m)} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4})\\ =& n^{-2}\sum_{i_3=i_4} c_{i_1i_2} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}),\end{aligned}$$

where $(cc_{i_1i_2(m)}) = 1$, if $|i_1 - i_2| = m$ or $n - m$ and $cc_{i_1i_2(m)} = 0$, otherwise, and $c_{i_1i_2} \in C_n$ as specified in (3).

In the following, we prove Property 1 (ii): T_w is invariant. By (A.1.3) and the LHS of (A.1.2), T_w can be expressed as

$$\begin{aligned}T_w = n^{-2}\sum_{i_3=i_4} c_{i_1i_2} &\Big[\text{sign}(R\Theta_{i1} - R\Theta_{i2}) - 2n^{-1}(R\Theta_{i1} - R\Theta_{i2})\big]\\ &\times\Big[\text{sign}(R\Phi_{i1} - R\Phi_{i2}) - 2n^{-1}(R\Phi_{i1} - R\Phi_{i2})\big].\end{aligned}$$

Since

$$AC_\Theta[\text{sign}(R\Theta_i - R\Theta_j)] = \sum_{i=1}^{n}[\text{sign}(R\Theta_i - R\Theta_j) - 2n^{-1}(R\Theta_i - R\Theta_j)]$$

and

$$AC_\Theta[R\Theta_i - R\Theta_j] = \frac{-2}{n}\Big[\sum_{l=1}^{n} C_l(R\Theta_i) - \sum_{l=1}^{n} C_l(R\Theta_j)\Big],$$

$AC_\Theta(T_w) = T_w$. Similarly, $AC_\Phi(T_w) = T_w$. Hence $AC(T_w) = T_w$. This is sufficient for Property 1 (ii).

Appendix 2: Proof of $E[n(T_w - \hat{T}_w)]^2 = O(n^{-1})$

We first show that $\hat{T}_w$ in (9) is the projection of T_w into the family of $\{\mathbf{X}_{i_1}, \mathbf{X}_{i_2}\}$, i.e.,

$$\hat{T}_w = \sum_{i_3=i_4} c_{i_1i_2} E[T_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \mid \mathbf{X}_{i_1}, \mathbf{X}_{i_2}],$$

where $c_{ij} = \Sigma_{m=1}^{M} k_m cc_{ij(m)}$ and $cc_{ij(m)} = 1$, if $|i-j| = m$ or $n-m$, and $cc_{ij(m)} = 0$, otherwise. After straightforward algebra, we have

$$\begin{aligned}&E[h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \mid \mathbf{X}_1 = (\theta_1, \phi_1), \mathbf{X}_2 = (\theta_2, \phi_2)]\\ &= \begin{cases} f_1(\theta_1, \theta_2) f_2(\phi_1, \phi_2), & (i_1, i_2) = (1,2) \text{ or } (2,1),\\ 0, & \text{otherwise,}\end{cases}\end{aligned}$$

where

$$\begin{cases} f_1(\theta_1, \theta_2) = \text{sign}(\theta_1 - \theta_2) - 2[F(\theta_1) - F(\theta_2)],\\ f_2(\phi_1, \phi_2) = \text{sign}(\phi_1 - \phi_2) - 2[G(\phi_1) - G(\phi_2)].\end{cases}$$

Let $\hat{h}_w(\mathbf{X}_1, \mathbf{X}_2) = f_1(\theta_1, \theta_2) f_2(\phi_1, \phi_2)$. Thus,

$$\sum_{i_3=i_4} E[h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \mid \mathbf{X}_1, \mathbf{X}_2] = 2(n-2)^2 \hat{h}_w(\mathbf{X}_1, \mathbf{X}_2),$$

since fixing i_1 and i_2, there are $(n-2)^2$ terms from summing over i_3 and i_4. This yields (9). Note that

$$\begin{aligned} &E[f_1(\Theta_1, \Phi_1) f_2(\Theta_2, \Phi_2) \mid \mathbf{X}_1 = (\theta_1, \phi_1)] \\ &= E_{\mathbf{X}_2}\Big[\{\mathrm{sign}(\theta_1 - \Theta_2) - 2[F(\theta_1) - F(\Theta_2)]\}\{\mathrm{sign}(\phi_1 - \Phi_2) - 2[G(\phi_1) - G(\Phi_2)]\}\Big] \\ &= \{2F(\theta_1) - 1 - 2[F(\theta_1) - 1/2]\}\{2G(\phi_1) - 1 - 2[G(\phi_1) - 1/2]\} \equiv 0. \end{aligned}$$

Thus, $\hat{h}_w$ is degenerate.

Alternatively, $\hat{T}_w$ may be expressed as

$$\hat{T}_w = n^{-2} \sum_{i_3=i_4} c_{i_1 i_2} \hat{h}_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}).$$

Let $H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) = h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) - \hat{h}_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2})$.

$$\begin{aligned} E[T_w - \hat{T}_w]^2 &= n^{-4} E\Big[\sum_{i_3=i_4} c_{i_1 i_2}(h_w - \hat{h}_w)\Big]^2 \qquad \text{(A.2.1)} \\ &= n^{-4} \sum_{i_3=i_4} \Big\{ c_{i_1 i_2}^2 E[H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4})]^2 \\ &\quad + \sum_{j_3=j_4} 2 c_{i_1 i_2} c_{j_1 j_2} E[H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) H(\mathbf{X}_{j_1}, \mathbf{X}_{j_2}, \mathbf{X}_{j_3}, \mathbf{X}_{j_4})]\Big\}. \end{aligned}$$

Let $\mathbf{i} = (i_1, i_2, i_3, i_4)$ and $\mathbf{j} = (j_1, j_2, j_3, j_4)$. Note that for $0, 1$ and 2 common indices in $\mathbf{i}$ and $\mathbf{j}$,

$$\begin{aligned} &EH(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) EH(\mathbf{X}_{j_1}, \mathbf{X}_{j_2}, \mathbf{X}_{j_3}, \mathbf{X}_{j_4}) = 0, \\ &EE[H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) H(\mathbf{X}_{i_1}, \mathbf{X}_{j_2}, \mathbf{X}_{j_3}, \mathbf{X}_{j_4}) \mid \mathbf{X}_{i_1}] = 0 \end{aligned}$$

and

$$EE[H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{j_3}, \mathbf{X}_{j_4}) \mid \mathbf{X}_{i_1}, \mathbf{X}_{i_2}] = 0.$$

Accordingly, the second term in the RHS of (A.2.1) vanishes except when there are 3 or 4 indices in common between $\mathbf{i}$ and $\mathbf{j}$.

Thus

$$\begin{aligned} E[T_w - \hat{T}_w]^2 &= n^{-4}\Big\{ \sum_{i_3=i_4} \Big[c_{i_1 i_2}^2 EH^2(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \\ &\quad + \sum_{j_4=1}^{n} 12 c_{i_1 i_2}^2 EH(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{j_4}) \\ &\quad + \sum_{j_2=1}^{n} 12 c_{i_1 i_2} c_{i_1 j_2} EH(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) H(\mathbf{X}_{i_1}, \mathbf{X}_{j_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \Big] \Big\}. \end{aligned}$$

Note that

$$|H(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4})| \le K < \infty,$$

and for $|i_1 - i_2| = m$ or $n - m$,

$$c_{i_1 i_2} = k_m = C_m n^{-2}, \text{where } C_m\text{'s are constants for each fixed } m.$$

Let $C_0 = \max_{1 \le m \le [n/2]} C_m$.

$$\begin{aligned} E[T_w - \hat{T}_w]^2 &\le \Big[\frac{KC_0}{n^2}\Big]^2 \Big\{n^2(n-1)(n-2)n^{-4} + 24n^2(n-1)(n-2)(n-3)n^{-4}\Big\} \\ &= O(n^{-3}). \end{aligned}$$

Appendix 3: Proof of Theorem 1

First, we wish to show that for fixed K

$$\sup_n E(WU_n - WU_{nK})^2 \le C_K \tag{A.3.1}$$

and

$$WU_{nK} \to_D G_K = \sum_{k=1}^{K} \alpha_k \Big[\sum_{m=1}^{\infty} \lambda_m (Y_{km}^2 - 1)\Big], \text{ as } n \to \infty, \tag{A.3.2}$$

where C_K is a constant depending on K and $C_K \to 0$ as $K \to \infty$, is true. Next, we note that as $K \to \infty$,

$$G_K \to_D G_0, \text{where } G_0 = G_\infty. \tag{A.3.3}$$

We then argue that Theorem 1 follows immediately from (A.3.1)-(A.3.3).

To show (A.3.1), we need (A.3.4) and (A.3.5) below. For each fixed k, Z_{ki}'s are independent r.v.'s, by independence of Z_i's and Theorem 3.3.1 of Chung (1974). Similarly, Z_{ki}^2's are independent. Furthermore, by the properties of orthonormal eigenfunctions $\{\phi(\cdot)\}$, $E\{\phi_k(Z)\phi_m(Z)\} = \delta_{km}$. Thus for $i \ne j$, we have

$$E\{Z_{ki}Z_{kj}\} = EZ_{ki}EZ_{kj} = 0. \tag{A.3.4}$$

$$E\{Z_{ki}^2 Z_{kj}^2\} = EZ_{ki}^2 EZ_{kj}^2 = 1. \tag{A.3.5}$$

If d_{ijn}'s satisfy Condition (iii), then by straightforward calculation, and by (A.3.4) and (A.3.5), we have

$$
\begin{aligned}
&E(WU_n - WU_{nK})^2 \\
&= 4E\Big\{\sum_{i<j} d_{ijn}^2\Big[\sum_{k=K+1}^{\infty} \alpha_k Z_{ki} Z_{kj}\Big]^2 \\
&\quad + 2\sum_{i<j<l<m} d_{ijn} d_{lmn}\Big[\sum_{k=K+1}^{\infty} \alpha_k Z_{ki} Z_{kj}\Big]\Big[\sum_{k=K+1}^{\infty} \alpha_k Z_{kl} Z_{km}\Big]\Big\} \\
&= 4\Big(\sum_{i<j} d_{ijn}^2\Big)\Big(\sum_{k=K+1}^{\infty} \alpha_k^2\Big) \\
&\le C_1\Big(\sum_{k=K+1}^{\infty} \alpha_k^2\Big), \text{ (uniform in } n \text{ by Condition (iii))}
\end{aligned}
$$

where C_1 is a constant. Let $C_K = C_1 \sum_{k=K+1}^{\infty} \alpha_k^2$, then (A.3.1) holds.

Next, we show (A.3.2) and (A.3.3) hold. For fixed K and n, by change of summation we can rewrite WU_{nK} as

$$WU_{nK} = \sum_{k=1}^{K} \alpha_k \sum_{i \neq j} d_{ijn} Z_{ki} Z_{kj}.$$

Let Y_{km}'s be N(0, 1) variates, for $1 \le k,\ m \le n$. For each fixed k, d_{ijn} and Z_{ki}'s satisfy the conditions of Theorem 1 of Verrill and Johnson (1988). Thus

$$\sum_{i \neq j} d_{ijn} Z_{ki} Z_{kj} \to_D \sum_{m=1}^{\infty} \lambda_m (Y_{km}^2 - 1),$$

which yields (A.3.2). Since $\sum_{k=1}^{\infty} \alpha_k < \infty$, (A.3.3) holds.

Now we are ready to show $WU_n \to_D G_0$. Note that

$$| WU_n - G_0 | \le | WU_n - WU_{nK} | + | WU_{nK} - G_K | + | G_K - G_0 | .$$

For any fixed K, letting n tend to infinity, by (A.3.1) and (A.3.2) we have $E(WU_n - WU_{nK})^2 \le C_K$ and $WU_{nK} - G_K \to_D 0$. Then letting K tend to infinity, we have $G_K - G_0 \to 0$. Further, since $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$,

$$C_K = C_1 \sum_{k=K+1}^{\infty} \alpha_k^2 \to 0 \quad \text{as } K \to \infty, \text{ for all } n.$$

Thus $WU_n \to_D G_0$.

Appendix 4: Proof of Property 2(i)

To prove statement (i), we first express $\hat{\Delta}_n$ in terms of h_w, namely

$$\hat{\Delta}_n = \binom{n}{3}^{-1} \frac{1}{3!} \sum_{p(3)} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_3}),$$

where h_w is defined in (5). From straightforward algebra, we have

$$T_n = \frac{3(n-3)}{n} U_w + \frac{3}{n}\hat{\Delta}_n,$$

where $U_w = \binom{n}{4}^{-1} \frac{1}{4!} \sum_{p(4)} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4})$. Once we show that $U_w = \frac{1}{3}\hat{\Delta}_n$, (i) holds. For $n = 4$, by (5) we have

$$\begin{aligned}
U_w =& \\
&\frac{2}{4!}\Big\{\text{csign}(\Theta_1,\Theta_2,\Theta_3)[\text{csign}(\Phi_1,\Phi_2,\Phi_4)-\text{csign}(\Phi_1,\Phi_3,\Phi_4)+\text{csign}(\Phi_2,\Phi_3,\Phi_4)] \\
&+ \text{csign}(\Theta_1,\Theta_2,\Theta_4)[\text{csign}(\Phi_1,\Phi_2,\Phi_3)+\text{csign}(\Phi_1,\Phi_3,\Phi_4)-\text{csign}(\Phi_2,\Phi_3,\Phi_4)] \\
&+ \text{csign}(\Theta_1,\Theta_3,\Theta_4)[-\text{csign}(\Phi_1,\Phi_2,\Phi_3)+\text{csign}(\Phi_1,\Phi_2,\Phi_4)+\text{csign}(\Phi_2,\Phi_3,\Phi_4)] \\
&+ \text{csign}(\Theta_2,\Theta_3,\Theta_4)[\text{csign}(\Phi_1,\Phi_2,\Phi_3)-\text{csign}(\Phi_1,\Phi_2,\Phi_4)+\text{csign}(\Phi_1,\Phi_3,\Phi_4)]\Big\} \\
&= \frac{1}{3}\binom{4}{3}^{-1} \sum_{i<j<k} \text{csign}(\Theta_i,\Theta_j,\Theta_k)\text{csign}(\Phi_i,\Phi_j,\Phi_k) = \frac{1}{3}\hat{\Delta}_n,
\end{aligned}$$

since $\text{csign}(x,y,z) = -\text{sign}(x-y)\text{sign}(y-z)\text{sign}(z-x)$. Thus (i) holds for $n = 4$.

Let h_{FL} be the kernel of $\hat{\Delta}_n$. Similarly, for $n \geq 5$, the sum of $\binom{n}{4} \times 2 \times 4$ terms of $\text{csign}(\Theta_i,\Theta_j,\Theta_k)\text{csign}(\Phi_i,\Phi_j,\Phi_k)$'s can be shown equal to $2 \times (n-3)\sum_{i<j<k} h_{FL}$. Thus

$$\begin{aligned}
U_w &= \binom{n}{4}^{-1} \frac{1}{4!} \sum_{p(4)} h_w(\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \mathbf{X}_{i_3}, \mathbf{X}_{i_4}) \\
&= \frac{2(n-3)}{n(n-1)(n-2)(n-3)} \sum_{i<j<k} h_{FL} = \frac{1}{3}\hat{\Delta}_n.
\end{aligned}$$

Acknowledgement

We wish to thank Steve Stigler for sharing his ideas concerning the invariance property of the Circular Kendall's tau in earlier discussions with Richard Johnson. G. Shieh is grateful to Brian Yandell, Nick Fisher and Paul Speckman for many helpful suggestions and to Tom Bennet for assistance in computation. We thank two referees and the editor for helpful suggestions which improved the paper. This research was supported by the NSF Grant SES-9022200 (for Frees) and Summer Research Fellowships from University of Missouri-Columbia (for Shieh).

References

Chung, K. L. (1974). *A Course in Probability Theory.* Academic Press, New York.

Downs, T. D. (1974). *Rotational Angular Correlation.* In *Biorhythms and Human Reproduction* (Edited by Ferin et al.), 97-104, John Wiley, New York.

Dunford, N. and Schwartz, J. T. (1963). *Linear Operators.* John Wiley, New York.

Epp, R. J., Tukey, J. W. and Watson, G. S. (1971). Testing unit vectors correlation. *J. Geophys. Res.* **76**, 8480-8483.

Fisher, N. I. and Lee, A. J. (1982). Nonparametric measures of angular-angular association. *Biometrika* **69**, 315-321.

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

Fisher, N. I. and Lee, A. J. (1986). Correlation coefficients for random variables on a unit sphere or hypersphere. *Biometrika* **73**, 159-164.

Gregory, G. G. (1977). Large sample theory for U-statistics and tests of fit. *Ann. Statist.* **5**, 110-123.

Hanson, B., Klink, K., Matsuura, K., Robeson, S. and Willmott, C. (1992). Vector correlation: Review, exposition, and geographic application. *Ann. Assoc. Amer. Geographers* **82**, 103-116.

Hillman, D. C. (1974). Correlation coefficients for ranked angular variates. In *Chronobiology* (Edited by L. E. Scheving, F. Halberg & J. E. Pauly), 723-730, Igaka Shoin, Tokyo.

Janson, S. (1984). The asymptotic distributions of incomplete U-statistics. *Z. Wahrschein. verw. Gebiete* **66**, 495-505.

Johnson, R. A. and Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *J. Roy. Statist. Soc. Ser. B* **39**, 222-229.

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

Jupp, P. E. and Spurr, B. D. (1985). Sobolev tests for independence of directions. *Ann. Statist.* **13**, 1140-1155.

Jupp, P. E. (1987). A nonparametric correlation coefficient and a two-sample test for random vectors or directions. *Biometrika* **74**, 887-890.

Jupp, P. E. and Mardia, K. V. (1989). A unified view of the theory of directional statistics, 1975-1988. *Internat. Statist. Rev.* **57**, 261-294.

Kendall, M. G. (1955). *Rank Correlation Methods.* Charles Griffin, London.

Kufner, A. and Kadlec, J. (1971). *Fourier Series.* Academia, Prague.

Lehmann, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks.* Holden Day, Oakland.

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

Puri, M. L. and Rao, J. S. (1977). Problems of association for bivariate circular data and a new test of independence. *Multivariate Analysis* IV, 513-522, North Holland.

Rivest, L.-P. (1982). Some statistical methods for bivariate circular data. *J. Roy. Statist. Soc. Ser. B* **44**, 81-90.

Rothman, E. D. (1971). Tests of coordinate independence for bivariate sample on a torus. *Ann. Math. Statist.* **42**, 1962-1969.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* John Wiley, New York.

Shapiro, C. P. and Hubert, L. (1979). Asymptotic normality of permutation statistics derived from weighted sums of bivariate functions. *Ann. Statist.* **7**, 788-794.

Shieh, S. R. (1990). Some extensions of U- and V-statistics. Unpublished Ph.D. dissertation, University of Wisconsin, Madison, WI.

Stephens, M. A. (1979). Vector correlation. *Biometrika* **66**, 41-48.

Verrill, S. and Johnson, R. A. (1988). Asymptotic distributions for quadratic forms with applications to censored data tests of fit. *Comm. in Statist.* **17**, 4011-4024.

Watson, G. S. and Beran, R. J. (1967). Testing a sequence of unit vectors for serial correlation. *J. Geophy. Res.* **72**, 5655-5659.

Weber, N. C. (1981). Incomplete degenerate U-statistics. *Scand. J. Statist.* **8**, 120-123.

Department of Statistics, University of Missouri, Columbia, MO 65211, U.S.A.
Department of Statistics, University of Wisconsin, Madison, WI 53706, U.S.A.

(Received August 1991; accepted December 1993)