SPEARMAN’S RHO AND KENDALL’S TAU FOR MULTIVARIATE DATA SETS

HAMANI EL MAACHE AND YVES LEPAGE
Université de Montréal

A class of U-statistics matrices is introduced to obtain the distribution of the matrices of the Spearman and Kendall correlation coefficients between the components of a random vector. These results are used to construct nonparametric tests of independence between two sets of variables based on three measures of multivariate relationship. The tests are illustrated by an example and a simulation study is performed to compare the tests based on Kendall’s matrix with those based on Spearman’s matrix.

1. Introduction

Let \( F(x) = F(x[1], x[2]) \) be the continuous c.d.f. (cumulative distribution function) of a random vector \( X = (X[1], X[2])', \) where \( x = (x(1), \ldots, x(m))' \in \mathbb{R}^m, \) \( m \geq 2, \) \( x[1] \in \mathbb{R}^p, \) \( x[2] \in \mathbb{R}^q \) \((p + q = m)\) and \( F[k](x[k]) \) \((k = 1, 2)\) denote the marginal c.d.f. of \( X[k]. \) The objective of this paper is to detect deviation from the null hypothesis of independence that is, to test \( H_0: F(x) = F(\alpha)F(\beta) \) against appropriate classes of alternatives \( H_{1,n}. \) A nonparametric approach to this problem was explored by Puri, Sen and Gokhale (1970) who defined a class of association parameters based on componentwise ranking. The statistic they proposed uses the elements of the matrix \( D_n = (D_{11}, D_{12}), \) where

\[
D_n^{(i,j)} = \frac{1}{n} \sum_{\alpha=1}^{n} J\left( \frac{R_{\alpha}^{(i)}}{n} \right) J\left( \frac{R_{\alpha}^{(j)}}{n} \right), \quad i, j = 1, \ldots, m.
\]

Here, \( R_{\alpha}^{(i)} \) is the rank of \( X_{\alpha}^{(i)}, \) that denote the \( i \)th coordinate of the vector \( X_{\alpha}; \) the symbol \( \alpha \) will run over the sample (from \( X \)) with \( \alpha = 1, \ldots, n \) and \( J \) represents an arbitrary standardized score function. Puri, Sen and Gokhale (1970) established the joint asymptotic multivariate normality of the vector formed by the elements of \( D_n. \)

When the score function is \( J(u) = J_0(u) = \sqrt{12}(u - \frac{1}{2}), \) then

\[
D_n^{(i,j)} = \frac{12}{n(n^2 - 1)} \sum_{\alpha=1}^{n} \left( R_{\alpha}^{(i)} - \frac{n + 1}{2} \right) \left( R_{\alpha}^{(j)} - \frac{n + 1}{2} \right), \quad i, j = 1, \ldots, m,
\]

AMS subject classifications: Primary 62H20; Secondary 62H15.

Keywords and phrases: U-statistics, Spearman rho correlation coefficient, Kendall tau correlation coefficient, measures of multivariate association, rank correlation, tests of independence.
which reduces to Spearman’s rank correlation with asymptotic mean given by Spearman’s coefficient (see Hoeffding, 1948, p. 318)

\[(1.3) \quad \rho_{ij} = 3 \int \left[2F^{(i)}(x^{(i)}) - 1\right]\left[2F^{(j)}(x^{(j)}) - 1\right] dF^{(i,j)}(x^{(i)}, x^{(j)}), \quad i, j = 1, \ldots, m,\]

where \(F^{(i)}(x^{(i)})\) and \(F^{(i,j)}(x^{(i)}, x^{(j)})\) denote the marginals c.d.f. of \(X_{\alpha}^{(i)}\) and \((X_{\alpha}^{(i)}, X_{\alpha}^{(j)})'\) respectively.

They based their test of independence on the statistic \(S^J = |D_n| \times (|D_{11}| |D_{22}|)^{-1}\), where \(|A|\) denotes the determinant of \(A\). They also showed that under \(H_0\), \(-n \log S^J \xrightarrow{L} \chi^2_{pq}\). With \(J_0\), the statistic \(S^J\) is a generalization of Spearman’s \(\rho\) for multivariate data sets.

Using the results of Puri, Sen and Gokhale (1970) with \(J_0(u)\), Cleroux, Lazraq and Lepage (1995) and Lazraq, Lepage and Cleroux (1995) proposed other tests of independence between two or more random vectors which are based on the measures of multivariate association proposed by Escoufier (1973) and Cramer and Niewodzicki (1979).

In the present paper, we present an approach based on an original concept of U-statistics matrix inspired from Hoeffding (1948) to the problem of detecting dependence between two random vectors. This theoretical tool allows us to deduce the asymptotic distribution of a general association matrix. The first application is to construct the association matrix with Kendall’s \(\tau\) and study its relationship with Spearman’s \(\rho\). We also propose nonparametric tests of independence between two random vectors based on three known measures of multivariate relationship with the Kendall and Spearman association matrices. We obtain the asymptotic distribution of the tests statistics under the null hypothesis and under a sequence of alternatives. In order to assess the behavior of the tests, a Monte Carlo study is performed to compare the empirical level and the empirical power of the tests based on Kendall’s matrix with those based on Spearman’s matrix.

Some multivariate generalizations of the Kendall’s \(\tau\) correlation coefficient have been studied in the literature by Hays (1960), Simon (1977) and Joe (1990). They have used the Kendall’s \(\tau\) correlation coefficient to test the total independence but not for the independence of two or more random vectors.

The paper is organized as follows. In Section 2, we give the asymptotic distribution of the matrices of U-statistics and deduce those for Spearman’s matrix and Kendall’s matrix. Section 3 is concerned with the three known measures of multivariate relationship: some properties and their asymptotic distributions under the null hypothesis and under a sequence of alternatives are given. In Section 4, we propose some tests of independence based on Spearman’s and Kendall’s matrices. We illustrate all the tests by an example.
Finally, Section 5 contains an empirical comparison of the new tests based on Kendall’s matrix with the competitors based on Spearman’s matrix. The results of this paper, can easily be extended to test the independence between several random vectors.

2. U-statistics matrix

Let $X_1, \ldots, X_n$ be $n$ independent random vectors, $X_\alpha = (X^{(1)}_\alpha, \ldots, X^{(m)}_\alpha)'$, $\alpha = 1, \ldots, n$, from an unknown continuous c.d.f. $F$. Let $\Phi^{(i,j)}(x_1, \ldots, x_{r(i,j)})$, for $i = 1, \ldots, p$ and $j = 1, \ldots, q$, be symmetric function with $r^{(i,j)} (r^{(i,j)} \in \mathbb{N})$ arguments. Let

$$U_n^{(i,j)} = \frac{1}{n} \sum_{\beta \in B} \Phi^{(i,j)}(X_{\beta_1}, \ldots, X_{\beta_{r(i,j)}}),$$

where $B = \{\beta = (\beta_1, \ldots, \beta_{r(i,j)}) \mid 1 \leq \beta_1 < \cdots < \beta_{r(i,j)} \leq n\}$, be a U-statistic for the parameter $\gamma^{(i,j)}$ of degree $r^{(i,j)}$ based on the symmetric kernel $\Phi^{(i,j)}$. Let

$$(2.1) \quad \Phi^{(i,j)}_1(x) = \mathbb{E} [\Phi^{(i,j)}(x, X_2, \ldots, X_{\beta_{r(i,j)}})], \quad \text{for } i = 1, \ldots, p \text{ and } j = 1, \ldots, q.$$

We note that $\mathbb{E} [U_n^{(i,j)}] = \mathbb{E} [\Phi^{(i,j)}_1(X)] = \gamma^{(i,j)}$ (see Hoeffding, 1948)). We now define the matrices of U-statistics $U_n$, of degrees $R$ and of parameters $\Gamma$ by respectively

$$U_n = \left( \begin{array}{c c c}
U^{(1,1)} & \cdots & U^{(1,q)} \\
\vdots & \ddots & \vdots \\
U^{(p,1)} & \cdots & U^{(p,q)}
\end{array} \right), \quad R = \left( \begin{array}{c c c}
r^{(1,1)} & \cdots & r^{(1,q)} \\
\vdots & \ddots & \vdots \\
r^{(p,1)} & \cdots & r^{(p,q)}
\end{array} \right)$$

and

$$\Gamma = \left( \begin{array}{c c c}
\gamma^{(1,1)} & \cdots & \gamma^{(1,q)} \\
\vdots & \ddots & \vdots \\
\gamma^{(p,1)} & \cdots & \gamma^{(p,q)}
\end{array} \right).$$

Consider vec $U_n$, as the vector formed by stacking the columns of $U_n$. The asymptotic multivariate normality of vec $U_n$ follows from Theorem 7.1 of Hoeffding (1948).

**Theorem 2.1.** If the kernel function $\Phi^{(i,j)}$ for the parameter $\gamma^{(i,j)}$ of degree $r^{(i,j)}$ is such that

$$\mathbb{E} [\Phi^{(i,j)}(X_1, \ldots, X_{r(i,j)})] = \gamma^{(i,j)} \quad \text{and} \quad \mathbb{E} [(\Phi^{(i,j)}(X_1, \ldots, X_{r(i,j)}))^2] < \infty,$$

for $i = 1, \ldots, p$ and $j = 1, \ldots, q$, then $\sqrt{n}(\text{vec } U_n - \text{vec } \Gamma) \xrightarrow{\mathcal{L}} \mathcal{N}_{pq}(0, \Omega)$ where the elements of $\Omega$ are given by

$$m^{(i,j,k,l)} = r^{(i,j)} r^{(k,l)} [\mathbb{E} [\Phi^{(i,j)}_1(X_1) \Phi^{(k,l)}_1(X_1)] - \gamma^{(i,j)} \gamma^{(k,l)}]$$
\( \Phi_1^{(i,j)}(x) \) and \( \Phi_1^{(k,l)}(x) \) are given in (2.1) for \( i, k = 1, \ldots, p \) and \( j, l = 1, \ldots, q \).

We can also deduce from Hoeffding (1948) that \( \text{vec } U_n \xrightarrow{P} \text{vec } \Gamma \).

### Spearman’s matrix

To express the rank correlation in terms of indicators, we define the signum function as

\[ s(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases} \]

Then we can define the U-statistic

\[ S_n^{(i,j)} = \frac{1}{n(n-1)/2} \sum_{1 \leq \alpha < \beta < \nu \leq n} \Psi^{(i,j)}(X_\alpha, X_\beta, X_\nu) \]

for Spearman’s coefficient \( \rho^{(i,j)} \) of degree 3 based on the kernel function

\[ \Psi^{(i,j)}(X_1, X_2, X_3) = \frac{1}{2} \sum_{1 \leq \alpha \neq \beta \neq \nu \leq 3} s(X_\alpha^{(i)} - X_\beta^{(i)}) s(X_\alpha^{(j)} - X_\nu^{(j)}). \]

Here, we have (see Hoeffding, 1948, p. 320)

\begin{align*}
(2.2) \quad \Psi^{(i,j)}(X_\alpha) & = [1 - 2F^{(i)}(X_\alpha^{(i)})][1 - 2F^{(j)}(X_\alpha^{(j)})] \\
& + 4 \int [F^{(i,j)}(x^{(i)}, X_\alpha^{(j)}) - F^{(i)}(x^{(i)})F^{(j)}(X_\alpha^{(j)})] dF^{(i)}(x^{(i)}) \\
& + 4 \int [F^{(i,j)}(X_\alpha^{(i)}, x^{(j)}) - F^{(i)}(X_\alpha^{(i)})F^{(j)}(x^{(j)})] dF^{(j)}(x^{(j)})
\end{align*}

where \( \Psi^{(i,j)}(x) = E[\Psi^{(i,j)}(x, X_2, X_3)] \). For \( i = j \), we have \( S_n^{(i,j)} = \rho^{(i,j)} = 1 \). Obviously \( S_n^{(i,j)} \) is an unbiased estimator of \( \rho^{(i,j)} \) while \( D^{(i,j)} \) given by (1.2) is not.

The matrix \( S_n = (S_n^{(i,j)})_{i,j=1,\ldots,m} \) will be called Spearman’s matrix for the parameter matrix \( P = (\rho^{(i,j)})_{i,j=1,\ldots,m} \). For all \( i \neq j \) the degree is 3 and zero for \( i = j \). The application of Theorem 2.1 leads immediately to the following theorem.

**Theorem 2.2.** The random vector \( \sqrt{n}(\text{vec } S_n - \text{vec } P) \) has a limiting \( m^2 \)-multivariate normal distribution \( N_{m^2}(0, \Sigma_S) \) where the elements of \( \Sigma_S \) are given by

\[
\Sigma_S^{(i,j),kl} = 9 \sum_{h=1}^{3} \sum_{h'=1}^{3} \text{Cov}(V_1^{(i,j),h}, V_1^{(k,l),h'})
\]

with \( i, j, k, l = 1, \ldots, m, \)

\[
V_1^{(i,j),1} = [1 - 2F^{(i)}(X_1^{(i)})][1 - 2F^{(j)}(X_1^{(j)})],
\]

\[
V_1^{(i,j),2} = 4 \int [F^{(i,j)}(x^{(i)}, X_1^{(j)}) - F^{(i)}(x^{(i)})F^{(j)}(X_1^{(j)})] dF^{(i)}(x^{(i)})
\]

and

\[
V_1^{(i,j),3} = 4 \int [F^{(i,j)}(X_1^{(i)}, x^{(j)}) - F^{(i)}(X_1^{(i)})F^{(j)}(x^{(j)})] dF^{(j)}(x^{(j)}).\]
Kendall’s matrix

Kendall’s \( \tau \) is a measure defined by the product moment correlation of signs of concordance,

\[
K_n^{(i,j)} = \frac{1}{\binom{n}{2}} \sum_{1 \leq \alpha < \beta \leq n} s(X^{(i)}_\beta - X^{(i)}_\alpha)s(X^{(j)}_\beta - X^{(j)}_\alpha),
\]

while Spearman’s rank correlation coefficient is the product moment correlation between \( F^{(i)}(X^{(i)}) \) and \( F^{(j)}(X^{(j)}) \) \( (i, j = 1, \ldots, m) \) (see Cléroux, Lazraq and Lepage, 1995, p. 719). Thus, Theorem 4.1 in Puri, Sen and Gokhale (1970) cannot be used to obtain the asymptotic multivariate normality of the elements of the Kendall’s matrix. The element \( K_n^{(i,j)} \) is a U-statistic of degree 2 based on the symmetric kernel

\[
\Phi^{(i,j)}(X_1, X_2) = s(X^{(i)}_2 - X^{(i)}_1)s(X^{(j)}_2 - X^{(j)}_1)
\]

for Kendall’s coefficient defined as

\[
\tau^{(i,j)} = 4 \int \int F^{(i,j)}(x^{(i)}, x^{(j)}) dF^{(i,j)}(x^{(i)}, x^{(j)}) - 1.
\]

Here also (see Hoeffding, 1948, p. 316), we have

\[
(2.3) \quad \Phi_1^{(i,j)}(X_\alpha) = 1 - 2F^{(i)}(X^{(i)}_\alpha) - 2F^{(j)}(X^{(j)}_\alpha) + 4F^{(i,j)}(X^{(i)}_\alpha, X^{(j)}_\alpha)
\]

\[
= [1 - 2F^{(i)}(X^{(i)}_\alpha)][1 - 2F^{(j)}(X^{(j)}_\alpha)]
\]

\[
+ 4[F^{(i,j)}(X^{(i)}_\alpha, X^{(j)}_\alpha) - F^{(i)}(X^{(i)}_\alpha)F^{(j)}(X^{(j)}_\alpha)]
\]

where \( \Phi_1^{(i,j)}(x) = \text{E}[\Phi^{(i,j)}(x, X_2)]. \) For \( i = j, \) we have \( K_n^{(i,j)} = \tau^{(i,i)} = 1. \)

The matrix \( K = (K_n^{(i,j)})_{i,j=1,...,m} \) will be called Kendall’s matrix for the parameter matrix \( \Lambda = (\tau^{(i,j)})_{i,j=1,...,m}. \) For all \( i \neq j, \) the degree is 2 while it is zero for \( i = j. \) The application of Theorem 2.1 leads immediately to the following theorem.

**Theorem 2.3.** The random vector \( \sqrt{n}(\text{vec } K_n - \text{vec } \Lambda) \) has a limiting \( m^2 \)-multivariate normal distribution \( N_{m^2}(O, \Sigma_K) \) where the elements of \( \Sigma_K \) are given by \( \sigma_K^{(i,j,k,l)} = 4 \sum_{h=1}^{2} \sum_{h'=1}^{2} \text{Cov}(U_1^{(i,j),h}, U_1^{(k,l),h'}) \) with \( i,j,k \) and \( l = 1, \ldots, m, \)

\[
U_1^{(i,j),1} = [1 - 2F^{(i)}(X^{(i)}_1)][1 - 2F^{(j)}(X^{(j)}_1)]
\]

and

\[
U_1^{(i,j),2} = 4[F^{(i,j)}(X^{(i)}_1, X^{(j)}_1) - F^{(i)}(X^{(i)}_1)F^{(j)}(X^{(j)}_1)].
\]
If we insert the rank \( R_{\alpha}^{(i)} \) of \( X_{\alpha}^{(i)} \) defined by

\[
R_{\alpha}^{(i)} = \frac{n+1}{2} + \frac{1}{2} \sum_{\beta=1}^{n} s(X_{\alpha}^{(i)} - X_{\beta}^{(i)})
\]

in (1.2), we have

\[
D_n = \frac{n-2}{n+1} \mathbf{s}_n + \frac{3}{n+1} K_n
\]

(see Hoeffding, 1948, p. 318). Then, \( \sqrt{n}(\text{vec } D_n - \text{vec } P) \) and \( \sqrt{n}(\text{vec } \mathbf{s}_n - \text{vec } P) \) have the same limiting distribution given by Theorem 2.2; we find here the result given by Puri, Sen and Gokhale (1970).

Let us now partition \( P \) and \( \Lambda \) and their analogue sample matrices \( \mathbf{s}_n \) and \( \mathbf{K}_n \) in following way:

\[
P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}, \quad K_n = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}
\]

and

\[
\mathbf{s}_n = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}
\]

where \( M_{21} (M = P, \Lambda, S \text{ or } K) \) is of order \( q \times p \). Now, we have \( P_{21} = \Lambda_{21} = O \) under \( H_0 \). Under \( H_0 \), \( X_{\alpha}^{(i)} \) and \( X_{\alpha}^{(j)} \) are independent for \( i = p+1, \ldots, m \) and \( j = 1, \ldots, p \).

**Theorem 2.4.** Under \( H_0 \) and when \( n \rightarrow \infty \), we have

\[
\sqrt{n} \text{vec } K_{21} \xrightarrow{\mathcal{L}} \mathbf{Z}^{(\tau)} \quad \text{where } \mathbf{Z}^{(\tau)} \text{ follows a } \mathcal{N}_{pq}(O, \frac{4}{9} P_{11} \otimes P_{22})
\]

and

\[
\sqrt{n} \text{vec } \mathbf{s}_{21} \xrightarrow{\mathcal{L}} \mathbf{Z}^{(\varrho)} \quad \text{where } \mathbf{Z}^{(\varrho)} \text{ follows a } \mathcal{N}_{pq}(O, P_{11} \otimes P_{22}).
\]

**Proof.** From Theorem 2.3, we note that under \( H_0 \) the random vector \( \sqrt{n} \text{vec } K_{21} \) has a limiting multivariate normal distribution \( \mathcal{N}_{pq}(O, A) \) where the elements of \( A \) are given for \( i, j, k, l = 1, \ldots, m \) by

\[
\sigma_{K}^{(i,j,k,l)} = 4E(U_{1}^{(i,j)} U_{1}^{(k,l)}) \\
= 4E[1 - 2F^{(i)}(X_{1}^{(i)})][1 - 2F^{(k)}(X_{1}^{(k)})] \
\times E[1 - 2F^{(j)}(X_{1}^{(j)})][1 - 2F^{(l)}(X_{1}^{(l)})] \\
= \frac{4}{9} \varrho^{(i,k)} \varrho^{(j,l)}.
\]

Thus, \( A = \frac{4}{9} P_{11} \otimes P_{22} \). In a similar way, we can obtain the limiting multivariate distribution of \( \sqrt{n} \text{vec } \mathbf{s}_{21} \) under \( H_0 \). \( \Box \)
We shall now study the asymptotic distribution of $M_{21}$ ($M = S$ or $K$) under a sequence of alternatives $\{H_{1:n}, n = 1, 2, \ldots\}$ (see Puri, Sen and Gokhale, 1970) which specifies that

$$H_{1:n} : F(x) = F^{[1]}(x^{[1]}) F^{[2]}(x^{[2]}) \left(1 + \frac{\Omega^{[1][2]}(F^{[1]}(x^{[1]}), F^{[2]}(x^{[2]}))}{\sqrt{n}}\right)$$

where $\Omega^{[1][2]}$ is some function of $(F^{[1]}(x^{[1]}), F^{[2]}(x^{[2]}))$ and $\Omega^{[1][2]} \neq 0$. $H_{1:n}$ implies that for $i = p + 1, \ldots, m$ and $j = 1, \ldots, p$,

$$F^{(i,j)}(x^{(i)}, x^{(j)}) = F^{(i)}(x^{(i)}) F^{(j)}(x^{(j)})$$

$$\times \left(1 + \frac{\Omega^{(i,j)}(F^{(i)}(x^{(i)}), F^{(j)}(x^{(j)}))}{\sqrt{n}}\right)$$

where $\Omega^{(i,j)}$ is a function of $(F^{(i)}, F^{(j)})$ and $\Omega^{(i,j)} \neq 0$; it also implies that

$$F^{(i,k,l)}(x^{(i)}, x^{(j)}, x^{(k)}, x^{(l)}) = F^{(i,k)}(x^{(i)}, x^{(k)}) F^{(j,l)}(x^{(j)}, x^{(l)})$$

$$\times \left(1 + \frac{\Omega^{(i,j,k,l)}(F^{(i,k)}(x^{(i)}, x^{(k)}), F^{(j,l)}(x^{(j)}, x^{(l)}))}{\sqrt{n}}\right)$$

where $F^{(i,k,l)}$ is the c.d.f. of the $(X^{(i)}, X^{(j)}, X^{(k)}, X^{(l)})$ for $j, l = 1, \ldots, p$; $i, k = p + 1, \ldots, m$, and $\Omega^{(i,j,k,l)} \neq 0$ is a function of $(F^{(i,k)}, F^{(j,l)})$.

Let for $i = p + 1, \ldots, m$ and $j = 1, \ldots, p$,

$$dF^{(i,j)} = f^{(i,j)}(x^{(i)}, x^{(j)}) dx^{(i)} dx^{(j)}$$

$$= \frac{\partial^2 F^{(i,j)}(x^{(i)}, x^{(j)})}{\partial x^{(i)} \partial x^{(j)}} dx^{(i)} dx^{(j)}$$

$$= dF^{(i)} dF^{(j)} \left(1 + \frac{1}{\sqrt{n}} \omega_{ij}(F^{(i)}(x^{(i)}), F^{(j)}(x^{(j)}))\right),$$

where the function $\omega_{ij}$ is obtained by differentiating (2.4)). In a similar way, let for $i = p + 1, \ldots, m$ and $j = 1, \ldots, p$,

$$f^{(i,k,l)}(x^{(i)}, x^{(j)}, x^{(k)}, x^{(l)})$$

$$= \frac{\partial^4 F^{(i,k,l)}(x^{(i)}, x^{(j)}, x^{(k)}, x^{(l)})}{\partial x^{(i)} \partial x^{(j)} \partial x^{(k)} \partial x^{(l)}}$$

$$= \frac{\partial^2 F^{(i,k)}(x^{(i)}, x^{(k)})}{\partial x^{(i)} \partial x^{(k)}} \frac{\partial^2 F^{(j,l)}(x^{(j)}, x^{(l)})}{\partial x^{(j)} \partial x^{(l)}} + \frac{1}{\sqrt{n}} \omega_{ij,k,l}.$$
and
\[ dF^{(i,k)} = f^{(i,k)}(x^{(i)}, x^{(k)}) dx^{(i)} dx^{(k)}. \]

Let \( B = (\beta^{(i,j)}) \) be the \( q \times p \) matrix where
\[ \beta^{(i,j)} = \int \int F^{(i)}(x^{(i)}) F^{(j)}(x^{(j)}) \Omega^{(i,j)}(F^{(i)}(x^{(i)}), F^{(j)}(x^{(j)})) dF^{(i,j)}. \]

Using (1.3) and \( \Psi^{(i,j)}_1 \) defined by (2.2) where
\[ E[\Psi^{(i,j)}_1(X_1)] = \varrho^{(i,j)} = 3 \int \int [2F^{(i)}(x^{(i)}) - 1][2F^{(j)}(x^{(j)}) - 1] dF^{(i,j)}, \]
we obtain under \( H_1:n \),
\[ \varrho^{(i,j)} = \frac{1}{3} \rho^{(i,j)} + 8 \int \int \left[ F^{(i,j)}(x^{(i)}, x^{(j)}) - F^{(i)}(x^{(i)}) F^{(j)}(x^{(j)}) \right] dF^{(i,j)} \\
= \frac{1}{3} \rho^{(i,j)} + \frac{8}{\sqrt{n}} \beta^{(i,j)} = \frac{12 \beta^{(i,j)}}{\sqrt{n}}. \]

In a similar way, using \( \Phi^{(i,j)}_1 \) defined by (2.3), we obtain under \( H_1:n \),
\[ \tau^{(i,j)} = \frac{1}{3} \rho^{(i,j)} + \frac{4}{\sqrt{n}} \beta^{(i,j)} = \frac{8 \beta^{(i,j)}}{\sqrt{n}}. \]

We thus have shown the following lemma.

**Lemma 2.1.** Under \( H_1:n \), we have
\[ \Lambda_{21} = \frac{8}{\sqrt{n}} B \quad \text{and} \quad P_{21} = \frac{12}{\sqrt{n}} B. \]

The next theorem gives the limiting distribution of \( K_{21} \) and \( S_{21} \) under the sequence \( H_1:n \).

**Theorem 2.5.** Under \( H_1:n \) and when \( n \to \infty \), we have
\[ \sqrt{n} \text{vec} K_{21} \overset{\mathcal{L}}{\to} Z^{(\tau)} \quad \text{where} \quad Z^{(\tau)} \text{ follows a } N_{pq}(8 \text{vec } B, \frac{4}{9} P_{11} \otimes P_{22}), \]
\[ \sqrt{n} \text{vec} S_{21} \overset{\mathcal{L}}{\to} Z^{(\rho)} \quad \text{where} \quad Z^{(\rho)} \text{ follows a } N_{pq}(12 \text{vec } B, P_{11} \otimes P_{22}). \]

**Proof.** From Theorem 2.3, the random vector \( \sqrt{n} \text{vec } K_{21} \) has a limiting multivariate distribution with mean vector \( E[\sqrt{n} \text{vec } K_{21}] = 8 \text{vec } B \) and covariance matrix \( \frac{4}{9} P_{11} \otimes P_{22} \) whose its elements are
\[ \sigma^{(i,j,k,l)}_K = 4 \text{Cov}(U^{(i,j),1}_1, U^{(k,l),1}_1) = \frac{4}{9} \varrho^{(i,k)} \varrho^{(j,l)}. \]
Using the expression for $dF^{(ij,kl)}$ given by equation (2.7), we have

$$
\text{Cov}(U^{(i,j),1}_1, U^{(k,l),1}_1) \\
= \text{E}[U^{(i,j),1}_1 U^{(k,l),1}_1] - \text{E}[U^{(i,j),1}_1] \text{E}[U^{(k,l),1}_1] \\
= \left( \int \int [1 - 2F^{(i)}(x^{(i)})][1 - 2F^{(k)}(x^{(k)})] dF^{(i,k)} \right) \\
\times \left( \int \int [1 - 2F^{(j)}(x^{(j)})][1 - 2F^{(l)}(x^{(l)})] dF^{(j,l)} \right) \\
+ \frac{1}{\sqrt{n}} \int \int \int \int [1 - 2F^{(i)}(x^{(i)})][1 - 2F^{(j)}(x^{(j)})][1 - 2F^{(k)}(x^{(k)})] \\
\times [1 - 2F^{(l)}(x^{(l)})] \omega_{ij,kl} \, dx^{(i)} \, dx^{(j)} \, dx^{(k)} \, dx^{(l)} \\
- \frac{8}{n} \beta^{(i,j)} \beta^{(k,l)} \\
= \frac{1}{9} \theta^{(i,k)} \theta^{(j,l)} \\
+ \frac{1}{\sqrt{n}} \int \int \int \int [1 - 2F^{(i)}(x^{(i)})][1 - 2F^{(j)}(x^{(j)})][1 - 2F^{(k)}(x^{(k)})] \\
\times [1 - 2F^{(l)}(x^{(l)})] \omega_{ij,kl} \, dx^{(i)} \, dx^{(j)} \, dx^{(k)} \, dx^{(l)} \\
- \frac{8}{n} \beta^{(i,j)} \beta^{(k,l)} \\
= \frac{1}{9} \theta^{(i,k)} \theta^{(j,l)} + O(n^{-1/2}).
$$

The result follows from Serfling (1980) (Lemma A, p. 20). In a similar way, we have the limiting distribution of $\sqrt{n} \text{vec} S_{21}$ from Theorem 2.2.

\[\square\]

3. Measures of association

We now apply the measures of multivariate relationship proposed by Escoufier (1973), Stewart and Love (1968) and Cramer and Nicewander (1979) to the Kendall and Spearman matrices.

For the Escoufier’s measure (1973), we have

$$
\text{RV}^\tau = \frac{\text{tr}(K_{12}K_{12}'_2)}{\sqrt{\text{tr}(K_{11}^2) \text{tr}(K_{22}^2)}} \quad \text{and} \quad \text{RV}^\theta = \frac{\text{tr}(S_{12}S_{12}')}{\sqrt{\text{tr}(S_{11}^2) \text{tr}(S_{22}^2)}}.
$$

The Stewart and Love’s measure (1968) gives

$$
\text{SL}^\tau = \frac{\text{tr}(K_{12}K_{22}^{-1}K_{12}')}{p} \quad \text{and} \quad \text{SL}^\theta = \frac{\text{tr}(S_{12}S_{22}^{-1}S_{12}')}{p}.
$$

Finally with the Cramer and Nicewander’s measure (1979), we have

$$
\text{CN}^\tau = \frac{\text{tr}(K_{11}^{-1}K_{12}K_{22}^{-1}K_{12}')}{p} \quad \text{and} \quad \text{CN}^\theta = \frac{\text{tr}(S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}')}{p}.
$$
The corresponding measures at the level of the population are defined by:

\[
\rho_{RV}^{(\tau)} = \frac{\text{tr}(\Lambda_{12}\Lambda_{12}')}{\sqrt{\text{tr}(\Lambda_{11}^2)\text{tr}(\Lambda_{22}^2)}} \quad \text{and} \quad \rho_{RV}^{(e)} = \frac{\text{tr}(P_{12}P_{12}')}{\sqrt{\text{tr}(P_{11}^2)\text{tr}(P_{22}^2)}}
\]

for the Escoufier’s measure,

\[
\rho_{SL}^{(\tau)} = \frac{\text{tr}(\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}')}{p} \quad \text{and} \quad \rho_{SL}^{(e)} = \frac{\text{tr}(P_{12}P_{22}^{-1}P_{12}')}{p}
\]

for the Stewart and Love’s measure,

\[
\rho_{CN}^{(\tau)} = \frac{\text{tr}(\Lambda_{11}^{-1}\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}')}{p} \quad \text{and} \quad \rho_{CN}^{(e)} = \frac{\text{tr}(P_{11}^{-1}P_{12}P_{22}^{-1}P_{12}')}{{p}}
\]

for the Cramer and Nicewander’s measure.

The main advantage of considering these transformed measures are that:

(a) the individual data may be ordinal variables, (b) the scale of measurement for each variable may be different, (c) the classical hypotheses of multivariate normality or ellipticity of the parent population may be omitted, (d) they lead to a robust procedure against outliers. Moreover, the three measures applied to Kendall’s matrix or Spearman’s matrix have the following properties:

(i) \( \rho M^{(\tau)} = \rho M^{(e)} = 0 \) if and only if \( P_{21} = \Lambda_{12} = 0 \), for \( M = RV, SL \) and CN.

(ii) when \( p = q = 1 \), the three measures reduce to the square of Kendall’s coefficient or to the square of Spearman’s coefficient between the variables \( X^{(1)} \) and \( X^{(2)} \).

(iii) \( 0 \leq \rho M^{(s)} \leq 1 \), for \( s = \tau, \varrho \) and \( M = RV, SL \) and CN. The sample analogue of the measures, \( M^{(s)} \), for \( s = \tau, \varrho \) and \( M = RV, SL \) and CN, have the same properties.

For the proof of these properties and other results on measures of multivariate relationship, the reader is referred to Lazraq and Cleroux (1988). The testing problem is now restated as \( H_0: \rho M^{(s)} = 0 \) versus \( \rho M^{(s)} > 0 \), for \( s = \tau, \varrho \) and \( M = RV, SL \) and CN.

In the following theorems we give the asymptotic distribution of our statistics under the null hypothesis and under a sequence of alternatives. We will show that they are represented as linear combinations of independent central \( \chi^2 \) and noncentral \( \chi^2 \) random variables respectively.
Theorem 3.1. Let $K_n$ and $S_n$ be Kendall’s and Spearman’s matrices respectively obtained from a sample of size $n$ drawn from a $m$-dimensional random vector with an arbitrary continuous c.d.f. $F(x)$. Then, under $H_0$ and when $n \to \infty$, we have

(i) 
$$nRV^{(r)} \xrightarrow{L} \frac{4}{9} \frac{1}{\sqrt{\text{tr}(\Lambda_{11}^2) \text{tr}(\Lambda_{22}^2)}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j U_{ij}^2,$$

(ii) 
$$nRV^{(q)} \xrightarrow{L} \frac{1}{\sqrt{\text{tr}(P_{11}^2) \text{tr}(P_{22}^2)}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j U_{ij}^2,$$

where the $U_{ij}$’s are iid $N(0,1)$, $i = 1, \ldots, p$; $j = 1, \ldots, q$, random variables and $\lambda_i$ and $\mu_j$ are the eigenvalues of $P_{11}$ and $P_{22}$ respectively.

(iii) 
$$nSL^{(r)} \xrightarrow{L} \frac{4}{9p} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i t_{ij}^{(2)} U_{ij}^2,$$

(iv) 
$$nSL^{(q)} \xrightarrow{L} \frac{1}{p} \sum_{i=1}^{p} \lambda_i Z_{q,i}^2,$$

where the $Z_{q,i}$’s are iid $X_q^2$, $i = 1, \ldots, p$, random variables with $q$ degrees of freedom, $\lambda_i$, $i = 1, \ldots, p$ are the eigenvalues of $P_{11}$ and $t_{ij}^{(2)}$, $j = 1, \ldots, q$, are the eigenvalues of $\Lambda_{11}^{-1}P_{12}$. 

(v) 
$$nCN^{(r)} \xrightarrow{L} \frac{4}{9p} \sum_{i=1}^{p} \sum_{j=1}^{q} t_{i}^{(1)} t_{j}^{(2)} U_{ij}^2,$$

(vi) 
$$nCN^{(q)} \xrightarrow{L} \frac{\Lambda_{pq}^2}{p},$$

where $t_{i}^{(1)}$, $i = 1, \ldots, p$, are the eigenvalues of $\Lambda_{11}^{-1}P_{12}$.

Proof. (i) Since $K_n$ converges in probability to $\Lambda$ as $n \to \infty$, the submatrices $K_{11}$ and $K_{22}$ converges in probability to $\Lambda_{11}$ and $\Lambda_{22}$ respectively as $n \to \infty$. Furthermore, under $H_0$, $\sqrt{n} \text{vec} K_{21}$ converges to $Z^{(r)}$ with distribution $N_{pq}(O, \frac{4}{9}P_{11} \otimes P_{22})$ Theorem 2.3. Since

$$n \text{tr}(K_{12}K_{21}) = (\sqrt{n} \text{vec} K_{21})'(\sqrt{n} \text{vec} K_{21}) \xrightarrow{L} Z^{(r)'} Z^{(r)},$$

we deduce using classical results on quadratic form (see Baldessari, 1967 or Johnson and Kotz, 1970) that,

$$nRV^{(r)} \xrightarrow{L} \frac{4}{9} \frac{1}{\sqrt{\text{tr}(\Lambda_{11}^2) \text{tr}(\Lambda_{22}^2)}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j U_{ij}^2,$$
where the \( U_{ij} \)'s are iid \( \mathcal{N}(0,1) \), \( i = 1, \ldots, p; j = 1, \ldots, q \), random variables and \( \lambda_i, \mu_j \) are the eigenvalues of \( P_{11}, P_{22} \) respectively.

Noting that \( n \text{tr}(K_{12}K_{22}^{-1}K_{21}) = (\sqrt{n} \text{vec } K_{21})'(I_p \otimes K_{22})^{-1}(\sqrt{n} \text{vec } K_{21}) \), we have

\[
n\text{SL}^{(\tau)} \overset{L}{\sim} \frac{4}{9p} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i t_j^{(2)} U_{ij}^2
\]

where \( t_j^{(2)}, j = 1, \ldots, q \), are the eigenvalues of \( \Lambda_{22}^{-1}P_{22} \). For the case \( n\text{CN}(\tau) \), we use

\[
n \text{tr}(K_{11}^{-1}K_{12}K_{22}^{-1}K_{21}) = (\sqrt{n} \text{vec } K_{21})'(K_{11} \otimes K_{22})^{-1}(\sqrt{n} \text{vec } K_{21}).
\]

The proofs are analogous when Spearman's matrix is used.

\[\square\]

**Theorem 3.2.** If the conditions of Theorem 3.1 are satisfied then under \( H_{1:n} \) and when \( n \rightarrow \infty \), we have

(i) \[
n\text{RV}^{(\tau)} \overset{L}{\sim} \frac{4}{9} \frac{1}{\sqrt{\text{tr}(\Lambda_{11}^2) \text{tr}(\Lambda_{22}^2)}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j U_{ij,1}^2,
\]

where the \( U_{ij,1} \)'s are independent \( \mathcal{N}(\delta_{ij}, 1) \), \( i = 1, \ldots, p; j = 1, \ldots, q \), random variables;

(ii) \[
n\text{RV}^{(\omega)} \overset{L}{\sim} \frac{1}{\sqrt{\text{tr}(P_{11}^2) \text{tr}(P_{22}^2)}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j U_{ij,2}^2,
\]

where the \( U_{ij,2} \)'s are independent \( \mathcal{N}(\sqrt{\frac{9}{4} \delta_{ij}^2}, 1) \), \( i = 1, \ldots, p; j = 1, \ldots, q \), random variables and \( \lambda_i, \mu_j \) are the eigenvalues of \( P_{11}, P_{22} \) resp. corresponding to the normalized eigenvectors \( a_i, b_j, \delta_{ij}^2 = 64 \text{tr}(B'b_jb_j'P_{22}^{-1}BP_{11}^{-1}a_i a_i') \) and \( B \) is the matrix defined in Lemma 2.1;

(iii) \[
n\text{SL}^{(\tau)} \overset{L}{\sim} \frac{4}{9p} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i t_j^{(2)} \chi_{1,ij}^2(\delta_{ij,1}^2),
\]

where the \( \chi_{1,ij}^2(\delta_{ij,1}^2) \)'s are independent chi-squared random variables with one degree of freedom, with \( \delta_{ij,1}^2 = \text{tr}(B'b_jb_j'P_{22}^{-1}BP_{11}^{-1}a_i a_i') \) as noncentrality parameter and \( p_j, j = 1, \ldots, q \), is the normalized eigenvector corresponding to the eigenvalue \( t_j^{(2)} \) of \( \Lambda_{22}^{-1}P_{22} \);

(iv) \[
n\text{SL}^{(\omega)} \overset{L}{\sim} \frac{1}{p} \sum_{i=1}^{p} \lambda_i \chi_{i(q)}^2(\delta_i^2),
\]
where the $\chi^2_{i(q)}(\delta^2)$'s, $i = 1, \ldots, p$, random variables are independent chi-squared random variables with $q$ degrees of freedom and noncentrality parameter defined by $\delta^2_i = 64 \text{tr}(B'^{-1}P^{-1}a_i a'_i)$;

\[(v)\] 
$$n \text{CN}(\tau) \xrightarrow{L} \frac{4}{9p} \sum_{i=1}^{p} \sum_{j=1}^{q} t_i^{(1)} t_j^{(2)} U_{ij,3},$$

where the $U_{ij}$'s are independent $\mathcal{N}(\Delta_{ij}, 1)$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, random variables with $\Delta_{ij}^2 = 64 \text{tr}(B'p_j P^{-1}BP'^{-1}d_id'_i)$ as noncentrality parameter and $d_i$, $i = 1, \ldots, p$, is the normalized eigenvector corresponding to the eigenvalue $t_i^{(1)}$ of $\Lambda_{11}^{-1}P_{11}$;

\[(vi)\] 
$$n \text{CN}(\rho) \xrightarrow{L} \frac{\chi^2_{pq}(\delta^2)}{p},$$

where the $\chi^2_{pq}$ random variable has $pq$ degrees of freedom and noncentrality parameter defined as $\delta^2 = \text{tr}(B'P^{-1}BP'^{-1})$.

The proof of this Theorem is analogous to Theorem 3.1, but a noncentrality parameter is introduced in the asymptotic distribution of $n\text{RV}(\tau)$, $n\text{SL}(\tau)$ and $n\text{CN}(\tau)$.

4. Tests of independence of two vectors

The results of the preceding section can be used to construct asymptotic tests of independence between two vectors. We will test for $M = \text{RV}, \text{SL}$ or $\text{CN}$ and $s = \tau$ or $\rho$, $H_0: \rho M(s) = 0$ against $\rho M(s) > 0$ at level $\alpha$ by rejecting $H_0$ if $nM(s) > c^{(s,M)}_{\alpha}$ where $c^{(s,M)}_{\alpha}$ is the $100(1 - \alpha)$th percentile of the corresponding distribution given in Theorem 3.1. Under $H_{1:n}$, $nM(s)$ converges in probability to $\rho M(s)$ for $M = \text{RV}, \text{SL}, \text{CN}$ and $s = \tau, \rho$ and thus the asymptotic power of each of these six tests converges to 1 when $n \to \infty$. Thus, each test is consistent.

The limiting distributions given in Theorem 3.1 are not easy to deal with and consequently, the percentiles will be computed by using Imhof’s algorithm (Imhof, 1961). Moreover, in these distributions, $P_{11}, P_{22}, \Lambda_{11}$ and $\Lambda_{22}$ are usually unknown, we thus use instead the estimators $S_{11}, S_{22}, K_{11}$ and $K_{22}$. Since the estimators are consistent, the asymptotic distributions remain unchanged.

Let us notice that the tests $nM(\tau)$ ($M = \text{RV}, \text{SL}$ and $\text{CN}$) based on the matrix of Kendall depend on the tests $M(\rho)$ based on the matrix of Spearman. For example, the asymptotic distribution of $n\text{RV}(\tau)$ and $n\text{RV}(\rho)$ use the same eigenvalues resulting from the submatrices $P_{11}$ and $P_{22}$ of Spearman’s matrix. They are asymptotically equivalent, up to a multiplicative coefficient which depends on Kendall matrix. In the case of total independence, this constant is $\frac{4}{9}$ and this is already mentioned by several authors (see, for example, Hájek and Šidák, 1967).
Description of the procedure

Given a sample of size $n$, $(X_1^{[1]}, X_1^{[2]})', \ldots, (X_n^{[1]}, X_n^{[2]})'$ where $X_i^{[1]}: p \times 1$ and $X_i^{[2]}: q \times 1$ for $i = 1, \ldots, n$.

**Step 1**: Compute $K_{11}, K_{22}, K_{12}, K_{21}$ and $S_{11}, S_{22}, S_{12}, S_{21}$.

**Step 2**: Compute the required eigenvalues from the consistent estimators.

**Step 3**: Compute $nM^s$ for $M = \text{RV}, \text{SL}, \text{CN}$ and $s = \tau, \varrho$.

**Step 4**: For each distribution given by Theorem 3.1, obtain the $100(1 - \alpha)$th percentile, $c_{\alpha}^{(s,M)}$, for $M = \text{RV}, \text{SL}, \text{CN}$ and $s = \tau, \varrho$, by using the Imhof (1961) algorithm.

**Step 5**: Reject $H_0$ at level $\alpha$ if $\rho M^s > c_{\alpha}^{(s,M)}$, for $M = \text{RV}, \text{SL}, \text{CN}$ and $s = \tau, \varrho$.

**Example.** The six tests are illustrated with sport data. The data consist of the 1984 Olympic track records of 55 nations for women as well as men (see Naik and Khattree, 1996). The data matrix for women is a $55 \times 7$ matrix with seven events represented: the 100 meters, 200 meters, 400 meters, 800 meters, 1500 meters, 3000 meters and marathon (which is 42195 meters). For the men the corresponding matrix is of order $55 \times 8$ differing from the women’s events in that the 3000 meters was excluded but 5000 meters and 10000 meters were included.

As noted by Naik and Khattree (1996), to test athletic performances of women and men, the appropriate variable that may be more relevant in this context is the speed, defined as the “distance covered per unit of time.” This variable succeeds in retaining the possibility of having different degrees of variability. We will therefore use the speed in the track events as the variable for the tests of independence between women and men performances. These two data sets are presented in Tables 1 and 2 of Naik and Khattree (1996).

First, we test the hypothesis $H_0$ of independence between $X^{[1]}$ and $X^{[2]}$ where $X^{[1]}$ is the vector formed by women performances and $X^{[2]}$ is the vector formed by men performances. We have $n = 55$, $p = 7$ and $q = 8$. Table 1 gives the value of the statistic, the 5% critical value and the observed critical value. Therefore, $H_0$ is strongly rejected.

5. Simulation study

In order to assess the behavior of the tests based on Kendall’s matrix, a Monte-Carlo study is performed to compare its empirical level and its empirical power with those of the three competitors based on Spearman’s matrix (see Cléroux, Lazraaq and Lepage, 1995).
Table 1. Tests of independence between women and men performances, the value of the statistic, the 5% critical value and the observed critical value.

<table>
<thead>
<tr>
<th>Matrix</th>
<th>Statistic</th>
<th>Value</th>
<th>Critical point $C_{0.05}$</th>
<th>Critical level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman</td>
<td>$RV_1(\rho)$</td>
<td>44.88</td>
<td>4.33</td>
<td>$1.19 \times 10^{-7}$</td>
</tr>
<tr>
<td></td>
<td>$SL_1(\rho)$</td>
<td>42.14</td>
<td>14.32</td>
<td>$1.19 \times 10^{-7}$</td>
</tr>
<tr>
<td></td>
<td>$CN_1(\rho)$</td>
<td>17.18</td>
<td>10.63</td>
<td>$1.25 \times 10^{-6}$</td>
</tr>
<tr>
<td>Kendall</td>
<td>$RV_1(\tau)$</td>
<td>37.16</td>
<td>2.77</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$SL_1(\tau)$</td>
<td>28.67</td>
<td>3.95</td>
<td>$1.19 \times 10^{-7}$</td>
</tr>
<tr>
<td></td>
<td>$CN_1(\tau)$</td>
<td>8.21</td>
<td>1.57</td>
<td>$1.19 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

All the simulation programs were written in FORTRAN programming language. For ease of comparison, the study is restricted to the case $p = 2$, $q = 3$ and the nominal level 1%. The number of repetitions at each setting is 10 000. Two types of underlying distributions are imposed. In the family of elliptic distributions, we consider a multivariate distribution $N_p(O, \Sigma)$ and an elliptic multivariate $t_5$. In the family of nonelliptic distributions, we consider a multivariate logistic $U$ (see Johnson, 1987) and a general multivariate distribution constructed as follows: each component of the vector $X$ is independently generated from the other, the first is $N_1(O, 1)$, the second is uniform on $[0, 1]$ minus 0.5 and multiplied by $\sqrt{12}$, the third is an exponential distribution (with parameter 1) minus 1, the fourth is a beta (with parameters 2 and 2) minus 0.5 and multiplied by $\sqrt{20}$ and finally the fifth is a gamma distribution (with parameters 1 and 4) minus 4 and divided by 2.

Under $H_0$, we generate two independent random vectors $X_1$ and $X_2$. For the alternative hypothesis, we consider the linear transformation $Y = CX$ where $C$ is such that $\Sigma = CC'$. The matrices considered here are

$$
\Sigma_{11} = I_2, \quad \Sigma_{22} = I_3, \quad \Sigma_{12} = \Sigma'_{21} = C_{00}, C_{10}, C_{15} \text{ and } C_{20}
$$

where the matrices $C_{xy}$ represent $2 \times 3$ matrices with all elements being the real number $0.xy$; for example, all elements of $C_{15}$ are equal to 0.15. This type of matrices was used and justified by Cléroux, Lazraq and Lepage (1995).

Table 2 summarizes the simulation results for the five distributions. In order to judge the empirical level of the asymptotic tests and their empirical power, an empirical level will be good if the nominal level 1% belongs to the 95% confidence interval. So that, for 10 000 repetitions, $C_{00}$ column must vary between 79 and 121.

The first observation is that for Kendall’s tests and Spearman’s tests, the empirical power of each test increases with departure from the null hypothesis that is when the value $xy$ of the matrices $C_{xy}$ increases. The empirical levels of $nM(\omega)$ are in general slightly conservative while that of
<table>
<thead>
<tr>
<th>Kendall</th>
<th>Spearman</th>
<th>Kendall</th>
<th>Spearman</th>
<th>Kendall</th>
<th>Spearman</th>
<th>Kendall</th>
<th>Spearman</th>
<th>Kendall</th>
<th>Spearman</th>
<th>Kendall</th>
<th>Spearman</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

For the multivariate distributions with $n = 2$ and $z = d$, the tests based on Spearman's matrix and Kendall's matrix at nominal level 1%
Kendall and Spearman for Multivariate Data Sets

\( nM^{(r)} (M = \text{RV, SL or CN}) \) are liberal. The tests \( nM^{(e)} \) have an empirical power slightly inferior to \( nM^{(r)} (M = \text{RV, SL or CN}) \). In each class of tests (Kendall or Spearman), we notice that the empirical power of the tests \( nSL^{(s)} (s = \tau \text{ or } \rho) \) is greater than the empirical power of the other tests, but when the underlying distribution is logistic, the empirical power of \( nRV^{(s)} (s = \tau \text{ or } \rho) \) is greater than the two others. In conclusion, the empirical power of each test, in a given class, depends on the underlying distribution. Nevertheless, one notices that the tests of Kendall’s class are empirically more powerful than the tests of Spearman’s class especially for small sample sizes and in the vicinity of the null hypothesis \( C_{00} \).

REFERENCES


---

**Hamani El Maache**  
Dép. de mathématiques et de statistique  
Université de Montréal  
C.P. 6128, succ. Centre-Ville  
Montréal QC H3C 3J7  
Canada  
elmaach@dms.umontreal.ca

**Yves Lepage**  
Dép. de mathématiques et de statistique  
Université de Montréal  
C.P. 6128, succ. Centre-Ville  
Montréal QC H3C 3J7  
Canada  
lepage@dms.umontreal.ca