Multivariate Nonparametric Tests of Independence

Y. Fan², P. Lafaye de Micheaux^{1,2}, S. Penev², D. Salopek²

¹Professor of Statistics *CREST-ENSAI, France* ²School of Mathematics and Statistics, *University of New South Wales, Australia*

Australian Statistical Conference

Hotel Realm, Canberra December 8, 2016

An example of non-monotone dependence

Birth rate against mortality rate in 145 countries (https://www.cia. gov/library/publications/the-world-factbook/geos/xx.html)



- ▶ Pearson's correlation is equal to 0.139, with a p-value of 0.095.
- ► Spearman's rho is equal to -0.035, with a p-value of 0.675
- ► Kendall's tau is equal to -0.032, with a p-value of 0.574.

We see a pattern \Rightarrow another tool to measure dependence is needed

[Bakirov et al., 2006], [Székely et al., 2007] developped a dependence measure between TWO random vectors, called distance correlation :

- $0 \le d\mathbb{C}\operatorname{or}(\mathbf{X}, \mathbf{Y}) \le 1$;
- $d\mathbb{C}or(\mathbf{x}, \mathbf{y}) = 1 \Rightarrow \mathbf{y} = \alpha + B\mathbf{x}$ (also true for the empirical);
- $d\mathbb{C}or(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x}$ and \mathbf{y} independent $\Rightarrow \mathbf{x}$ and \mathbf{y} non correlated;
- can detect any form of dependence !

Note : The empirical version of $d\mathbb{C}or(\mathbf{x}, \mathbf{y})$, noted $dcor(\mathbf{x}, \mathbf{y})$, has a simple formula built using empirical characteristic functions.

Implemented in the energy R package.

Independence test based on the distance correlation

Let be given **two** observed samples $X : n \times p$ and $Y : n \times q$, each constituted of *n* realizations of the random vectors **x** and **y**. Define :

$$a_{kl} = |\mathbf{x}_k - \mathbf{x}_l|_p, \ A_{kl} = a_{kl} - \bar{a}_{k\bullet} - \bar{a}_{\bullet l} + \bar{a}_{\bullet \bullet}, \ b_{kl} = |\mathbf{y}_k - \mathbf{y}_l|_q, \ B_{kl} = b_{kl} - \bar{b}_{k\bullet} - \bar{b}_{\bullet l} + \bar{b}_{\bullet \bullet},$$

$$dcov^2(\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}) = n^{-2}\sum_{k,l=1}^n A_{kl}B_{kl} \text{ and } s_2 = n^{-2}\sum_{k,l=1}^n a_{kl}n^{-2}\sum_{k,l=1}^n b_{kl}.$$

We reject, at significance level at most α (use of a bound, or of random permutations), the hypothesis of independence between x and y if

$$\sqrt{n} \times dcov(\boldsymbol{X}, \boldsymbol{\mathcal{Y}})/\sqrt{s_2} > \Phi^{-1}(1-\alpha/2).$$

Is it necessary to go further than pairwise dependence?

$$\begin{array}{c} x1 \\ x2 \\ y \end{array} \right\} 3 \text{ sample } (n \times 1) \text{ i.i.d. } \mathcal{N}(0,1) \text{ and } x3 = \left\{ \begin{array}{c} +abs(y), & \text{if } x1*x2 > 0; \\ -abs(y), & \text{otherwise.} \end{array} \right.$$





... but ...

Values of x3 such that x1*x2 > 0 are all positive ! 0.6408445 1.6013778 1.7164682 0.4069357 1.7282346 ··· Values of x3 such that x1*x2 <= 0 are all negative ! -0.7778154 -1.6473925 -0.1542662 -1.1756313 -1.1021281 ···



If you give me the sign of x1*x2, I can give you the sign of x3!

Going further than pairwise dependence



Hence, here the interaction with a third variable is relevant to understand the relation between birth rate and mortality rate.

* Credit : G. Boglioni

Towards mutual dependence between p sub-vectors

We thus need a test of independence between several variables.

We have generalized the *distance correlation* of [Bakirov et al., 2006] and [Székely et al., 2007] to obtain a test procedure for investigating the mutual dependence between $p \ge 2$ vectors. (First improvement over the distance correlation test)

Let $\underline{\mathbf{x}}_1, \ldots, \underline{\mathbf{x}}_n$ be a random sample, namely *n* independent copies of

$$\underline{\mathbf{x}} = \begin{pmatrix} \mathbf{x}^1 \\ \vdots \\ \mathbf{x}^p \end{pmatrix} \in \mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_p}.$$

Similarly, we note $\underline{t} = (t_1^T, \dots, t_p^T)^T \in \mathbb{R}^{q_1} \times \dots \times \mathbb{R}^{q_p}$ to denote a column vector of length $q := \sum_{\ell=1}^p q_\ell$.

We want to test the mutual independence of the p sub-components of \underline{x} . This is equivalent to test the null hypothesis

$$\mathcal{H}_0: \varphi(\underline{t}) = \prod_{\ell=1}^p \varphi_\ell(t_\ell), \quad \forall \underline{t} \in \mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_p},$$

where $\varphi(\underline{t})$ is the true (unknown) characteristic function of \underline{x} at point \underline{t} and $\varphi_{\ell}(t_{\ell})$ is the true (unknown) characteristic function of \mathbf{x}^{ℓ} at point t_{ℓ} , $\ell = 1, \ldots, p$.

A natural nonparametric empirical estimator of $\varphi(\underline{t})$ is

$$\hat{\varphi}_n\left(\underline{t}\right) = n^{-1} \sum_{j=1}^n e^{\underline{i}\underline{t}^{\mathsf{T}}\underline{\mathbf{X}}_j} = n^{-1} \sum_{j=1}^n \prod_{\ell=1}^p e^{it_\ell^{\mathsf{T}}\mathbf{X}_\ell^{\mathsf{T}}}.$$

Similarly, a nonparametric estimator of $\varphi_\ell(t_\ell)$ is given by

$$\hat{\varphi}_{n,\ell}(\boldsymbol{t}_{\ell}) = n^{-1} \sum_{j=1}^{n} e^{i\boldsymbol{t}_{\ell}^{\mathsf{T}} \mathbf{X}_{j}^{\ell}}.$$

The test statistics (location+orthogonal invariant) we propose for testing \mathcal{H}_0 are

$$n\mathsf{T}_{n}(w) = \left\| \sqrt{n}\mathsf{D}_{n} \right\|_{w}^{2} = n \int_{\mathbb{R}^{q_{1}} \times \cdots \times \mathbb{R}^{q_{p}}} \left| \mathsf{D}_{n}\left(\underline{t}\right) \right|^{2} w\left(\underline{t}\right) d\underline{t},$$

where

$$\mathsf{D}_n\left(\underline{t}\right) = \hat{\varphi}_n\left(\underline{t}\right) - \prod_{\ell=1}^p \hat{\varphi}_{n,\ell}(\boldsymbol{t}_\ell) \in \mathbb{C}$$

and where $w(\underline{t}) = \prod_{\ell=1}^{p} v(t_{\ell})$ are some properly chosen weight functions.

We obtained the following closed-form expression for our new measure of multivariate dependence

$$nT_{n}(w) := \int_{\mathbb{R}^{q_{1}}\times\cdots\times\mathbb{R}^{q_{p}}} \left| \sqrt{n} \left(\hat{\varphi}_{n}\left(\underline{t} \right) - \prod_{\ell=1}^{p} \hat{\varphi}_{n,\ell}(t_{\ell}) \right) \right|^{2} \prod_{\ell=1}^{p} v(t_{\ell}) dt$$
$$= n^{-1} \sum_{j,j'=1}^{n} \prod_{\ell=1}^{p} \xi_{j,j',\ell} - 2 \left\{ \sum_{j=1}^{n} \prod_{\ell=1}^{p} \left[n^{-1} \sum_{j'=1}^{n} \xi_{j,j',\ell} \right] \right\} + n \prod_{\ell=1}^{p} \left[n^{-2} \sum_{j,j'=1}^{n} \xi_{j,j',\ell} \right]$$

where

$$\xi_{j,j',\ell} = \int_{\mathbb{R}^{q_\ell}} \cos\left(t_\ell^\mathsf{T}\left(\mathsf{x}_j^\ell - \mathsf{x}_{j'}^\ell\right)\right) v(t_\ell) dt_\ell \in \mathbb{R}.$$

We also worked on other choices for the weight function with which we also obtained closed form expression. (Second improvement over the distance correlation test which considered only one specific weight function)

To construct the rejection region of the test, it would be great to know the (asymptotic) null distribution of $nT_n(w)$. We have done this in two steps.

Step 1.

Theorem.1

If $\mathbf{x}^1, \ldots, \mathbf{x}^p$ are mutually independent, then $\sqrt{n} D_n = \sqrt{n} \left(\hat{\varphi}_n - \prod_{\ell=1}^p \hat{\varphi}_{n,\ell} \right)$ converges (weak convergence of processes) in $C(\mathbb{R}^q, \mathbb{C})$ to a zero mean complex Gaussian process D having covariance function given by

$$C(\underline{s},\underline{t}) = \mathbb{E}\left[\mathsf{D}(\underline{s})\overline{\mathsf{D}}(\underline{t})\right]$$
$$= \prod_{\ell=1}^{p} \varphi_{\ell}(-\underline{t}_{\ell} + \underline{s}_{\ell}) - \prod_{\ell=1}^{p} \left[\varphi_{\ell}(-\underline{t}_{\ell})\varphi_{\ell}(\underline{s}_{\ell})\right] \left\{1 - p + \sum_{\ell=1}^{p} \frac{\varphi_{\ell}(-\underline{t}_{\ell} + \underline{s}_{\ell})}{\varphi_{\ell}(-\underline{t}_{\ell})\varphi_{\ell}(\underline{s}_{\ell})}\right\}$$

and pseudo-covariance function given by

$$P(\underline{s},\underline{t}) = \mathbb{E}\left[\mathsf{D}(\underline{s})\mathsf{D}(\underline{t})\right]$$
$$= \prod_{\ell=1}^{p} \varphi_{\ell}(t_{\ell} + s_{\ell}) - \prod_{\ell=1}^{p} \left[\varphi_{\ell}(t_{\ell})\varphi_{\ell}(s_{\ell})\right] \left\{1 - p + \sum_{\ell=1}^{p} \frac{\varphi_{\ell}(t_{\ell} + s_{\ell})}{\varphi_{\ell}(t_{\ell})\varphi_{\ell}(s_{\ell})}\right\}.$$

Sketch of proof: To prove that $\sqrt{n}D_n$ converges in $C(\mathbb{R}^q, \mathbb{C})$ to a complex Gaussian process D, we use

Proposition .1 (Proposition 14.6 of [Kallenberg, 2002])

Let X, X_1, X_2, \ldots be random elements in C(T, S) (endowed with the topology of uniform convergence on compacts), where S is a metric space and T is locally compact, second-countable, and Hausdorff (abbreviated as lcscH). Then $X_n \xrightarrow{d} X$ iff the convergence holds for the restrictions to arbitrary compact subsets $K \subset T$.

The metric ρ of uniform convergence on all compact sets on $C(\mathbb{R}^q,\mathbb{C})$ is defined by

$$\rho(x,y) := \sum_{j=1}^{\infty} 2^{-j} \frac{\rho_j(x,y)}{1 + \rho_j(x,y)}, \quad \text{where } \rho_j(x,y) = \sup_{\|\underline{t}\| \le j} |x(\underline{t}) - y(\underline{t})|$$

is the usual sup norm.

Sketch of proof (cont'd) : It thus suffices to show that $r_j(\sqrt{n}D_n) \xrightarrow{d} r_j(D)$ in order to show the convergence $\sqrt{n}D_n \xrightarrow{d} D$, where, for all mapping $x \in C(\mathbb{R}^q, \mathbb{C}), r_j(x)$ is the restriction of x to the ball \mathcal{B}_i^q of radius j.

We prove a straightforward generalization of [Csörgo, 1985, Theorem p. 294], who worked on the same processes as we did (defined for random variables, not random vectors though) but restricted to a compact set.

We also need his (technical) condition (*) to be satisfied (mild tail condition on the underlying joint distribution).

It then remains to compute the moments of the limiting Gaussian, which is not too difficult.

Theorem.2

We have

$$\begin{split} nT_n(w) &= \|\sqrt{n}\mathsf{D}_n\|_w^2 = \int_{\mathbb{R}^{q_1}\times\cdots\times\mathbb{R}^{q_p}} |\sqrt{n}\mathsf{D}_n\left(\underline{t}\right)|^2 w\left(\underline{t}\right) d\underline{t} \stackrel{L}{\longrightarrow} \sum_{k=1}^{\infty} \frac{\lambda_k}{2} \Big((1+|p_k|)\xi_k \\ &+ (1-|p_k|)\eta_k \Big) \end{split}$$

Here $\xi_k, \eta_k, k = 1, 2, ..., m, ...$ are independent pairs of independent chisquare random variables, each with one degree of freedom and λ_k (respectively f_k) are the eigenvalues (respectively eigenfunctions) of the integral operator O defined by

$$O(f)(\underline{\mathbf{x}}) = \int_{\mathbb{R}^q} f(\underline{\mathbf{y}}) C(\underline{\mathbf{x}}, \underline{\mathbf{y}}) w(\underline{\mathbf{y}}) d\underline{\mathbf{y}},$$

The values p_k are defined as

$$p_k = \int f_k(\underline{s}) f_k(-\underline{s}) w(\underline{s}) d\underline{s}.$$

Sketch of proof: We use the nice properties of our limiting covariance function $C(\cdot, \cdot)$ (hermitian, positive definite) and a generalization [Ferreira and Menegatto, 2012] (to \mathbb{R}^q) of the Mercer's theorem which states that **there exists** a sequence of non-negative real numbers $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ and an orthonormal (with respect to $\langle \cdot, \cdot \rangle_w$) family $\{f_k\}$ such that

$$C(\underline{s},\underline{t}) = \sum_{k=1}^{\infty} \lambda_k f_k(\underline{s}) \overline{f_k(\underline{t})},$$

where $\langle f, g \rangle_w = \int f(\underline{t}) \overline{g(\underline{t})} w(\underline{t}) d\underline{t}$. This enables us to write :

$$\int C\left(\underline{s},\underline{t}\right) f_{j}\left(\underline{t}\right) w\left(\underline{t}\right) d\underline{t} = \int \sum_{k=1}^{\infty} \lambda_{k} f_{k}(\underline{s}) \overline{f_{k}\left(\underline{t}\right)} f_{j}\left(\underline{t}\right) w\left(\underline{t}\right) d\underline{t}$$
$$= \sum_{k=1}^{\infty} \lambda_{k} f_{k}(\underline{s}) \int f_{j}\left(\underline{t}\right) \overline{f_{k}\left(\underline{t}\right)} w\left(\underline{t}\right) d\underline{t}$$
$$= \sum_{k=1}^{\infty} \lambda_{k} f_{k}(\underline{s}) \delta_{jk} = \lambda_{j} f_{j}(\underline{s}).$$

 \Rightarrow They are solutions of the integral operator O.

ASC 2016

Sketch of proof (cont'd) : Now, an expansion of the complex Gaussian process D on our $\{f_k\}$ orthonormal basis gives :

$$D\left(\underline{t}\right) = \sum_{k=1}^{\infty} \langle D, f_k \rangle f_k(\underline{t}).$$

So

$$\int |\mathsf{D}(\underline{t})|^2 w(\underline{t}) d\underline{t} = \int \sum_{k=1}^{\infty} \sum_{k'=1}^{\infty} \langle \mathsf{D}, f_k \rangle \overline{\langle \mathsf{D}, f_{k'} \rangle} f_k(\underline{t}) \overline{f_{k'}(\underline{t})} d\underline{t} = \sum_{k=1}^{\infty} |\langle \mathsf{D}, f_k \rangle|^2$$

where

$$\langle \mathsf{D}, f_k \rangle = \int \mathsf{D}(\underline{t}) \overline{f_k(\underline{t})} w(\underline{t}) d\underline{t} \sim CN_1(0, \lambda_k, \lambda_k \overline{p}_k)$$

(the complex normal distribution) where $p_k = \int f_k(\underline{s}) f_k(-\underline{s}) w(\underline{s}) d\underline{s}$ and where

$$|\langle \mathsf{D}, f_k \rangle|^2 \sim \frac{\lambda_k}{2} \left((1 + |p_k|)\chi_1^2 + (1 - |p_k|)\chi_1^2 \right)$$

from [Ducharme et al., 2016, Theorem 3.3].

We have thus obtained the (asymptotic) null distribution of our test statistic $nT_n(w)$: infinite weighted sum of independent chi-square r.v.

Unfortunately, since

$$C\left(\underline{s},\underline{t}\right) = \prod_{\ell=1}^{p} \varphi_{\ell}(-t_{\ell} + s_{\ell}) - \prod_{\ell=1}^{p} \left[\varphi_{\ell}(-t_{\ell})\varphi_{\ell}(s_{\ell})\right] \left\{1 - p + \sum_{\ell=1}^{p} \frac{\varphi_{\ell}(-t_{\ell} + s_{\ell})}{\varphi_{\ell}(-t_{\ell})\varphi_{\ell}(s_{\ell})}\right\}$$

is unknown (depends on the unknown marginals $\varphi_{\ell}(\cdot)$), then **the weights** λ_k 's, which are solutions (together with the f_k 's) of

$$\lambda f(\underline{s}) = \int_{\mathbb{R}^q} f(\underline{t}) C(\underline{s}, \underline{t}) w(\underline{t}) d\underline{t}$$

are also unknown and one cannot compute critical values for the test (quantiles of the distribution $\sum_{k=1}^{\infty} \frac{\lambda_k}{2} \left((1+|p_k|)\chi_1^2 + (1-|p_k|)\chi_1^2 \right)$, using the CompQuadForm **R** package)

 \Rightarrow We need to (This will be the third improvement over the distance correlation test)

- estimate these λ_k 's:
- show that the use of these estimates $\hat{\lambda}_{nk}$ has no (asymptotic) impact.

Theorem.3

For all $K \ge 1$, we have

$$\sum_{k=1}^{K} \frac{\lambda_{k}}{2} ((1+|p_{k}|)\xi_{k} + (1-|p_{k}|)\eta_{k}) - \sum_{k=1}^{K} \frac{\hat{\lambda}_{k}}{2} ((1+|\hat{p}_{k}|)\xi_{k} + (1-|\hat{p}_{k}|)\eta_{k}) \xrightarrow{P} 0,$$

where the $\hat{\lambda}_k$ (and $\hat{f}_k)$ are the solutions of a discretized version of

$$\lambda f(\underline{\mathbf{x}}) = \int_{\mathbb{R}^q} f(\mathbf{y}) C(\underline{\mathbf{x}}, \underline{\mathbf{y}}) w(\underline{\mathbf{y}}) d\underline{\mathbf{y}}$$
(1)

where the unknown $C(\underline{x}, \underline{y})$ is replaced with

$$\hat{C}_n(\underline{\mathbf{x}},\underline{\mathbf{y}}) = \prod_{\ell=1}^p \hat{\varphi}_{n,\ell}(-\mathbf{x}_\ell + \mathbf{y}_\ell) - \prod_{\ell=1}^p \left[\hat{\varphi}_{n,\ell}(-\mathbf{x}_\ell) \hat{\varphi}_{n,\ell}(\mathbf{y}_\ell) \right] \left\{ 1 - p + \sum_{\ell=1}^p \frac{\hat{\varphi}_{n,\ell}(-\mathbf{x}_\ell + \mathbf{y}_\ell)}{\hat{\varphi}_{n,\ell}(-\mathbf{x}_\ell) \hat{\varphi}_{n,\ell}(\mathbf{y}_\ell)} \right\}$$

Sketch of proof : One can write

$$\sum_{j=1}^{N} \omega_j f(\underline{\mathbf{y}}_j) C(\underline{\mathbf{x}}, \underline{\mathbf{y}}_j) \xrightarrow[N \to \infty]{} \int_{\mathbb{R}^d} f(\underline{\mathbf{y}}) C(\underline{\mathbf{x}}, \underline{\mathbf{y}}) w(\underline{\mathbf{y}}) d\underline{\mathbf{y}} = \lambda f(\underline{\mathbf{x}}),$$

where the ω_j 's are the weights of some cubature and where the \underline{y}_j 's are nodes of the cubature. In the sequel, we suppose that $(\omega_j, \underline{y}_j)$ can also be given as $(1/N, \underline{y}_j)$ (i.e., a Monte-Carlo integration). So now we can write

$$\sum_{j=1}^{N} \omega_j f(\underline{\mathbf{y}}_j) C(\underline{\mathbf{y}}_i, \underline{\mathbf{y}}_j) \approx \lambda f(\underline{\mathbf{y}}_i) \qquad i = 1, \dots, N.$$

This is the discretized version of the integral operator problem.

Sketch of proof (cont'd) : Let $f_j = f(\underline{y}_j)$, j = 1, ..., N and $f = (f_1, ..., f_N)^T$. The system above can be written in matrix form as

$$Cf = \lambda f \tag{2}$$

where C is the $N \times N$ (possibly random) matrix $\left(C(\underline{y}_i, \underline{y}_j)\right)_{i,j}$.

Using results in [Rosasco et al., 2010], we can prove that the solutions (eigenvalues and eigenvectors) of (2) are good approximations (when N gets large) of those of (1) to estimate the p_k .

Sketch of proof (cont'd) : But, since *C* is unknown, we replace problem (2) with

$$\widehat{C}_n f = \lambda f \tag{3}$$

where \widehat{C}_n is the $N \times N$ matrix $(\widehat{C}_n(\underline{y}_i, \underline{y}_j))_{i,j}$. Notice *n* is not *N*.

Since \hat{C}_n is a strongly consistent estimator of C, we have that each element of \widehat{C}_n converges (w.p.1, $n \to \infty$) to the corresponding element of C.

We can then use our Theorem 4 on next slide to prove that, for each N fixed, the solutions of (3) converge (w.p.1, $n \to \infty$) to the solutions of (2).

Note : Even if N is large, we need only to compute the first largest eigenvalues of the above Hermitian matrix (this is possible using the zhpevx Lapack Fortran subroutine).

Theorem.4

Let $N \in \mathbb{N}^*$ be a fixed integer. Let $A_n = (a_{ij,n})$ be a sequence of $(N \times N)$ stochastic complex matrices defined on some sample space Ω , with (possibly complex) eigenvalues $\lambda_{(1),n}, \dots, \lambda_{(N),n}$, counting mutiplicity, ordered by the increasing value of their modulus. Let $A = (a_{ij})$ be a $(N \times N)$ positive definite stochastic complex matrix defined on Ω with (real positive) eigenvalues $0 < \lambda_{(1)} \leq \dots \leq \lambda_{(N)}$, counting multiplicity. We note $f_{k,n}$ and f_k , $k = 1, \dots, N$, the associated eigenvectors. If $A_n \xrightarrow{a.s.} A$, i.e.

$$\forall i = 1, \dots, N, \ \forall j = 1, \dots, N, \ a_{ij,n} \xrightarrow{a.s.} a_{ij}, \qquad \text{when } n \to \infty,$$

then

$$\lambda_{(k),n} \xrightarrow{a.s.} \lambda_{(k)}, \quad \text{and} \quad \|(I_N - P_{A,k})\boldsymbol{f}_{k,n}\|_2 \xrightarrow{a.s.} 0, \quad \forall k = 1, \dots, N,$$

where $P_{A,k}$ is the orthogonal projector on the space spanned by all the orthonormal eigenvectors of A associated with λ_k .

Sketch of Proof : Use characteristic polynomials and a general form of Rouché's Theorem in complex analysis (e.g., [Rudin, 1987, p. 229]).

We implemented this test in the **R** package IndependenceTests.

Let us apply it on our data set.

```
> require("IndependenceTests")
> X <- read.table("world data.txt", header = TRUE) ; attach(X)</pre>
> X <- as.matrix(X) ; n <- nrow(X)</pre>
> head(X[,1:3])
     mortality natality qdp.capita
                25.47
[1,]
        14.89
                              3000
[2,]
        14.46
                10.72
                             8000
[3,]
       14.44
                8.92
                             18400
[4,] 14.33
               33.38
                             1500
[5.1
       14.31
                10.00
                             24500
[6,]
       14.28
               36.60
                              2800
> set.seed(1) ; res <- mdcov(X, vecd = c(1, 1), a = 1, weight.choice = 1, N = 300, K = 200)</pre>
> res$mdcov
[1] 3.505156
> res$pvalue
[1] 5.66429e-10 # We reject the pairwise independence
> set.seed(1) : res <- mdcov(X, vecd = c(1, 1, 1), a = 1, weight.choice = 1, N = 300, K = 200)</pre>
> res$mdcov
[1] 1.029831
> res$pvalue
[1] 0.00016071 # We reject the threewise independence
```

Note : our procedure generalizes the distance covariance.

A few perspectives

- Perform simulations to evaluate the power of this new test procedure against some competitors.
- ► Study optimality/efficiency, e.g. using results in [Nikitin, 1995].
- Build a test of multivariate serial dependence.

Thank you for your attention !



Ē	Bakirov, N. K., Rizzo, M. L., and Székely, G. J. (2006).	Positivity,
	A multivariate nonparametric test of independence. Journal of Multivariate Analysis, 97(8) :1742–1756.	Kallenberg
	Csörgo, S. (1985).	Foundation Springer, 2
	Testing for independence by the empirical characteristic function.	Nikitin, Y.
	Journal of Multivariate Analysis, 16 :290–299.	Asymptotic Cambridge
	Ducharme, G., Lafaye de Micheaux, P., and Marchina, B. (2016).	Rosasco, L
	The complex multinormal distribution, quadratic forms in complex random vectors and an omnibus goodness-of-fit	On learning Journal of
_	Annals of the Institute of Statistical Mathematics, 68(1).	Rudin, W.
	Fan, Y., Lafaye de Micheaux, P., Penev, S., and Salopek,	Real and C McGraw-H
	D. (2017). Multivariate nonparametric test of independence. Journal of Multivariate Analysis, 153 :189–210.	Székely, G. Measuring
	Ferreira, J. and Menegatto, V. (2012).	tances. The Annal

. ion of Mercer's theory to L^p . 16(2):197-212.

, O. (2002).

ns of Modern Probability. nd edition.

(1995).

c Efficiency of Nonparametric Tests. University Press.

., Belkin, M., and De Vito, E. (2010).

g with integral operators. Machine Learning Research, 11 :905-934.

(1987).

Complex Analysis. HH.

., Rizzo, M., and Bakirov, N. (2007).

and testing dependence by correlation of dis-

Is of Statistics, 35(6) :2769-2794.