# *Measuring and testing mutual dependence for functional data*

Tomasz Górecki, **Mirosław Krzyśko**, Waldemar Wołyński

Faculty of Mathematics and Computer Science
Adam Mickiewicz University, Poznań, Poland

The XLV annual conference "Statystyka Matematyczna"
Będlewo, 2 - 6 December 2019

# Outline

# Introduction

Let's consider the problem of testing mutual independence for random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K$, $K \geq 2$, $\boldsymbol{X}_i \in L_2^{p_i}(I)$, $i = 1, \ldots, K$, i.e. we are testing the following hypothesis:

$$H_0: \boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K$$

vs

$$H_1: \neg H_0.$$

This problem is connected with examining the significance of the coefficient of mutual dependency.

In our presentation we will present methods of measuring mutual independence for functional data. We will consider two or more vector random processes. Based on the obtained measures we will construct permutational tests to test mutual independence for functional data.

Let us assume that $\boldsymbol{X} \in L_2^p(I)$ is a random process, where $L_2(I)$ is a Hilbert space of square integrable functions on the interval $I$. Additionally, we also assume that

$$\mathrm{E}(\boldsymbol{X}(t)) = \mathbf{0}, \quad t \in I.$$

# Smoothing

We will further assume that each component $X_g$ of the process $\boldsymbol{X}$ can be represented by a finite number of basis functions $\{\varphi_e\}$:

$$X_g(t) = \sum_{e=0}^{B_g} \alpha_{ge}\varphi_e(t), s \in I, g = 1, 2, ..., p.$$

The degree of smoothness of function $X_g$ depends on the value $B_g$ (a small value causes more smoothing of the function).

# Smoothing

We introduce the following notation:

$$\boldsymbol{\alpha} = (\alpha_{10}, ..., \alpha_{1B_1}, ..., \alpha_{p0}, ..., \alpha_{pB_p})^\top,$$

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \boldsymbol{\varphi}_1^\top(t) & \mathbf{0}^\top & \dots & \mathbf{0}^\top \\ \mathbf{0}^\top & \boldsymbol{\varphi}_2^\top(t) & \dots & \mathbf{0}^\top \\ \dots & \dots & \dots & \dots \\ \mathbf{0}^\top & \mathbf{0}^\top & \dots & \boldsymbol{\varphi}_p^\top(t) \end{bmatrix},$$

where $\boldsymbol{\varphi}_k(t) = (\varphi_0(t), \varphi_1(t), \dots, \varphi_{B_k}(t))^\top$, $k = 1, 2, \dots, p$ are orthonormal basis functions of space $L_2(I)$.
Using this matrix notation the random process $\boldsymbol{X}$ can be represented as

$$\mathbf{X}(t) = \boldsymbol{\Phi}(t)\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{K+p}$, $K = B_1 + B_2 + \dots + B_p$.

# Notation

Let $\boldsymbol{X} \in L_2^p(I)$ i $\boldsymbol{Y} \in L_2^q(I)$ be random processes with the following representations:

$$\mathbf{X}(t) = \boldsymbol{\Phi}_1(t)\boldsymbol{\alpha}, \quad \mathbf{Y}(s) = \boldsymbol{\Phi}_2(s)\boldsymbol{\beta}, \quad t, s \in I.$$

Additionally, let the covariance matrix of processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the form

$$\boldsymbol{\Sigma}(t,s) = \begin{bmatrix} \boldsymbol{\Sigma}_{XX}(t,s) & \boldsymbol{\Sigma}_{XY}(t,s) \\ \boldsymbol{\Sigma}_{YX}(t,s) & \boldsymbol{\Sigma}_{YY}(t,s) \end{bmatrix}, \ t, s \in I.$$

Then

$$\begin{aligned} \boldsymbol{\Sigma}_{XX}(t,s) &= \boldsymbol{\Phi}_1(t)\boldsymbol{\Sigma}_{\alpha\alpha}\boldsymbol{\Phi}_1^\top(s), \\ \boldsymbol{\Sigma}_{XY}(t,s) &= \boldsymbol{\Phi}_1(t)\boldsymbol{\Sigma}_{\alpha\beta}\boldsymbol{\Phi}_2^\top(s), \\ \boldsymbol{\Sigma}_{YY}(t,s) &= \boldsymbol{\Phi}_2(t)\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\Phi}_2^\top(s), \end{aligned}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\alpha\alpha} & \boldsymbol{\Sigma}_{\alpha\beta} \\ \boldsymbol{\Sigma}_{\beta\alpha} & \boldsymbol{\Sigma}_{\beta\beta} \end{bmatrix}$$

is a covariance matrix of vectors $\boldsymbol{\alpha}$ i $\boldsymbol{\beta}$.

# Functional coefficient $\rho V$

Functional coefficient $\rho V$ is a nonnegative number given by

$$\rho V_{\boldsymbol{X},\boldsymbol{Y}} = \frac{\|\boldsymbol{\Sigma}_{XY}\|_F}{\sqrt{\|\boldsymbol{\Sigma}_{XX}\|_F \|\boldsymbol{\Sigma}_{YY}\|_F}},$$

where

$$\|\boldsymbol{\Sigma}_{XY}\|_F = \sqrt{\int_I \int_I \mathrm{tr}(\boldsymbol{\Sigma}_{XY}^{\top}(t,s)\boldsymbol{\Sigma}_{XY}(t,s))dtds}.$$

# Functional coefficient $\rho V$

If processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the form

$$\mathbf{X}(t) = \mathbf{\Phi}_1(t)\boldsymbol{\alpha}, \quad \mathbf{Y}(s) = \mathbf{\Phi}_2(s)\boldsymbol{\beta}, \quad t, s \in I,$$

then

$$\rho V_{\boldsymbol{X},\boldsymbol{Y}} = \rho V_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Coefficient $\rho V$ for random vectors was proposed by Escoufier (1973).

*Escoufier, G.Y. (1973): Le traitement des variables vectorielles, Biometrics 29, 751-760.*

# Functional coefficient $\rho V$

Let's assume that the joint distribution of random vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is $p+q$ dimensional normal distribution.

Problem of testing the null hypothesis

$$H_0: \boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}$$

is equivalent to the problem of testing the null hypothesis

$$H_0: \rho V_{\boldsymbol{X},\boldsymbol{Y}} = 0.$$

which is equivalent to the problem of testing the null hypothesis

$$H_0: \rho V_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0$$

for a pair of random vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. To verify this hypothesis we use a permutational test.

# Functional distance correlation

For random vectors $\boldsymbol{\alpha} \in \mathbb{R}^p$ and $\boldsymbol{\beta} \in \mathbb{R}^q$ Székely et al. (2007) introduced the distance correlation (dCorr) as a nonnegative number given by

$$\mathrm{dCorr}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{\mathrm{dCov}(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\sqrt{\mathrm{dCov}(\boldsymbol{\alpha}, \boldsymbol{\alpha})\,\mathrm{dCov}(\boldsymbol{\beta}, \boldsymbol{\beta})}},$$

where

$$\mathrm{dCov}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \|\phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\lambda}, \boldsymbol{\mu}) - \phi_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})\phi_{\boldsymbol{\beta}}(\boldsymbol{\mu})\|_w,$$

and

$$\|f\|_w = \sqrt{\iint |f(\boldsymbol{\lambda}, \boldsymbol{\mu})|^2 w(\boldsymbol{\lambda}, \boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu}}.$$

*Székely, G.J., Rizzo, M.L., Bakirov, N.K. (2007): Measuring and testing dependence by correlation of distances, The Annals of Statistics 35, 2769-2794.*

# Functional distance correlation

For jointly distributed random processes $\boldsymbol{X} \in L_2^p(I)$ and $\boldsymbol{Y} \in L_2^q(I)$, let

$$\phi_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{l},\boldsymbol{m}) = \mathrm{E}\{\exp[i < \boldsymbol{l},\boldsymbol{X} > + i < \boldsymbol{m},\boldsymbol{Y} >]\}$$

be the joint characteristic function of $(\boldsymbol{X},\boldsymbol{Y})$. If processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the form

$$\mathbf{X}(t) = \boldsymbol{\Phi}_1(t)\boldsymbol{\alpha}, \quad \mathbf{Y}(s) = \boldsymbol{\Phi}_2(s)\boldsymbol{\beta}, \quad t,s \in I,$$

and if

$$\boldsymbol{l}(s) = \boldsymbol{\Phi}_1(s)\boldsymbol{\lambda}, \quad \boldsymbol{m}(t) = \boldsymbol{\Phi}_2(t)\boldsymbol{\mu},$$

then

$$\phi_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{l},\boldsymbol{m}) = \mathrm{E}\{\exp[i\boldsymbol{\lambda}'\boldsymbol{\alpha} + i\boldsymbol{\mu}'\boldsymbol{\beta}]\} = \phi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\lambda},\boldsymbol{\mu}).$$

# Functional distance correlation

We define the functional distance correlation in the following way

$$\mathrm{dCorr}(\boldsymbol{X}, \boldsymbol{Y}) = \mathrm{dCorr}(\boldsymbol{\alpha}, \boldsymbol{\beta}).$$

*Górecki T., Krzyśko M., Wołyński W. (2017): Correlation analysis for multivariate functional data. In: Data Science, Studies in Classification, Data Analysis, and Knowledge Organization, Springer International Publishing, 243-258.*

*Górecki T., Krzyśko M., Wołyński W. Independence test and canonical correlation analysis based on the alignment between kernel matrices for multivariate functional data. Artifical Inteligence Review (in press).*

# Functional distance correlation

Problem of testing the null hypothesis

$$H_0: \boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}$$

is equivalent to the problem of testing the null hypothesis

$$H_0: \mathrm{dCorr}(\boldsymbol{X}, \boldsymbol{Y}) = 0,$$

which is equivalent to the problem of testing the null hypothesis

$$H_0: \mathrm{dCorr}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$$

for a pair of random vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. To verify this hypothesis we use a permutational test.

## $K > 2$ case

Let $\boldsymbol{X}_1 \in L_2^{p_1}(I), \boldsymbol{X}_2 \in L_2^{p_2}(I), \ldots, \boldsymbol{X}_K \in L_2^{p_K}(I)$ be random processes with the following representation:

$$\mathbf{X}_1(t) = \boldsymbol{\Phi}_1(t)\boldsymbol{\alpha}_1, \mathbf{X}_2(t) = \boldsymbol{\Phi}_2(t)\boldsymbol{\alpha}_2, \ldots, \mathbf{X}_K(t) = \boldsymbol{\Phi}_K(t)\boldsymbol{\alpha}_K, \ t \in I.$$

Additionally, let the matrix of covariance for vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_K$ has the form:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \cdots & \boldsymbol{\Sigma}_{1K} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \cdots & \boldsymbol{\Sigma}_{2K} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\Sigma}_{K1} & \boldsymbol{\Sigma}_{K2} & \cdots & \boldsymbol{\Sigma}_{KK} \end{bmatrix}.$$

Assuming joint $p_1 + p_2 + \cdots + p_k$ dimensional normal distribution of vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_K$, the problem of testing the null hypothesis

$$H_0: \boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K$$

is equivalent to the problem of testing the null hypothesis

$$H_0: \sum_{i<j} \|\boldsymbol{\Sigma}_{ij}\|_F = 0.$$

# Coefficient of mutual correlation $\rho MV$

We define coefficient of mutual correlation $\rho MV$ as a positive number given by

$$\rho^2 MV = \frac{2}{K(K-1)} \sum_{i<j} \rho^2 V(\boldsymbol{X}_i, \boldsymbol{X}_j).$$

Assuming joint $p_1 + p_2 + \cdots + p_k$ dimensional normal distribution of vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_K$, the problem of testing the null hypothesis

$$H_0: \boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K$$

is equivalent to the problem of testing the null hypothesis

$$H_0: \rho MV = 0.$$

# Measures of multiple independence

Let $\mathrm{Corr}(\boldsymbol{X}, \boldsymbol{Y})$ be some measure of dependence for random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ with property: $\mathrm{Corr}(\boldsymbol{X}, \boldsymbol{Y}) = 0$ if and only if random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent.

Let $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K)$, and

$$\boldsymbol{X}_{c+} = (\boldsymbol{X}_{c+1}^\top, \ldots, \boldsymbol{X}_K^\top)^\top, \ c = 1, \ldots, K-1,$$

$$\boldsymbol{X}_{c-} = (\boldsymbol{X}_1^\top, \ldots, \boldsymbol{X}_{c-1}^\top, \boldsymbol{X}_{c+1}^\top, \ldots, \boldsymbol{X}_K^\top)^\top, \ c = 1, \ldots, K.$$

We define the coefficients of multiple independence as

$$\mathcal{R}(\boldsymbol{X}) = \frac{1}{K-1} \sum_{c=1}^{K-1} \mathrm{Corr}^2(\boldsymbol{X}_c, \boldsymbol{X}_{c+}),$$

and

$$\mathcal{S}(\boldsymbol{X}) = \frac{1}{K} \sum_{c=1}^{K} \mathrm{Corr}^2(\boldsymbol{X}_c, \boldsymbol{X}_{c-}).$$

*Theorem*

$$\boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K \iff \mathcal{R}(\boldsymbol{X}) = 0,$$

$$\boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K \iff \mathcal{S}(\boldsymbol{X}) = 0.$$

*Remark*

*In the place of* Corr *we can put eg.* $\rho V$ *or* dCorr.

Hence, the problem of testing the null hypothesis

$$H_0: \boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp \boldsymbol{X}_K$$

is equivalent to the problem of testing the null hypothesis

$$H_0: \mathcal{R}(\boldsymbol{X}) = 0 \ (\mathcal{S}(\boldsymbol{X}) = 0).$$

# Example

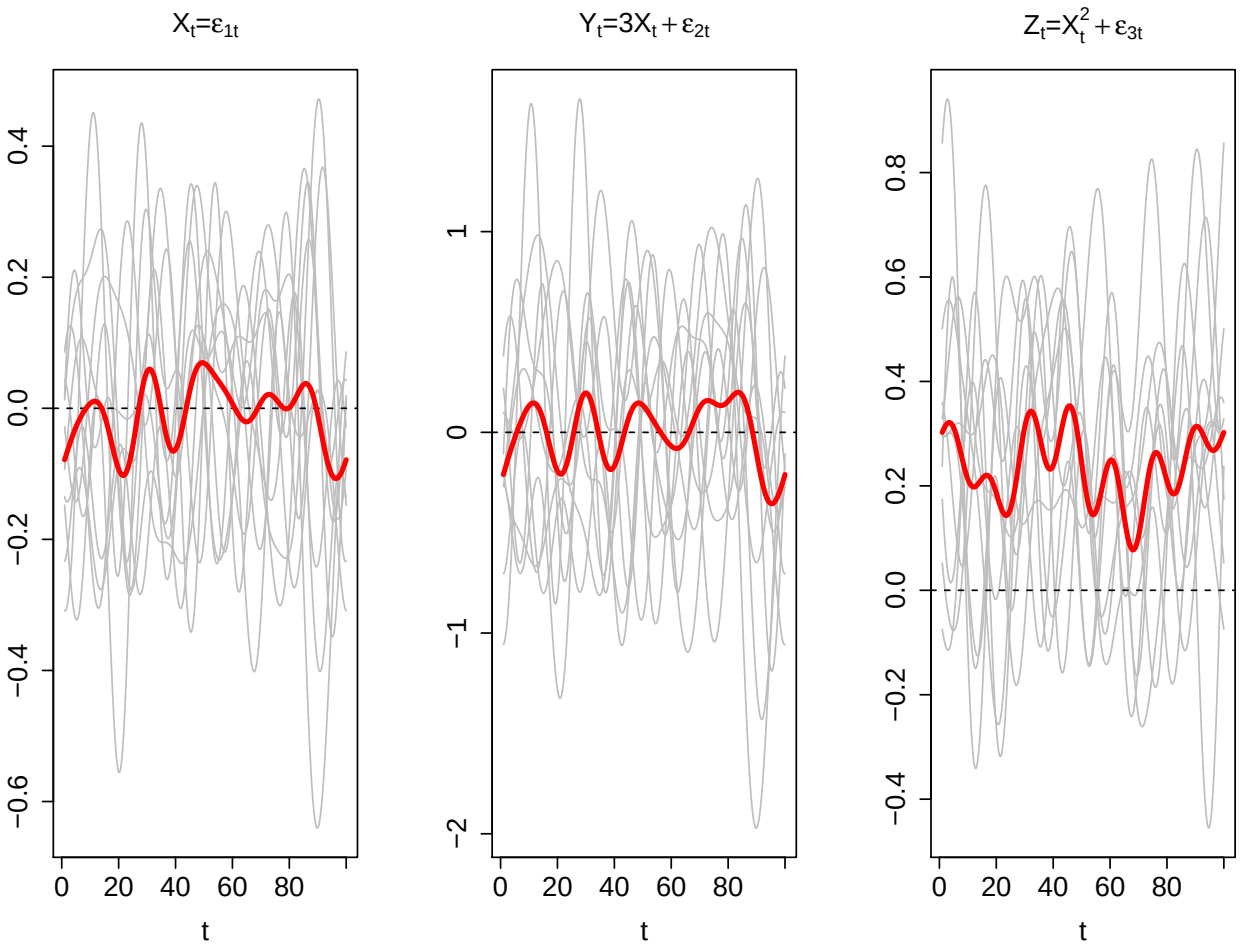
Let

$$X_t = \varepsilon_{1t},$$
$$Y_t = 3X_t + \varepsilon_{2t},$$
$$Z_t = X_t^2 + \varepsilon_{3t},$$

where $\varepsilon_{1t}$, $\varepsilon_{2t}$ and $\varepsilon_{3t}$ are independent random variables with $N(0, 0.25)$ distribution. We generated 1000 random realizations for each process with length 100. To smooth the data we used Fourier series with 15 elements.

Processes $X_t$ and $Y_t$ are linearly dependent. Processes $X_t$, $Z_t$ and $Y_t$, $Z_t$ are non-linearly dependent.

# Example




10 randomly selected realizations of processes $X_t$, $Y_t$ and $Z_t$ (functional means in red).

# Example

| Coefficient | Processes | $p$-value |
|---|---|---|
| $\rho MV$ | $X_t, Y_t, Z_t$ | 0.014 |
| $\mathcal{R} - \rho V$ | $X_t, Y_t, Z_t$ | 0.006 |
| $\mathcal{S} - \rho V$ | $X_t, Y_t, Z_t$ | 0.027 |
| $\mathcal{R}$ – dCorr | $X_t, Y_t, Z_t$ | 0.007 |
| $\mathcal{S}$ – dCorr | $X_t, Y_t, Z_t$ | 0.016 |
| $\rho V$ | $X_t$ vs $Y_t$ | 0.001 |
| | $X_t$ vs $Z_t$ | 0.367 |
| | $Y_t$ vs $Z_t$ | 0.481 |
| dCorr | $X_t$ vs $Y_t$ | 0.001 |
| | $X_t$ vs $Z_t$ | 0.003 |
| | $Y_t$ vs $Z_t$ | 0.025 |