UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 13-33
Statistics and Econometrics Series 29
December 2013

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

SPEARMAN COEFFICIENT FOR FUNCTIONS

Dalia Valencia[a], Rosa E. Lillo[b], and Juan Romo[b]

[a]Departamento de Estadística, Universidad Carlos III de Madrid, 28911, Leganés, Spain
[b] Departamento de Estadística, Universidad Carlos III de Madrid, 28903, Getafe, Spain

Abstract----------
We present a notion of Spearman's coefficient for functional data that extends the classical bivariate concept to situations where the observed data are curves generated by a stochastic process. Since Spearman's coefficient for bivariate samples is based on the natural data ordering in dimension one, we need to consider a data order in the functional context where a natural order between functions does not exist. The development uses a pre-order inspired in a depth definition but considering a down-up ordering instead of a center-outwards ordering of the sample. We show some of the main characteristics of Spearman's coefficient for functions and propose an independence test with a bootstrap methodology. We illustrate the performance of the new coefficient with both simulated and real data.

Keywords: Spearman's coefficient, dependence, functional data, grades.

Email addresses: daliajazmin.valencia@uc3m.es (Dalia Valencia), rosaelvira.lillo@uc3m.es (Rosa E. Lillo), juan.romo@uc3m.es (Juan Romo)

Spearman Coefficient for functions

Dalia Valencia* Rosa E. Lillo† Juan Romo‡

December 12, 2013

Abstract

We present a notion of Spearman's coefficient for functional data that extends the classical bivariate concept to situations where the observed data are curves generated by a stochastic process. Since Spearman's coefficient for bivariate samples is based on the natural data ordering in dimension one, we need to consider a data order in the functional context where a natural order between functions does not exist. The development uses a pre-order inspired in a depth definition but considering a down-up ordering instead of a center-outwards ordering of the sample. We show some of the main characteristics of Spearman's coefficient for functions and propose an independence test with a bootstrap methodology. We illustrate the performance of the new coefficient with both simulated and real data.

1 Introduction

Functional data analysis (FDA) has recently become a topic of interest in statistics, having a wide range of applications in chemometrics, medicine, meteorology, economics and analysis of images, among others, where it can be assumed that the observed data are functions generated by a stochastic process. Despite the fact that several multivariate methods are not usually well suited for functional datasets, many multivariate techniques have inspired advances in FDA, for example, to quantify the relationship of dependence between two or more groups of functional data.
Despite the fact that several multivariate methods are not usually well suited for functional datasets, many multivariate techniques have inspired advances in FDA, for example, to quantify the relationship of dependence between two or more groups of functional data.

*daliajazmin.valencia@uc3m.es
†rosaelvira.lillo@uc3m.es
‡juan.romo@uc3m.es

The observations represented by curves come from a real-valued stochastic process in continuous time, $X(t)\ \ t \in [0,T]$. Most of the statistical analysis with functional data consider just the univariate case (see Ramsay and Silverman [26] and Ferraty and Vieu [9]), where a path of $X(t)$ is represented by a single curve. However, less attention has been paid to the multidimensional case, where a path of $X(t)$ is a set of p curves. Berrendero et al. [1] studied principal component analysis for multivariate functional data and more recently, Jacques and Preda [15] introduced the cluster analysis for multivariate functional data. Ramsay and Silverman [26] show some real examples and techniques for the statistical treatment of bivariate functional data. Specifically, the study of the dependence between two or more groups of functional data is being increasingly developed. For example, Opgen-Rhein and Strimmer [24] proposed an estimator of the dynamical correlation that is based on the concept of dynamical correlation introduced by Dubin and Müller [4] for longitudinal data, which provides a measure of similarity between pairs of functional observations. He et al. ([12], [13]) proposed a natural way of finding the canonical correlation for functions, previously introduced by Leurgans et al. [16]. They found significant difficulties such as the covariance operator not being invertible, since it is a compact operator that is not generally invertible in infinite dimensional Hilbert space. Li and Chow [17] provided a generalization of Pearson's correlation coefficient for functional data that allows a measure of agreement to be introduced. This measure is called the concordance correlation coefficient and was used to evaluate the reproducibility of repeated-paired curve data. Valencia et al. [29] defined Kendall's τ coefficient for functions considering pre-orders that permit the sorting of the functional observations and calculation of the concordant and discordant pairs of a bivariate sample of curves. Ramsay and Silverman [26] also consider a dependence functional measure called the cross-correlation function. This measure provides a surface that evaluates point by point the usual Pearson correlation between the corresponding values of the pair of curves in two given points.

In this paper, we focus on numeric dependence measures for functions with the first contribution being the definition of a Spearman coefficient that extends the classical bivariate concept, based on the ranks of the observations of the sample. However, the main difficulty that we found in extending this coefficient to the functional setting is that there is not a natural order among functions. Thus, our first task is to consider a suitable way to sort the observations depending on the relative position of the curve within the sample. There are some alternatives for sorting the curves; one of them is based on the notion of depth that measures the centrality of a curve with respect to the group to which it belongs, so depth provides a way of ordering data from the center outwards. Different notions of depth have been studied for functional data (see for example, Fraiman and Muñiz [8], Cuevas et al. [2], López-Pintado and Romo [19], [20]) and each definition gives rise to different ways of ordering the curves. However, alternative definitions of ordering can also be interesting; for example, in Valencia et al. [29] functions are compared depending on their maximun values or on their total area below the curves. In this paper, we have used the pre-order introduced in López-Pintado and Romo [20] and the way of sorting the functions used in Martín-Barragan et al. [21], who provided a way of sorting the data in a down-up direction based on the concepts of hypograph and epigraph of a function. This pre-order takes into consideration more the particular

structure of the data. We also introduce the notion of grade for functions that it is useful to develop the theoretical background necessary to properly define the Spearman coefficient. The main properties of this coefficient as a well-defined dependence measure are also derived. To our knowledge, an independence test for functional data has not been proposed in the literature. Here, we try to fill this gap and present an independence test based on a bootstrap methodology suitable to be applied with some of the numeric dependence coefficients previously introduced in the literature.

The paper is organized as follows. In Section 2, we recall some concepts about Spearman's coefficient for bivariate samples necessary to understand the extension to the functional context. Section 3 presents the main definitions that allow functions to be sorted. In Section 4, we introduce Spearman's coefficient for functions and study its properties. A simulation study and a robustness analysis is carried out in Section 5. In Section 6 the independence test is provided as well as a simulation study. Several examples with real data are shown in Section 7. Finally, Section 8 gathers the main conclusions.

2 Preliminaries

Spearman's coefficient is a non-parametric measure of association between two random variables. It is defined as the Pearson correlation coefficient between the ranks of the sample, being useful when the data are distribution free, so it is not necessary to assume the assumption of normality (Pearson [23], Hauke and Kossowski [11]). It is well known that it presents significant advantages over the Pearson coefficient: (1) It is a more robust coefficient (less sensitive to outliers) and (2) Spearman's coefficient is a better indicator than the Pearson correlation for determining whether a relationship exists between two variables when the relationship is nonlinear.

One of the definitions of the Spearman coefficient between two random variables is the following: Let $(X_1, Y_1), (X_2, Y_2)$ and let (X_3, Y_3) be three independent copies of the random vector (X, Y) with joint distribution function F_{XY} and marginals F_X and F_Y, respectively. The Spearman coefficient between the variables X and Y (denoted by ρ_s) is defined as

$$\rho_s = 3[P\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_3) < 0\}].$$

As we can see, Spearman's coefficient is proportional to the difference between the probability of concordance and the probability of discordance for two vectors (X_1, Y_1) and (X_2, Y_3). The Kendall τ is also based on the concordance probability and it is well known that both coefficients measure non linear dependence from a non-parametric point of view. (For further details see Nelsen [22]).

However, we are interested in the equivalent definition of ρ_s given by calculating the Pearson coefficient between the uniform random variables $U = F_X(X)$ and $V = F_Y(Y)$; that is,

$$\rho_s = \rho_p[U, V] = \frac{E(UV) - E(U)E(V)}{\sqrt{Var(U)}\sqrt{Var(V)}}, \tag{1}$$

where ρ_p denotes the Pearson coefficient. The random variables U and V are called the "grades" of X and Y and the realizations u of U and v of V can be obtained evaluating the realizations x of X and y of Y in the distribution functions F_X and F_Y, respectively. Therefore, $u = F_X(x)$ and $v = F_Y(y)$ can also be called the grades of x and y. These grades can be seen as the population definition analogs of ranks (see Nelsen [22], page 169). If the distribution functions are unknown, then the grades of x and y can be estimated through the empirical distribution, i.e., $\hat{u} = \hat{F}_X(x)$ similar to $\hat{v}$ and hence we can calculate the sample version of this coefficient by calculating the sample version of the Pearson coefficient between the estimated grades. For this reason, Spearman's coefficient is also called *the grade correlation coefficient*. Observe that the grades are values that are always in $[0, 1]$ and they are bounded independently of the support of the random variables. Therefore, an estimation of the Spearman coefficient is less sensitive in the presence of outliers than an estimation of the Pearson coefficient and, most importantly, ρ_s is well defined for all pairs of random variables, whereas ρ_p needs the random variables to have a finite second moment.

The definition of ρ_s based on grades inspires the development provided in this paper: defining a Spearman coefficient for functions extending the definition of grades for functions. This is done in the following section.

Spearman's coefficient satisfies some general and intuitive properties required for any reasonable dependence measure. For example, the sign of ρ_s indicates the direction of association between X and Y, so that if Y increases when X increases, Spearman's coefficient will be positive. Now, if Y tends to decrease when X increases, Spearman's coefficient is negative. A Spearman's coefficient with value zero indicates that there is not a clear tendency for Y to either increase or decrease when X increases and its value is zero if the variables are independent. Spearman's coefficient increases in magnitude as X and Y become closer to being perfect monotone functions of each other. When X and Y are perfectly monotonically related (positive perfect dependence), Spearman's coefficient becomes 1. Therefore, Spearman's coefficient informs about the dependence, either positive or negative, between the random variables.

3 Grades for functional data

The possible concept of grade for functions may be linked to the relative position of a curve in the sample which implicitly implies defining an ordering among functions. There are some alternatives to sorting curves, one of them based on the notion of depth that measures the centrality of a curve with respect to the group to which it belongs; thus, depth provides a way of ordering data from center outwards. Different notions of depth have been studied for functional data (see for example, Fraiman and Muñiz [8], Cuevas et al.[2], López-Pintado and Romo [19], [20]) and each definition leads to different ways of ordering the curves. However, alternative definitions of ordering can also be interesting; for example, in Valencia et al. [29] the curves are ordered depending respectively on values of their maximum or their area below the curves in order to define a Kendall tau coefficient for functions. Martín-Barragan et al. [21] apply the concept of epigraphs and hypographs of a function to define some indexes that are useful for sorting curves in a down-up direction, even when

the curves cross.

To define the grades of the curves, we will follow some concepts introduced in López-Pintado and Romo [20] and the idea of ordering implemented in Martín-Barragan et al.[21]. In López-Pintado and Romo [20], two concepts called the *Inferior Length* and the *Superior Length* of a curve, are defined as the foundation of a depth definition and these concepts are used to introduce a new boxplot for functional data in Martín-Barragan et al. [21]. In order to make our paper self contained, we briefly define the previous concepts.

Let $C(I)$ be the space of the continuous functions defined in a compact interval I. Consider a stochastic process $X(t)$ with distribution P and whose sample paths are in $C(I)$. Let $x_1(t), \ldots, x_n(t)$ be a sample of curves from P. The graph of a function x is the subset of the plane $G(x) = \{(t, x(t)), t \in I\}$. The hypograph, written as hyp, and the epygraph, written as epi, of a function x in $C(I)$ are given respectively by

$$hyp(x) = \{(t, y) \in I \times \mathbb{R} : y \leq x(t)\},$$
$$epi(x) = \{(t, y) \in I \times \mathbb{R} : y \geq x(t)\}.$$

A natural form of ordering curves is pointwise, which means that a curve x is greater than another curve y if, and only if, $hyp(y) \subset hyp(x)$ or $epi(x) \subset epi(y)$, for all $t \in I$. However, in practical situations the curves in a sample can be crossed and hence the natural ordering in these cases does not work. An alternative for ordering curves can be developed by using two concepts, the Inferior Length and the Superior Length of a curve with respect to a stochastic process $X(t)$:

$$IL(x) = \frac{1}{\lambda(I)} E[\lambda\{t \in I : x(t) \geq X(t)\}],$$
$$SL(x) = \frac{1}{\lambda(I)} E[\lambda\{t \in I : x(t) \leq X(t)\}],$$

where λ stands for the Lebesgue measure on $\mathbb{R}$. The inferior length $IL(x)$ can be interpreted as the "proportion of time" that the stochastic process $X(t)$ is smaller than x and the superior length $SL(x)$ is the "proportion of time" that the stochastic process $X(t)$ is greater than x.

These notions are behind the definitions of the grades of a stochastic process $X(t)$ with respect to another process $\widetilde{X}_t$ as follows:

Definition 1 *Let $X(t)$ and $\widetilde{X}(t)$ be two stochastic processes. Then,*

$$IL\text{-}grade(X(t))_{\widetilde{X}(t)} = \frac{1}{\lambda(I)} E_{\widetilde{X}(t)}[\lambda\{t \in I : X(t) \geq \widetilde{X}(t)\}],$$
$$SL\text{-}grade(X(t))_{\widetilde{X}(t)} = \frac{1}{\lambda(I)} E_{\widetilde{X}(t)}[\lambda\{t \in I : X(t) \leq \widetilde{X}(t)\}].$$

Observe that $IL\text{-}grade$ or $SL\text{-}grade$ assigns a value between [0,1] to each process. We note that if the $X(t)$ and $\widetilde{X}(t)$ have the same distribution, we then eliminate

$\widetilde{X}(t)$ from the definitions of *IL-grade* and *SL-grade* to avoid hard notation.

If we consider a sample of functional data, $x_1(t), \ldots, x_n(t)$ and fix any curve $x = x(t)$ of the data set, the sample version of both *IL-grade* and *SL-grade* can be easily obtained by substituting the expectation by the sample mean, respectively

$$IL_n\text{-}grade(x) = \frac{1}{n\lambda(I)} \sum_{i=1}^{n} \lambda\{t \in I : x(t) \geq x_i(t)\},$$
$$SL_n\text{-}grade(x) = \frac{1}{n\lambda(I)} \sum_{i=1}^{n} \lambda\{t \in I : x(t) \leq x_i(t)\}.$$

It is noteworthy that IL_n-*grade*(x) or SL_n-*grade*(x) has been viewed as the relative position of a curve with respect to the sample. Also, note that the curves can be ordered by sorting the values of IL_n-*grade* or SL_n-*grade* for each one of them. That is,

Definition 2 *Consider functional observations $x_1(t), \ldots, x_n(t)$ of a stochastic process X. Then,*

$$x_i(t) \preceq x_j(t) \equiv IL_n\text{-}grade(x_i) \leq IL_n\text{-}grade(x_j).$$

A similar definition can be obtained by replacing the IL_n-*grade* with SL_n-*grade*.

The relation given in Definition 2 meets important properties such as reflectivity and transitivity, but, unfortunately, it does not satisfy the antisymmetry property. Therefore, the relation is a pre-order, which is less restrictive than a partial order and allows us to compare any pair of functions in the sample. Observe that if the curves do not cross each other, Definition 2 corresponds to the pointwise order.

To illustrate this pre-order, observe the example in Figure 1 that shows the IL_n-*grade* assigned to each function in a sample of four functions. The blue curve has the smallest IL_n-*grade* because the proportion of time that it is above any other curve is smaller than the value assigned to any curve in the same sample. The black curve has the largest IL_n-*grade* value assigned, since in this case the time proportion is greater than any other. The proportions assigned to each curve are what we call the grade of the curve regarding the sample. Note that the largest functional grade in the sample may not be one unless the curve with the highest functional grade does not cross with any other, which means that it will be largest point-to-point than them. Once the grades are introduced, we can define Spearman's coefficient for functions in a parallel way to (1).

4 Spearman's coefficient for functional data

In this section, we define the concept of Spearman's coefficient in the functional context in order to quantify the dependence in a bivariate data set of functions. Taking into account Definition (1), we define a Spearman coefficient for two stochastic processes as the Pearson coefficient between the random variables *IL-grade*$(X(t))$ and *IL-grade*$(Y(t))$; that is,

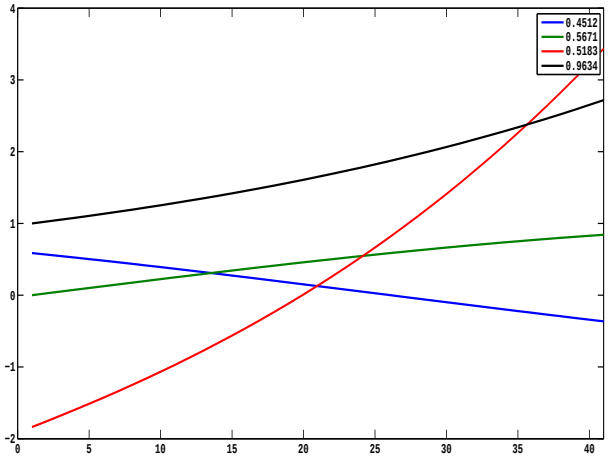

Figure 1: Grades for functions

Definition 3 (Spearman coefficient for stochastic processes) *Let* $(X(t), Y(t))$ *be bivariate stochastic process whose paths are continuous functions on an interval* $I \subset \mathbb{R}$*. Then, Spearman's coefficient of* $(X(t), Y(t))$ *is:*

$$\rho_s(X(t), Y(t)) \equiv \rho_p(IL\text{-}grade(X(t)), IL\text{-}grade(Y(t))), \tag{2}$$

where ρ_p *denotes the Pearson correlation coefficient and IL-grade*$(\cdot)$ *is the grade associated to a stochastic process given in Definition 1.*

In the same way, the sample version of ρ_s is the following:

Definition 4 (Spearman's coefficient for functions) *Let*

$$(\mathbf{x}, \mathbf{y}) = \{(x_1(t), y_1(t)), \ldots, (x_n(t), y_n(t))\}$$

be a bivariate functional sample from $(X(t), Y(t))$*. Then, the Spearman coefficient related to the data set and denoted by* $\widehat{\rho}_s$ *is defined by*

$$\widehat{\rho}_s \equiv \widehat{\rho}_p(IL_n\text{-}grade(\mathbf{x}), IL_n\text{-}grade(\mathbf{y})), \tag{3}$$

where,

$$IL_n\text{-}grade(\mathbf{x}) = \{IL_n\text{-}grade(x_1), IL_n\text{-}grade(x_2), \ldots, IL_n\text{-}grade(x_n)\}$$
$$IL_n\text{-}grade(\mathbf{y}) = \{IL_n\text{-}grade(y_1), IL_n\text{-}grade(y_2), \ldots, IL_n\text{-}grade(y_n)\}.$$

Another definition of Spearman's coefficient for functions can be obtained by replacing IL_n-*grade* by SL_n-*grade*. In order to illustrate how the Spearman coefficient works, we have taken a small bivariate set of four curves and calculated the corresponding coefficient. Figure 2 shows the pairs of curves, each pair represented by its own color. We can see that the curves in a group are organized in a different way than their respective partner in the other group. Observe that the order of the curves in first group seems to have a more or less opposite direction with respect to the other group. Therefore, Spearman's coefficient for functional data is small, indicating to us that the association between the groups of curves is weak and negative.

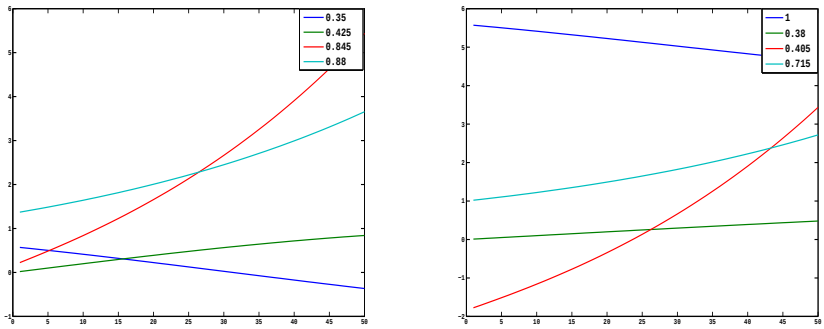

Figure 2: Spearman's coefficient for functional data, $\widehat{\rho}_s = -0.2994$.

4.1 Properties of Spearman's coefficient for functional data

As commented in Section 2, Spearman's coefficient for bivariate data satisfies certain desirable properties required for a dependence measure (see Xu et al. [28]). In this section, we prove that Spearman's coefficient for stochastic processes also possesses such properties. Let $(X(t), Y(t))$ be a bivariate stochastic process and ρ_s be Spearman's coefficient as in Definition 3. Then ρ_s satisfies the following properties:

1. $\rho_s(X(t), Y(t)) = \rho_s(Y(t), X(t))$. (Symmetry).
2. $-1 \leq \rho_s(X(t), Y(t)) \leq 1$.
3. $\rho_s(X(t), g(X(t))) = 1$, for any monotone increasing function g.
4. $\rho_s(X(t), g(X(t))) = -1$, for any monotone decreasing function g.
5. Spearman's coefficient for functions is invariant under strictly increasing transformations of the functional variables; that is,

$$\rho_s(\alpha(X(t)), \beta(Y(t))) = \rho_s(X(t), Y(t)).$$

   for any α and β being strictly increasing functions.
6. If $X(t)$ and $Y(t)$ are stochastically independent then $\rho_s(X(t), Y(t)) = 0$.
7. The sample Spearman's coefficient is a consistent estimator of the population coefficient.

The proofs of properties 1 and 2 are trivial from the definition of ρ_s. The proof of properties 3, 4 and 5 are based on the following:

$$\begin{aligned}
\text{IL-grade}(g(X(t)))_{\widetilde{X}(t)} &= \frac{1}{\lambda(I)} E_{\widetilde{X}(t)}[\lambda\{t \in I : g(X(t)) \geq g(\widetilde{X}(t))\}] \\
&= \frac{1}{\lambda(I)} E_{\widetilde{X}(t)}[\lambda\{t \in I : X(t) \geq \widetilde{X}(t)\}] \\
&= \text{IL-grade}(X(t))_{\widetilde{X}(t)},
\end{aligned}$$

for any monotone increasing function g. The proof of property 6 is based on that, if $X(t)$ and $Y(t)$ are independent then *IL-grade*$(X(t))$ and *IL-grade*$(Y(t))$ are also independent. Therefore, $\rho_s(X(t), Y(t)) = \rho_p(\text{IL-grade}(X(t)), \text{IL-grade}(Y(t))) = 0$ by the well known property of the Pearson coefficient. The last property holds since, as n goes to infinity,

$$\frac{\sum_{i=1}^{n} IL_n\text{-}grade(x_i)}{n} \xrightarrow{\text{a.s}} E[IL\text{-}grade(X(t))],$$

where $x_1, \ldots, x_n$ is a sample from $X(t)$. Finally, since $\widehat{\rho}_p$ is a consistent estimator, also $\widehat{\rho}_s$ is.

5 Simulation study

In this section we show how Spearman's coefficient works in several simulated data sets and we establish comparisons with other dependence measures introduced previously in the literature. Specifically, we consider the canonical correlation, the dynamical correlation, Pearson's coefficient for functional data and Kendall's τ for functions.

Extension of the canonical correlation to functional data has been proposed in Leurgans et al. [16], which pointed to the need for regularization in order to provide a greater interpretability of the results and useful information from the data. As is commented in [26], canonical correlation analysis seeks to investigate which modes of variability in two sets of curves are most associated with one another. As usual, assume that n observed pairs of data curves (x_i, y_i) are available for argument t in some finite interval I, and all integrals are taken over I. The problem is finding a pair of functions (ξ, η) called canonical variates weight that maximize the penalized squared sample correlation.

$$\widehat{\rho}_c(\xi, \eta) = \frac{\left\{cov\left(\int \xi x_i, \int \eta y_i\right)\right\}^2}{\left\{var\left(\int \xi x_i\right) + \lambda||D^2\xi||^2\right\}\left\{var\left(\int \eta y_i\right) + \lambda||D^2\eta||^2\right\}},$$

where λ is a positive smoothing parameter and $||D^2 f||^2 = \int \left(D^2 f\right)^2$, the integrated squared curvature of f that quantifies its roughness.

Having a pair of canonical variables with fairly smooth weight functions and correlations that are not excessively low can be achieved by choosing the appropriate smoothing parameter. This parameter can be chosen subjectively but also can be selected through a cross-validation score if an automatic procedure is required.

Another dependence measure that we consider is the dynamical correlation introduced by Dubin and Müller [4] as a simple and efficient non-parametric correlation measure for multivariate longitudinal data. They interpreted dynamical correlation as a measure of the average concordant or discordant behavior of pairs of random trajectories, in the sense that "*if both trajectories tend to be mostly on the same side of their time average (a constant), then the dynamical correlation is positive; if the opposite occurs, then the dynamical correlation is negative*". We will use in the paper the following estimator of the dynamical correlation proposed in Opgen-Rhein and Strimmer that is a slightly revised version of the dynamical correlation introduced in

Dubin and Müller [4],

$$\widehat{\rho}_d = \frac{1}{n-1}\sum_{i=1}^{n}\langle x_i^s(t), y_i^s(t)\rangle,$$

where $x^s(t) = \frac{x^c(t)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\langle x_i^c(t), x_i^c(t)\rangle}}$, and $x^c(t)$ are functions centered in space and time simultaneously, i.e.,

$$x^c(t) = x(t) - \langle \overline{x}(t), 1\rangle, \quad \text{where} \quad \overline{x}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t),$$

and $\langle\cdot\rangle$ is the usual inner product for functions $\langle x(t), y(t)\rangle = \int_I x(t)y(t)dt$. As we can see, $\widehat{\rho}_d$ is an estimator of the population dynamical correlation

$$\rho_d = E\left\langle X^S(t), Y^S(t)\right\rangle,$$

that can been viewed as an average of individual correlations in contrast with the canonical correlation, which is the solution of a maximization problem.

Li and Chow [17] introduced the following generalization of Pearson's correlation coefficient for two stochastic processes $X(t)$ and $Y(t)$,

$$\rho(X(t), Y(t)) = \frac{\langle X(t) - E(X(t)), Y(t) - E(Y(t))\rangle}{\|X(t) - E(X(t))\|\|Y(t) - E(Y(t))\|}, \tag{4}$$

where the inner product is defined as $\langle X(\cdot), Y(\cdot)\rangle = E\int X(t)Y(t)w(t)dt$, and the norm is that induced by the inner product. In our study, we set $w(t) = 1$ assigning the same weight for each t.

Finally, we establish comparisons with Kendall's τ for functions introduced in Valencia et al. [29]. It is well known that Kendall's tau for bivariate data measures a form of dependence known as concordance. Therefore, it was necessary to define concordance in the functional setting which implicitly implies considering pre-orders for functions. To carry out this task, the following pre-orders for functions were considered in [29]:

- $x \prec_1 y \equiv \max_{t\in I} x(t) < \max_{t\in I} y(t)$.
- $x \prec_2 y \equiv \int_I (x(t) - y(t))dt < 0$.

Informally, a pair of random variables are concordant if "large" values of one tend to be associated with "large" values of the other and "small" values of one with "small" values of the other. To be more precise, let (x_1, y_1) and let (x_2, y_2) denote two observations of a bivariate stochastic process $(X(t), Y(t))$. We say that (x_1, y_1) and (x_2, y_2) are concordant if:

$$x_1 \prec_i x_2 \text{ and } y_1 \prec_i y_2, \text{ or } x_2 \prec_i x_1 \text{ and } y_2 \prec_i y_1,$$

for any of the pre-orders $i = 1, 2$ previously considered. Similarly, (x_1, y_1) and (x_2, y_2) are discordant if

$$x_1 \prec_i x_2 \text{ and } y_2 \prec_i y_1 \text{ or } x_2 \prec_i x_1 \text{ and } y_1 \prec_i y_2.$$

These definitions allow us to introduce the sample version of the measure of association known as Kendall's τ in terms of concordance as follows,

$$\hat{\tau} \equiv \frac{\#\ concordant\ \ pairs - \#\ \ discordant\ \ pairs}{\binom{n}{2}}. \tag{5}$$

To the extent that this coefficient takes into account all the possible observation pairs that may be compared with each other to calculate the number of both concordant and discordant pairs, $\hat{\tau}$ is a more global dependence measure than the dynamical correlation where each standardized curve is only compared with its corresponding pair. This makes the dynamical correlation capture changes only at an individual performance level, and makes the Kendall's coefficient detect changes at a more collective level. It is also shown in Valencia et al. [29] that τ is an more robust estimator than the other dependence measures.

For the data given in Figure 2, the different association measures have the values:

$$\widehat{\tau}_1 = 0, \quad \widehat{\tau}_2 = -0.33, \quad \widehat{\rho}_c = 0.83, \quad \widehat{\rho}_d = -0.13, \quad \widehat{\rho}_p = -0.2374$$

Note that Kendall's τ built with the pre-order of maximum and denoted as $\widehat{\tau}_1$ is zero since there are as many concordant pairs as discordant pairs. The canonical correlation $\widehat{\rho}_c$ has a very large and positive value since it is always positive and does not allow the direction of the dependency to be identified. The dynamical correlation $\widehat{\rho}_d$, Kendall's τ built with the pre-order of the integral $\widehat{\tau}_2$ and Pearson's correlation coefficient for functional data $\widehat{\rho}_p$ have negative values that reflect the direction of weak dependence shown in the data set as well as Spearman's coefficient ($\widehat{\rho}_s = -0.2994$).

We have simulated 50 realizations from different processes $X(t) = f_1(t, Z_1)$ and $Y(t) = f_2(t, Z_2)$, where (Z_1, Z_2) represents the random part of the processes. In the remainder of this section, we assume (Z_1, Z_2) to be a normal bivariate with correlation σ_{12}. For each pair $(f_1,\ f_2)$, we use a different correlation σ_{12}.

Table 1 shows the sample means of different association measures for the simulated samples with $n = d = 50$ and 100 replications. We have also included the standard deviation (between parenthesis). We can see that both coefficients, the Spearman and Kendall, properly reflect the cases where the pairs of functions present perfect co-monotonicity or counter-monotonicity, (see rows 3, 4 and 5 in Table 1). As we know, the canonical correlation is always positive, i.e., it does not capture the direction of the dependence. Note from the definition of the dynamical correlation that, it just reflects individual changes between the pairs of functions rather than among groups. On the other hand, Pearson's coefficient does not work well when the dependence relations are not lineal, as in cases 4 and 5.

We have also analyze the sensitivity of $\widehat{\rho}_s$ with respect to the size n. We will use the following two pairs of stochastic process that correspond with row 1 in Table 1 with $\sigma_{12} = 0.8$ and $\sigma_{12} = 0.1$:

$$X(t) = (t + Z_1)^3 + (t + Z_1)^2 + 3(t + Z_1), \quad Y(t) = (t + Z_2)^2 + \frac{7}{8}(t + Z_2) - 10.$$

Table 1: Dependence measures in simulated data

	$X(t) = f_1(t, Z_1)$	$Y(t) = f_2(t, Z_2)$	σ_{12}	$\bar{\hat{\rho}}_s IL$	$\bar{\hat{\rho}}_s SL$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\rho}}_c$	$\bar{\hat{\rho}}_d$	$\bar{\hat{\rho}}_p$
1	$(t+Z_1)^3 + (t+Z_1)^2 + 3(t+Z_1)$	$(t+Z_2)^2 + \frac{7}{8}(t+Z_2) - 10$	0.8	0.667 (0.0811)	0.6596 (0.0882)	0.4861 (0.0657)	0.4874 (0.0711)	0.7448 (0.0898)	0.7098 (0.1139)	0.6943 (0.1055)
2	$\sin(t+Z_1)$	$\cos(t+Z_2)$	-0.7	0.4354 (0.1244)	0.445 (0.1407)	0.3084 (0.0923)	0.2774 (0.0835)	0.5367 (0.1004)	0.3605 (0.11)	0.4022 (0.1189)
3	$(t+Z_1)^2$	$(t+Z_1)^4$	1	1 (0)	1 (0)	1 (0)	1 (0)	0.9566 (0.0118)	0.922 (0.0125)	0.9179 (0.0127)
4	$(t+Z_1)^2 + 7(t+Z_1) + 2$	$((t+Z_2)^2 + 7(t+Z_2) + 2)^3$	1	0.9997 (0.0029)	1 (0)	1 (0)	1 (0)	0.9989 (0)	0.7779 (0.0347)	0.7688 (0.0278)
5	$(t+Z_1)^2 + 7(t+Z_1) + 2$	$1 - ((t+Z_2)^2 + 7(t+Z_2) + 2)^3$	1	-1 (0)	-1 (0)	-1 (0)	-1 (0)	0.999 (0.0009)	-0.78 (0.0275)	-0.7644 (0.0285)
6	$\exp(t+Z_1)$	$(t+Z_2)^3 + (t+Z_2)^2 + 3(t+Z_2)$	0.6	0.5802 (0.0967)	0.5546 (0.1072)	0.4047 (0.0811)	0.4138 (0.0751)	0.5098 (0.1431)	0.5682 (0.1301)	0.5193 (0.1559)
7	$\exp(t+Z_1)^2$	$\cos(t+Z_2)$	-0.8	0.4417 (0.1195)	0.4430 (0.1198)	0.3097 (0.0922)	0.2982 (0.1035)	0.3101 (0.07)	0.0408 (0.1458)	0.0846 (0.1697)
8	$\sin(t+Z_1)$	$(t+Z_2)^2$	0.4	0.1706 (0.1331)	0.1458 (0.1307)	0.1080 (0.1035)	0.1059 (0.1021)	0.3382 (0.1132)	0.1647 (0.0916)	0.1173 (0.1175)
9	$(t+Z_1)^2 + 9(t+Z_1) - 5$	$\cos(3t+Z_2)$	1	-0.935 (0.0176)	-0.9327 (0.0199)	-0.7198 (0.0853)	-0.9476 (0.0358)	0.9334 (0.0458)	-0.7244 (0.0562)	-0.6976 (0.0708)
10	$\exp(t^2+Z_1)$	$(t+Z_2)^2 - 8t + Z_2$	0.9	0.7743 (0.0634)	0.7892 (0.0608)	0.3621 (0.1078)	0.5991 (0.0706)	0.8544 (0.0485)	0.4620 (0.1215)	0.8309 (0.0616)
11	$\exp(t+Z_1)$	$\sin(t+Z_2)$	0	0.05 (0.1467)	0.0051 (0.1508)	-0.0076 (0.1004)	0.0087 (0.0883)	0.1438 (0.0861)	0.0560 (0.1275)	-0.0209 (0.1221)

We have considered $n = 25,\ 50,\ 100,\ 150$ and 1000 with $d = 50$. Table 2 shows that the changes in $\widehat{\rho}_s$ are negligible and it is stable with respect to the sample size. Table 3 illustrates the sensitivity with respect to d. Now, fix $n = 50$, and move $d = 25,\ 50,\ 100,\ 150$ and 1000 points. It is noteworthy that the coefficients present good stability with respect to the number of points taken to discretize the functions. We point out that we have made the sensitivity analysis with other models, but the conclusions are the same for the models reported.

Table 2: Sensitivity to sample size

sample size	Model 1 $\widehat{\widehat{\rho}}_s IL$	Model 1 $\widehat{\widehat{\rho}}_s SL$	Model 2 $\widehat{\widehat{\rho}}_s IL$	Model 2 $\widehat{\widehat{\rho}}_s SL$
25	0.6492	0.6612	0.077	0.0781
	(0.1270)	(0.1301)	(0.2030)	(0.2137)
50	0.6697	0.6748	0.0732	0.0993
	(0.0881)	(0.0686)	(0.1426)	(0.1369)
100	0.6709	0.6534	0.0883	0.0754
	(0.0559)	(0.0617)	(0.0945)	(0.0998)
150	0.6598	0.6668	0.0626	0.0685
	(0.0448)	(0.0495)	(0.0847)	(0.0789)
1000	0.6699	0.6724	0.0767	0.0807
	(0.0177)	(0.0204)	(0.0341)	(0.0348)

Table 3: Sensitivity to the number of points in the discretization

numbers of points	Model 1 $\widehat{\widehat{\rho}}_s IL$	Model 1 $\widehat{\widehat{\rho}}_s SL$	Model 2 $\widehat{\widehat{\rho}}_s IL$	Model 2 $\widehat{\widehat{\rho}}_s SL$
25	0.6542	0.6542	0.0647	0.0647
50	0.6542	0.6542	0.0648	0.0648
100	0.6546	0.6546	0.0648	0.0648
150	0.6548	0.6548	0.0646	0.0646
1000	0.6548	0.6548	0.0648	0.0648

5.1 Robustness

Spearman's coefficient is a more appropriate association measure than Pearson's correlation when the data are ordinal or non-normally distributed or a tiny fraction of outliers exists. In this section, we analyze this last point. That is, we check if Spearman's coefficient for functions fulfills the robustness property by contaminating a sample with the three different types of outliers commonly used in the functional context: shape outliers, magnitude outliers and shape-magnitude outliers. The method to contaminate data is similar to that implemented in Valencia et al.[29] where the objective was to show the robustness of Kendall's τ for functions. The procedure is summarized briefly as follows: we have simulated fifty paths of the stochastic processes,

$$X(t) = \exp(t + Z_1),\ \ Y(t) = (t + Z_2)^3 + (t + Z_2)^2 + 3(t + Z_2),\ \sigma_{12} = 0.6, \qquad (6)$$

and the types of outliers to be considered are:

- Shape outliers. Changing the argument, t to $(1 - t)$.

- Magnitude outliers. Adding a constant to the original process, $X(t)$ to $X(t)+k$. In our case we will use $k = 60$.
- Shape-magnitude outliers. Changing the argument and adding a constant to the original function, $X(t)$ to $X(1-t)+k$.

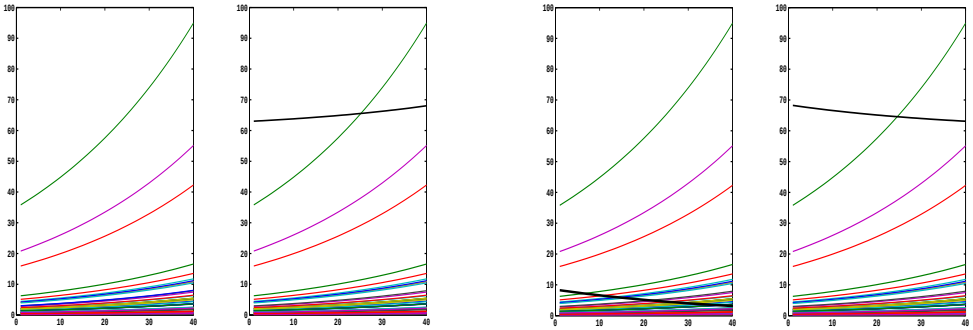

Figure 3: left to right: original data, a magnitude outlier, a shape outlier, a shape-magnitude outlier.

Figure 3 shows a data set generated from stochastic process $X(t) = \exp(t + Z_1)$ and the same data set but contaminated with different types of outliers, which is represented with a black curve.

Contaminated data are considered in processes (6),but introducing outliers in the following way:

1. Contaminating just the group of curves that comes from $X(t)$.
2. Contaminating both groups of curves $(X(t), Y(t))$ in the same position.
3. Contaminating both groups of curves that come from $X(t)$ and $Y(t)$ but in different positions.

Table 4 shows the variation of the coefficients when the outliers are introduced. Each measure is calculated before contaminating the data (row 1). Once the data are contaminated, we report the relative variation of the association measure with respect to its value in the uncontaminated data set.

We can see that Kendall's τ is the most robust coefficient in most cases. However, Spearman's coefficient also exhibits a good degree of robustness, even being more robust in general than the canonical correlation, dynamical correlation and the Pearson correlation coefficient for functions. The values in red are those where the Spearman coefficients present smaller variation than the other measures except for Kendall's τ. We highlight that the robustness analysis has been made with other models $(X(t), Y(t))$ and the same conclusions can be drawn.

Table 4: Variation of the coefficients in presence of a different number of outliers

Contaminated Groups	Type of Outliers	Nº outlyers	$\widehat{\rho}_{IL}$	$\widehat{\rho}_{SL}$	$\widehat{\tau}_1$	$\widehat{\tau}_2$	$\widehat{\rho}_d$	$\widehat{\rho}_c$	$\widehat{\rho}_p$
none	none	0	0.6213	0.6213	0.4547	0.4547	0.5491	0.5449	0.5367
$X(t)$	Shape	1	0.0067	0.0067	0	0	0.00036	0.0027	0.0007
$X(t)$	Shape	2	0.0069	0.0069	0	0	0.042	0.1213	0.0007
$X(t)$, $Y(t)$ same position	Shape	1	0.010	0.010	0	0	0	0.006	0.0015
$X(t)$, $Y(t)$ same position	Shape	2	0.0094	0.0094	0	0	0	0.7511	0.0018
$X(t)$, $Y(t)$ different position	Shape	1	0.0086	0.0086	0	0	0.0009	0.0011	0.00037
$X(t)$, $Y(t)$ different position	Shape	2	0.0072	0.0072	0	0	0.046	0.1477	0.0005
$X(t)$	Magnitude	1	0.045	0.045	0.035	0.039	0.28	0	0.313
$X(t)$	Magnitude	2	0.039	0.039	0.025	0.028	0.066	0.035	0.5446
$X(t)$, $Y(t)$ same position	Magnitude	1	**0.053**	**0.053**	0.0646	0.075	0.227	0.6505	0.2457
$X(t)$, $Y(t)$ same position	Magnitude	2	**0.055**	**0.055**	0.078	0.086	0.47	0.7414	0.3547
$X(t)$, $Y(t)$ different position	Magnitude	1	0.074	0.074	0.072	0.082	0.418	0.008	0.436
$X(t)$, $Y(t)$ different position	Magnitude	2	0.079	0.079	0.075	0.086	0.383	0.017	0.7315
$X(t)$	Shape-magnitude	1	0.045	0.045	0.035	0.039	0.2811	0.001	0.312
$X(t)$	Shape-magnitude	2	**0.039**	**0.039**	0.025	0.028	0.092	0.043	0.5438
$X(t)$, $Y(t)$ same position	Shape-magnitude	1	**0.053**	**0.053**	0.064	0.075	0.227	0.689	0.2467
$X(t)$, $Y(t)$ same position	Shape-magnitude	2	**0.055**	**0.055**	0.086	0.086	0.4775	0.7973	0.3551
$X(t)$, $Y(t)$ different position	Shape-magnitude	1	0.074	0.074	0.072	0.082	0.419	0.0014	0.4373
$X(t)$, $Y(t)$ different position	Shape-magnitude	2	0.079	0.079	0.075	0.086	0.404	0.034	0.730

6 Independence test for functional data

In the literature on association measures, it is usual to provide an independence test to check if the corresponding coefficient used to measure dependence can be considered zero or not (see for example Gibbons [10] and Wilcox [27] for more details). This section deals with the design of a test when data are curves and the hypotheses are:
$H_0: \ \rho_s = 0.$
$H_1: \ \rho_s \neq 0.$
Since the asymptotic distribution for ρ_s is not known when the data set are functions, an alternative methodology is necessary to find the critical region associated with the statistics $\widehat{\rho}_s$. We will use a bootstrap approach to estimate the statistics distribution, (see Efron [5], Efron and Tibshirani [6], Davison and Hinkley[3], for more information).

Given a sample of functions $(\mathbf{x}, \mathbf{y})$ of size n, B bootstrap samples of size n are obtained by resampling from $(\mathbf{x}, \mathbf{y})$ under the null hypothesis; that is, there is no association between the components of the stochastic process (X, Y) that generated the data set $(\mathbf{x}, \mathbf{y})$. The steps necessary to obtain the *p-value* of the test are summarized in Table 5, where $\widehat{\rho}_s(\mathbf{x}, \mathbf{y})$ is the sampled value of ρ_s and $\widehat{\rho}_s(\mathbf{x}^*, \mathbf{y}^*)$ is its corresponding value for the bootstrap sample. The decision rule is to reject H_0 if $p\text{-}value \leq \alpha$, where α is the significance level. We fix $\alpha = 0.05$ in the following.

To illustrate the results of the bootstrap test, we come back with the simulated

Table 5: Bootstrap test

1. Input: a sample of functions $(\mathbf{x}, \mathbf{y})$ from a stochastic process (X, Y) and α-level.
2. Find $\widehat{\rho}_s(\mathbf{x}, \mathbf{y})$.
3. Obtain under H_0 a bootstrap sample $(\mathbf{x}^*, \mathbf{y}^*)$ of size n from $(\mathbf{x}, \mathbf{y})$.
4. Calculate $\widehat{\rho}_s(\mathbf{x}^*, \mathbf{y}^*)$.
5. Repeat 3 and 4 a sufficient number of times (B).
6. Find $p\text{-}value = \frac{\sum_{i=1}^{B} \mathbf{I}_{[\widehat{\rho}_s(\mathbf{x}_i^*, \mathbf{y}_i^*) \geq \widehat{\rho}_s(\mathbf{x}, \mathbf{y})]}}{B}$.
7. Output: Reject H_0, if $p\text{-}value < \alpha\text{-}level$.

data of in Table 1. We fix a sample of size $n = 50$ and apply the previous test with $B = 2500$. For each case, both $\widehat{\rho}_s IL$ and the *p-value* are displayed in Table 6. Note how the test is consistent when the simulated models are curves generated from stochastic processes with positive or negative perfect dependence. In these cases, the test produces *p-value* equal to zero. We can also observe that when the groups of curves have a high correlation coefficient the *p-value* is smaller than 0.05 so that the null hypothesis is rejected. Likewise, when the groups of curves have a low correlation coefficient, the *p-value* is larger than 0.05 and then the null hypothesis is not rejected.

In order to make comparisons, Table 6 also shows the results of applying the same hypothesis test but considering the statistics $\widehat{\tau}_1$ and $\widehat{\tau}_2$, defined previously. The canonical correlation, dynamical correlation and Pearson correlation coefficient for functions are not considered because these coefficients show a very wide casuistry for which they equal zero. Hence, simulating bootstrap samples under the null hypothesis (independence) is not a good strategy for these coefficients where many anomalies are observed.

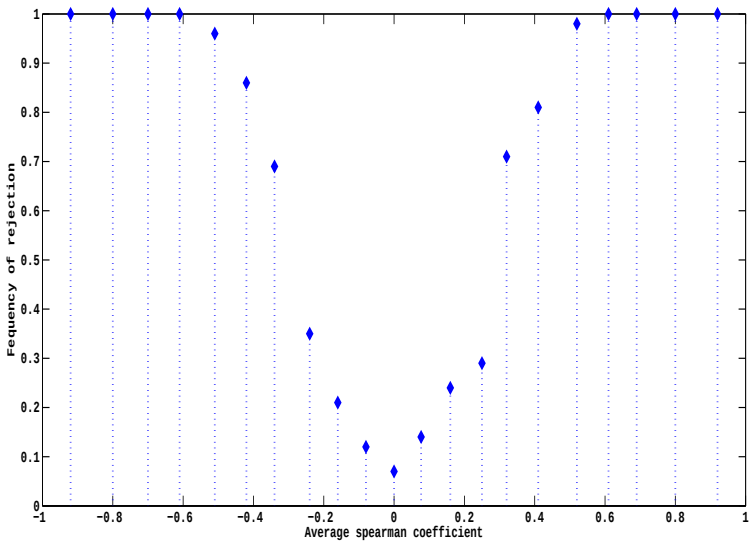

Figure 4: Power test

We now analyze the power of the test with a simulation study. First, we consider

Table 6: Hypothesis test

$X(t)$	$Y(t)$	σ_{12}	$\widehat{\rho}_s IL$	*p-value*	$\widehat{\tau}_1$	*p-value*	$\widehat{\tau}_2$	*p-value*
$(t+Z_1)^2$	$(t+Z_1)^4$	1	1	0	1	0	1	0
$(t+Z_1)^2+7(t+Z_1)+2$	$((t+Z_1)^2+7(t+Z_1)+2)^3$	1	0.9996	0	0.9967	0	0.9967	0
$(t+Z_1)^3+(t+Z_1)^2+3(t+Z_1)$	$(t+Z_2)^2+\frac{7}{8}(t+Z_2)-10$	0.8	0.7197	0	0.4645	0	0.4645	0
$\exp(t+Z_1)$	$(t+Z_2)^3+(t+Z_2)^2+3(t+Z_2)$	0.6	0.6213	0	0.4547	0	0.4547	0
$\sin(t+Z_1)$	$\cos(t+Z_2)$	-0.7	0.4840	0.0002	0.3763	0.0002	0.3127	0.0012
$\sin(t+Z_1)$	$(t+Z_2)^2$	0.4	0.3178	0.0241	0.2212	0.0230	0.2180	0.0244
$\cos(t+Z_1)$	$(t+Z_2)^2-9(t+Z_2)$	0.2	0.0583	0.6813	0.0351	0.7244	0.0351	0.7122
$\exp(t+Z_1)^2$	$5(t-Z_2)^3-3(t+Z_2)+9$	-0.2	0.0442	0.7587	0.0155	0.8812	0.0155	0.8826
$(t+Z_1)^3$	$(t+Z_2)^2+4(t+Z_2)-7$	-0.5	-0.6804	0	-0.4906	0	-0.4906	0
$(t+Z_1)^3+(t+Z_1)^2$	$(t+Z_2)^2-2(t+Z_2)$	-0.9	-0.8815	0	-0.5527	0	-0.7012	0
$(t+Z_1)^2+7(t+Z_1)+2$	$1-((t+Z_1)^2+7(t+Z_1)+2)^3$	1	-0.9938	0	-0.9837	0	-0.9755	0
$(5/9)(t+Z_1)^3$	$48-(5/9)(t+Z_1)^3$	0	-1	0	-1	0	-1	0

Table 7: Relationships between the coefficients, frequency of rejection (fr) and σ_{12}

σ_{12}	-1	-0,9	-0,8	-0,7	-0,6	-0,5	-0,4	-0,3	-0,2	-0,1	0
$\bar{\widehat{\rho}}_s$	-0,92	-0,8	-0,7	-0,61	-0,51	-0,42	-0,34	-0,24	-0,16	-0,08	0
fr	1	1	1	1	0,96	0,86	0,69	0,35	0,21	0,12	0,07
σ_{12}	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	
$\bar{\widehat{\rho}}_s$	0,077	0,16	0,25	0,32	0,41	0,52	0,61	0,69	0,8	0,92	
fr	0,14	0,24	0,29	0,71	0,81	0,98	1	1	1	1	

a bivariate sample of 50 curves generated from the process $[\exp(t+Z_1), \sin(t+Z_2)]$, being (Z_1, Z_2) a normal bivariate with zero mean and correlation σ_{12}. Given that there exists a certain relationship between σ_{12} and $\widehat{\rho}_s$, we consider different values of σ_{12} in the interval $[-1, 1]$ in order to obtain values of $\widehat{\rho}_s$ over all the interval too. For a given σ_{12}, we generate 100 times a sample of $[\exp(t+Z_1), \sin(t+Z_2)]$, calculate $\widehat{\rho}_{si}$, $i = 1, \ldots, 100$ and its corresponding mean $\bar{\widehat{\rho}}_s$. Finally, we show in Figure 6 the frequency of rejection of the null hypothesis versus $\bar{\widehat{\rho}}_s$. The bootstrap part of each iteration is made with $B = 2500$. Table 7 shows the relationship between the coefficient $\bar{\widehat{\rho}}_s$, σ_{12} and the frequency of rejection for the test. As we can see, as $|\bar{\widehat{\rho}}_s|$ increases, the frequency of rejection also increases which ensures the reliability of the test.

To end this section, we have carried out a sensitivity analysis of the test with respect to the bootstrap sample size B and the number of points used to discretize the functions. Table 8 shows the rejection frequency of the null hypothesis for different values of the $\bar{\widehat{\rho}}_s$. We can conclude that the size of bootstrap samples does not significatively affect the frequency of rejection, whereas the power test improves as d increases, which is due to more information being available about the original process.

Table 8: Sensitivity analysis with respect to B and d

$\widehat{\rho}_s$	size of the bootstrap sample					number of points			
	500	1000	1500	2000	2500	25	50	100	150
0.5989	1	0.99	1	1	1	0.99	0.97	1	1
-0.2447	0.43	0.33	0.35	0.36	0.51	0.12	0.51	0.63	0.89
0.012	0.05	0.06	0.04	0.03	0.02	0	0.02	0.23	0.3
0.2516	0.47	0.5	0.38	0.42	0.36	0.13	0.36	0.69	0.73
0.6968	1	1	1	1	1	1	1	1	1
0.7969	1	1	1	1	1	1	1	1	1

7 Application to real data sets

We consider three real data sets. The first one is composed of daily temperature and precipitation per year in 35 Canadian weather stations (see Ramsay and Silverman[26]). We also have the same data set by months. The sample size is 35. The objective in this first example is to measure the association between temperature and precipitation. The second data set corresponds to monthly temperatures in four

cities of Canada from 1985 until 2004 (taken from the web page [30]). The data consist of 20 curves (one per year/city) with 12 observation points per curve where we are interested in analyzing the possible pattern of spatial correlation among cities in relation to their temperatures. Finally, the third data set is part of the original data from the web page [31]. It consists of five groups of phonemes SH, IY, DCL, AA, and AO; each group contains 400 log-periodograms (functions) discretized in 150 frequencies (points). Each of the log-periodograms corresponds to a different speaker. In this example, we look for possible associated phonemes. These three data sets have been extensively used in the literature in functional data analysis (Epifanio-López [7], Jacques and Preda [15], Li and Yu [18], López-Pintado and Romo [19], [20], Valencia et al.[29]), and in particular, Epifanio-López [7], Li and Yu [18], for other purposes such as classification.

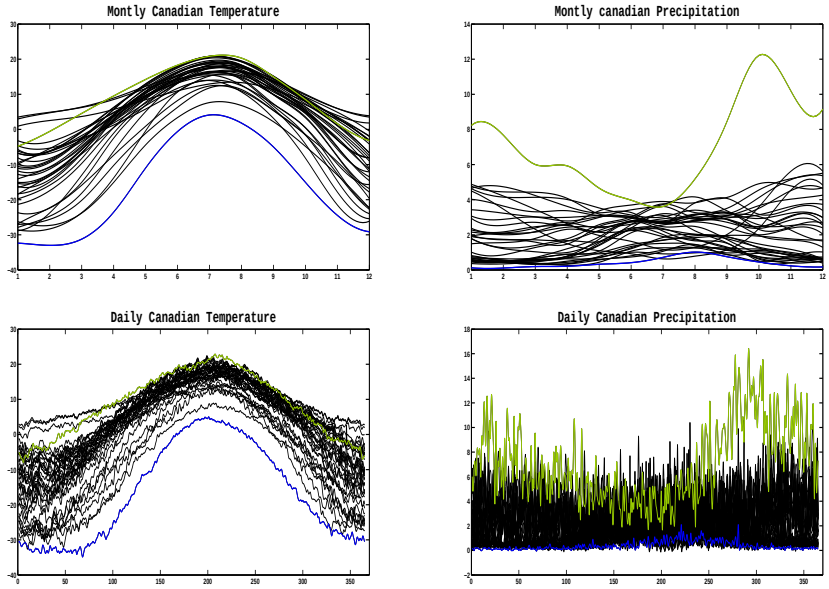

Figure 5: Monthly and daily temperature and precipitation of Canada.

Table 9: Association test for temperature and precipitation data

Data 1	Data 2	$\widehat{\rho}_s IL$	p-$value$	Decision	$\widehat{\tau}_1$	p-$value$	$\widehat{\tau}_2$	p-$value$
Daily temperature	Daily precipitation	**0.6043**	**0.0002**	reject H_0	0.0807	0.5050	**0.4958**	**0**
Monthly temperature	Monthly precipitation	**0.5764**	**0.0004**	reject H_0	0.1378	0.2438	**0.4622**	**0.0002**
Montreal	Resolute	**0.6041**	**0.0050**	reject H_0	0.2368	0.1468	**0.3316**	**0.0394**
Montreal	Prince Rupert	-0.0612	0.7940	accept H_0	0.0632	0.6914	-0.026	0.883
Montreal	Fort San John	0.1160	0.6220	accept H_0	-0.0579	0.7398	0.0684	0.6902
Resolute	Prince Rupert	-0.1836	0.4322	accept H_0	-0.2316	0.1620	-0.1158	0.4850
Resolute	Fort San John	0.0168	0.9516	accept H_0	0.0895	0.59	0.0316	0.8668
Prince Rupert	Fort San John	0.3280	0.1474	accept H_0	0.0842	0.6092	0.1789	0.2780

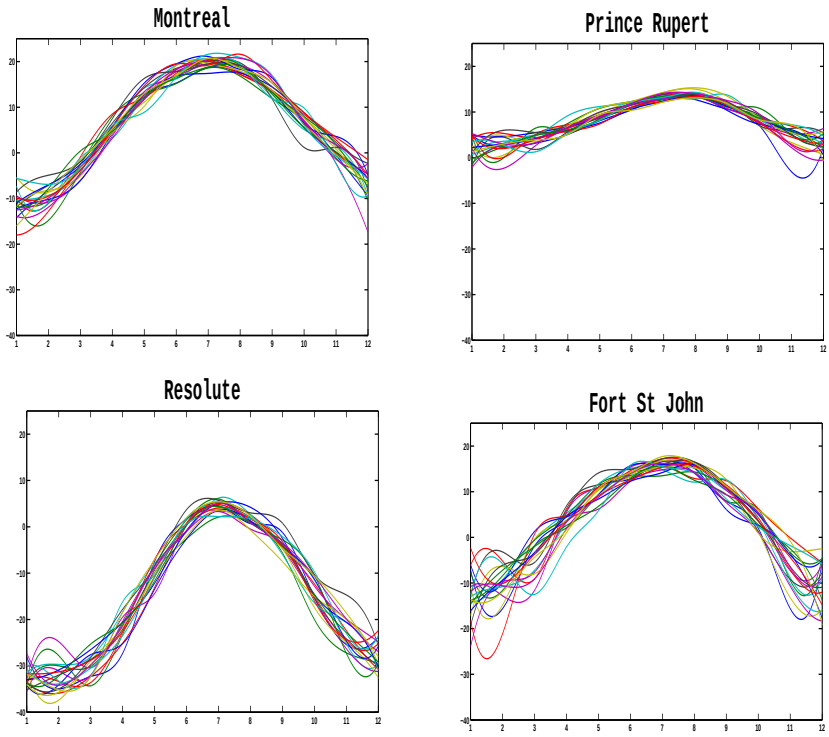

Figure 6: Temperatures of 4 cities in Canada.

Figure 5 shows monthly and daily data of temperature and precipitation in Canada. Green curves are the highest and the blue curves are the smallest in the sense of the IL_n-*grade* ordering. Table 9 shows the values of the Spearman coefficient $\widehat{\rho}_s IL$, the p-value related to the association test with 10000 bootstrap samples and the corresponding decision with $\alpha = 0.05$. The association test for the other coefficients is also shown in Table 9. As we can see, the null hypothesis is rejected for all cases except when using $\widehat{\tau}_1$. Remember that $\widehat{\tau}_1$ is based on the pre-order induced by the maximum of the curves, which is more sensitive to outliers and reflects worse than the other pre-order a summary of the curves shapes. Therefore, we can say that the temperature and the precipitation in Canada have a significant association, which was expected because they are strongly linked to climatological phenomena.

In relation to the data sets of Canadian cities, only Montreal and Resolute present significant dependence for both Spearman's and Kendall's τ (with pre-order of the integral) coefficients. We have tried to find a physical explanation for this fact but these two cities do not share the same kind of weather, nor do they have a similar latitude or other factors that directly relate them, so the significant dependence may be due to the similarity with respect to shape and position of the curves per year (see Figure 6). However, the positive association between Montreal and Resolute does not hold when we pass the same test with $\widehat{\tau}_1$. Hence, spatial correlation is not observed for these four cities.

Table 10 shows the results of the association tests for the phoneme data. We include also the p-values for each test. Note that, in general, the dependence between the phonemes is very small for all measures, being only statistically significative for the phonemes AA and SH with the coefficients $\widehat{\rho}_s IL$, $\widehat{\tau}_1$, and $\widehat{\tau}_2$. This may be due to the position and shape of the curves. We can see that the shape of the curves of the phoneme SH is in general different when compared to other phonemes. Indeed, it can be easily observed that a certain negative dependence could exist (see Figure 7). This fact is reflected in the sign of the coefficients since they are negative in most cases where the phoneme SH is evaluated. It can be seen that in this case, the shape of the two pairs of curves exhibits opposite behavior.

Table 10: Phoneme Data

Phoneme 1	Phoneme 2	$\widehat{\rho}_s IL$	*p-value*	Decision	$\widehat{\tau}_1$	*p-value*	$\widehat{\tau}_2$	*p-value*
AA	AO	0.078	0.1144	accept H_0	0.0257	0.4536	0.0604	0.0692
AA	SH	**−0.100**	**0.0464**	reject H_0	**−0.0675**	**0.048**	**−0.0763**	**0.0192**
AA	IY	0.058	0.2664	accept H_0	0.0004	0.9624	0.0459	0.1504
AA	DCL	0.010	0.791	accept H_0	−0.0186	0.6056	0.003	0.9174
AO	SH	−0.040	0.422	accept H_0	0.0079	0.8088	−0.0245	0.4744
AO	IY	0.010	0.845	accept H_0	0.0386	0.2336	0.0086	0.7944
AO	DCL	−0.020	0.696	accept H_0	−0.0053	0.8920	0.00045	0.9840
SH	IY	−0.025	0.64	accept H_0	−0.0479	0.1592	−0.0179	0.5832
SH	DCL	0.027	0.547	accept H_0	0.0188	0.5832	0.0109	0.7616
IY	DCL	−0.019	0.691	accept H_0	0.0271	0.4256	−0.0079	0.8320

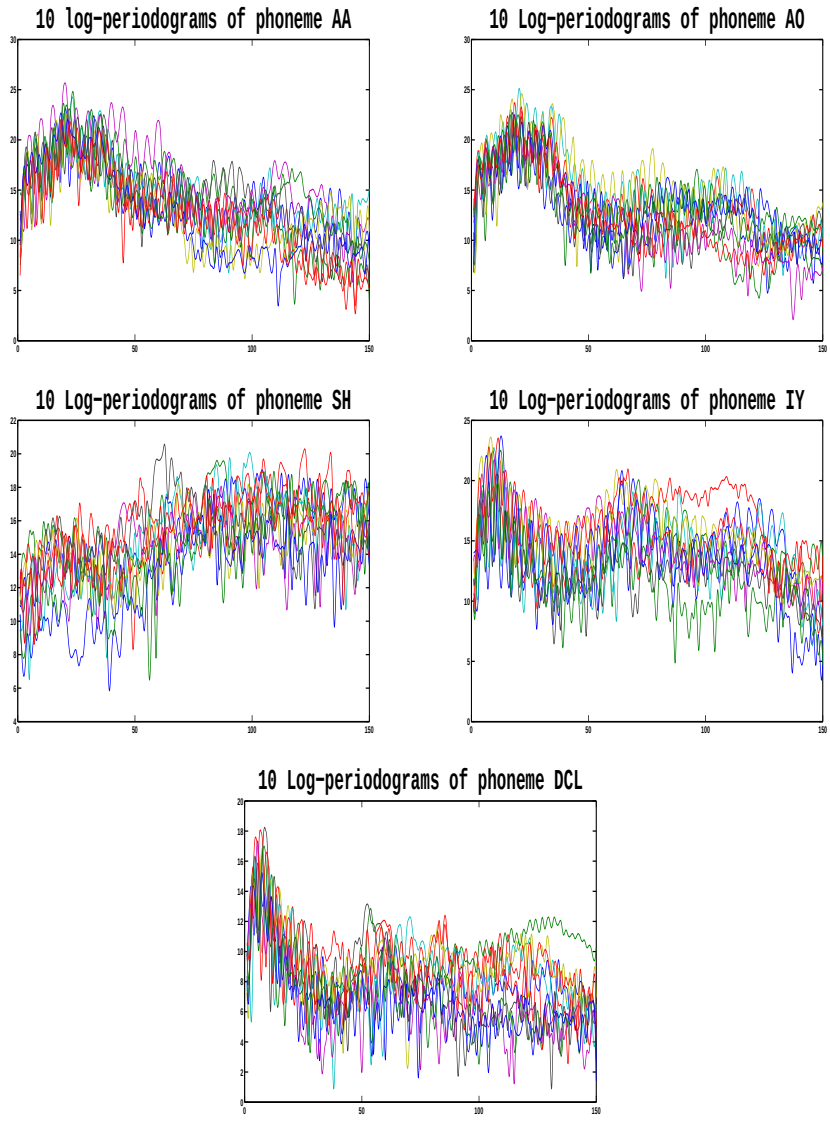

Figure 7: Log-periodograms of phonemes AA, AO, SH, IY and DCL.

8 Conclusions

We have introduced a new association coefficient to measure dependence between functions when a bivariate sample of functional data is considered. Specifically, a natural extension of the usual Spearman coefficient is provided by ranking the functions using two kinds of ordering for the curves: the Inferior Length and the Superior Length. These orderings among curves allowed us to adapt the definition of grade examined in Nelsen [22] but for the functional context and so, Spearman's coefficient can be defined as usual is, the Pearson correlation among grades. We have also proved that Spearman's coefficient has a good theoretical and practical properties. The simulation study and real examples provided in the paper show the good performance of the Spearman coefficient as well as its robustness.

We have also introduced a bootstrap independence test to assess the significance of the association between two groups of curves. Tests of this type also allow us to quantify the statistical significance of some conjectures made on the basis of exploratory analysis. We have illustrated with simulated data the power of this test.

We focused in this paper on an univariate dependence measure, but it could be of interest to explore other possible options such as a functional dependence measure as an alternative to the functional correlation introduced in Ramsay and Silverman [26]. In addition, note how all the univariate dependence measures are linked to a curves ordering. Thus, other pre-orders for curves can provide alternative dependence measures that can be useful for visualizing association in data sets.

Acknowledgment

This research has been supported by Spanish Ministry of Science and Innovation project (ECO2011-25706) and project (SES2007)

References

[1] Berrendero, J. Justel, A. and Svarc, M. (2011). Principal components for multivariate functional data. *Computational Statistics and Data Analysis* 55, pp 2619-263.

[2] Cuevas, A. Febrero, M. and Fraiman, R. (2007). Robust estimation and classication for functional data via projection-based depth notions. *Computational Statistics* 22, pp 481-496.

[3] Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. New York, Cambridge University Press.

[4] Dubin, J. A. and Müller, H. G. (2005). Dynamical Correlation for Multivariate Longitudinal Data. *Journal of the American Statistical Association* 100, pp 872-881.

[5] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *the Annals of Statistics* 7(1), pp 1-26.

[6] Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. New York, Chapman and Hall.

[7] Epifanio-López, I. (2008). Shape Descriptors for classification of functional data. *Technometrics* 50(3), pp 284-294.

[8] Fraiman, R. and Muniz, C. (2001). Trimmed Means for Functional Data. *Test* 10(2), pp 419-440.

[9] Ferraty, F. and Vieu P. (2006). *Nonparametric Functional Data Analysis : Theory and Practice*. New York, Springer.

[10] Gibbons, J.D. (1993). *Nonparametric Measures of Association*. Newbury Park (California), Sage Publications.

[11] Hauke, J. and Kossowski, T. (2007). Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data. Proceedings of the MAT TRIAD Conference, Bedlewo, Poland.

[12] He, G. Müller, H.G. and Wang, J. L. (2000). Extending Correlation and Regression from Multivariate to Functional Data. *Asymptotics in Statistics and Probability*, pp 197-210.

[13] He, G. Müller, H.G. Wang, J.L. (2004). Methods of canonical analysis for functional data. *Journal of Statistical Planning and Inference* 122, pp 141-159

[14] Hoeffding, W. (1948). A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics* 19(3), pp 293-325.

[15] Jacques, J. and Preda, C. (2012). Model-based clustering of multivariate functional data. Preprint HAL n°00713334, in revision for *Computational Statistics and Data Analysis*.

[16] Leurgans, S.E., Moyeed, R.A., and Silverman, B.W. (1993). Canonical correlation analysis when data are curves. *Journal of the Royal Statistical Society B* 55, pp 725-740.

[17] Li, R. and Chow, M. (2005). Evaluation of reproducibility for paired functional data. *Journal of Multivariate Analysis* 93, pp 81-101.

[18] Li, B. and Yu, Q. (2008). Classification of functional data: A segmentation approach. *Computational Statistics and Data Analysis* 52(10), pp 4790-4800.

[19] López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *Journal of the American Statistical Association* 104, pp 718-734.

[20] López-Pintado, S. and Romo, J. (2011). A half-region depth for functional data. *Computational Statistics and Data Analysis* 55, pp 1679-1695.

[21] Martín-Barragan, B. Lillo, R. and Romo, J. (2012) Functional Boxplots based on Half-regions. Submitted.

[22] Nelsen, R. B. (2006). *An Introduction to Copulas 2nd Edition*. New York, Springer Series in Statistics.

[23] Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Royal Society* 187, pp 253-318.

[24] Opgen-Rhein, R. and Strimmer, K. (2006). Inferring Gene Dependency Networks from Genomic Longitudinal Data: a Functional data Approach. *REVSTAT* 4, pp 53-65.

[25] Ramsay, J. O. Silverman, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. New York, Springer.

[26] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis 2nd Edition*. New York, Springer Verlag.

[27] Wilcox, R. R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*, 3ra edition. San Diego, Elsevier.

[28] Xu, W. Hou, Y. Hung, Y. and Zou, Y. (2010). Comparison of Spearman's rho and Kendall's tau in Normal and Contaminated Normal Models. arXiv:1011.2009v1 [cs.IT] 9 Nov 2010.

[29] Valencia, D. Lillo, R. and Romo J. (2013). A Kendall correlation coefficient for functional dependence. *Working paper*. Statistics and Econometrics Series 28. Universidad Carlos III de Madrid.

[30] http://www.tutiempo.net/clima/Canada/CA.html

[31] http://www-stat.stanford.edu/ tibs/ElemStatLearn/