MOX–Report No. 3/2015

Distribution-Free Interval-Wise Inference for Functional-on-Scalar Linear Models

Abramowicz, K.; de Luna, S.; Hger, C.; Pini, A.; Schelin, L.; Vantini, S.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Distribution-Free Interval-Wise Inference for Functional-on-Scalar Linear Models

K.Abramowicz[a], S. de Luna[a], C. Häger[b], A. Pini[c], L. Schelin[d], S. Vantini[c]

[a] Department of Mathematical Statistics, Umeå University, Umeå Sweden

[b] Department of Community Medicine and Rehabilitation, Physiotherapy, Umeå University, Umeå Sweden

[c] MOX– Department of Mathematics, Politecnico di Milano, Milan Italy

[d] Department of Statistics, Umeå University, Umeå Sweden

konrad.abramowicz@math.umu.se
sara.de.luna@math.umu.se
charlotte.hager@umu.se
alessia.pini@polimi.it
lina.schelin@math.umu.se
simone.vantini@polimi.it

Abstract

We introduce a distribution-free procedure for testing a functional-on-scalar linear model with fixed effects. The procedure does not only test the global hypothesis on all the domain, but also selects the intervals where statistically significant effects are detected. We prove that the proposed tests are provided with an asymptotic interval-wise control of the family-wise error rate, i.e., the probability of falsely rejecting any interval of true null hypotheses. The procedure is then applied to one-leg hop data from a study on anterior cruciate ligament injury. We compare knee kinematics of three groups of individuals, taking individual-specific covariates into account.
Keywords: Functional Data, Permutation Test, Interval Testing Procedure, ANCOVA.

1 Introduction

Functional data analysis (FDA) is a relatively new, dynamically developing, research area within the field of statistics. In recent literature, linear models for functional data have been widely studied (see, e.g., Fan and Zhang 2000; Abramovich and Angelini 2006; Cardot et al. 2007; Reiss et al. 2010; Gertheiss et al. 2013; Abramowicz et al. 2014).

Motivated by the analysis of the dependence of knee kinematics on subject-specific covariates, in this paper we consider a functional-on-scalar linear model. In detail, we model a functional response with a set of covariates multiplied by functional parameters. Such model finds its application in a wide range of research fields where modern techniques enable collection of high-resolution data. In this context, many of the empirically relevant questions do not only address the effect of covariates on a functional response, but also require identification of significant domain subsets. In this work, data from a follow-up study after anterior cruciate ligament injury are analyzed. We study knee kinematics during one-leg hop, comparing individuals treated with physiotherapy or surgery with healthy controls. Previous studies suggest a difference in the movement patterns between the groups (Tengman et al. 2014; Abramowicz et al. 2014). In this paper, we investigate if this difference is only due to group effect, or if it can be explained by means of additional individual-specific covariates such as jump length, gender, age, and body mass index (BMI).

We focus on a distribution-free method and therefore propose to use a least squares method for parameter estimation. Parameter estimation of the functional model is handled by first representing the functional response and the functional regression parameters in terms of a suitable functional basis. The functional estimation problem is thus decomposed into a family of corresponding linear models of univariate response variables, one for each of the coefficients (components) of the basis expansion. Hence, least squares estimation methods for linear models with univariate response variables can be used to estimate the functional linear model (see Section 2).

Forming valid tests of various hypotheses about the functional regression parameters, with control of the error rate, is not straightforward. One solution adopted in the literature is to develop global tests for the parameters of the model. Such tests investigate if a covariate has a significant effect on the response, but does not provide any domain selection (Cuevas et al. 2004; Abramovich and Angelini 2006; Antoniadis and Sapatinas 2007; Cardot et al. 2007; Schott 2007; Cuesta-Albertos and Febrero-Bande 2010; Zhang and Liang 2014). Another approach, proposed in Fan and Zhang (2000); Reiss et al. (2010); Ramsay and Silverman (2005), is to provide point-wise confidence bands for the functional parameters. The results indicate in which parts of the domain the covariates have an effect, but not at which significance level. As clearly discussed in Ramsay and Silverman (2005, pp. 243–244), point-wise limits are not equivalent to confidence regions for the entire estimated curves. Assuming that data are expressed through a functional basis, inference can be based directly on the expansion coefficients, as proposed by Spitzner et al. (2003). In the latter work, single-component tests are performed, and their p-values adjusted with the Bonferroni-Holm procedure (Holm 1979). In this way, results are compensated for the many dependent tests performed on the same data set. A drawback with this procedure is that it is typically too conservative, and needs a relevant dimensional reduction of data in order to detect significant functional parameters.

In our work, we follow the same line of research proposed by Spitzner et al. (2003), introducing a less conservative p-values adjustment, which rely on the properties of functional data, and does not require any dimensional reduction of the functional data set. The continuous nature of functional models expressed in terms of a basis expansion such as B-splines, typically implies that neighbouring basis coefficients present a positive dependence. Combinations of neighbouring component-wise (dependent) tests thus have the potential to more easily detect parts of the domain where a functional regression parameter is significantly different from zero. Therefore, in our paper we restrict multiple comparisons of component-wise tests to intervals of neighbouring components and use the Interval Testing Procedure (ITP) introduced by Pini and Vantini (2013), which is based on single- and multiple-component tests. The single-component tests are based on Freedman and Lane permutation schemes (Freedman and Lane 1983), which do not rely on any distributional assumptions. Further, we use a Non-Parametric Combination (NPC) procedure to obtain simultaneous tests on intervals of components. The NPC procedure is a computationally efficient procedure which preserves the exactness and consistency properties of single-component tests. For further details, we refer to Pesarin and Salmaso (2010). Using the ITP, for each basis component we obtain an adjusted p-value, which is used to select the significant component intervals. Such tests are provided with an interval-wise control of the Family Wise Error Rate (FWER). In detail, this control implies that the probability of falsely rejecting any interval of basis components associated to true null hypotheses is controlled at the desired significance level. We prove that the proposed tests are exact or asymptotically exact.

The paper is outlined as follows: in Section 2, we describe the functional-on-scalar linear model, discussing the methodology proposed for functional parameter estimation and inference. Section 3 reports the theoretical properties of the proposed methodology. The proofs of theorems in Section 3 are reported in Appendix A. In Section 4, we report the results of the analysis of kinematic data. Finally, Appendix B reports some details on the Freedman and Lane permutation scheme, while Appendix C briefly describes the NPC procedure. All computations and plots have been created using R (R Core Team 2014).

2 Methodology

2.1 The functional-on-scalar linear model

Suppose we have observed a sample of n continuous random functions $\{y_i(t)\}_{i=1,...,n}$, over time t: $t \in [a, b]$. We want to study the following functional-on-scalar linear model:

$$y_i(t) = \beta_0(t) + \sum_{l=1}^{L} \beta_l(t) x_{li} + \varepsilon_i(t), \quad \forall i = 1, ..., n, \tag{1}$$

where $x_{1i}, ..., x_{Li} \in \mathbb{R}$ are known scalar covariates and $\beta_l(t)$, $l = 0, ..., L$, are the fixed functional regression parameters. The errors $\varepsilon_i(t)$, $t \in [a, b]$ are *i.i.d.* (with respect to units) zero-mean random functions (not necessarily Gaussian) with finite total variance, i.e.,

$$\int_a^b \mathbb{E}\left[\varepsilon_i(t)\right]^2 dt < \infty, \quad \forall i = 1, ..., n. \tag{2}$$

We assume that, for each $i = 1, \ldots, n$, $y_i(t)$ can be expressed in terms of basis functions $\{\phi^{(k)}(t)\}_{k=1}^p$, i.e.,

$$y_i(t) = \sum_{k=1}^{p} {y_i}^{(k)} \phi^{(k)}(t).$$

Whenever functional data are described through a basis expansion, we can perform inference directly on the set of coefficients representing the data. Therefore, we can project the model (1) on the functional space spanned by the basis:

$$\sum_{k=1}^{p} {y_i}^{(k)} \phi^{(k)}(t) = \sum_{k=1}^{p} {\beta_0}^{(k)} \phi^{(k)}(t) + \sum_{l=1}^{L} \sum_{k=1}^{p} {\beta_l}^{(k)} \phi^{(k)}(t) x_{li} + \sum_{k=1}^{p} {\varepsilon_i}^{(k)} \phi^{(k)}(t),$$

for all $t \in [a, b]$ and $i = 1, \ldots, n$, which leads to:

$$\sum_{k=1}^{p} \left[{y_i}^{(k)} - {\beta_0}^{(k)} - \sum_{l=1}^{L} {\beta_l}^{(k)} x_{li} - {\varepsilon_i}^{(k)} \right] \phi^{(k)}(t) = 0, \quad \forall t \in [a, b], \forall i \in 1, \ldots, n. \tag{3}$$

Since $\{\phi^{(k)}(t)\}_{k=1}^p$ is a basis, equation (3) holds if

$${y_i}^{(k)} = {\beta_0}^{(k)} + \sum_{l=1}^{L} {\beta_l}^{(k)} x_{li} + {\varepsilon_i}^{(k)}, \quad \forall k = 1, ..., p, \forall i \in 1, \ldots, n \tag{4}$$

holds. Therefore, we can express model (1) as a family of p scalar-on-scalar linear models, with errors pertaining to the same sample unit i possibly dependent. Moreover, we have that:

$$0 = \mathbb{E}[\varepsilon_i(t)] = \mathbb{E}\left[\sum_{k=1}^{p} {\varepsilon_i}^{(k)} \phi^{(k)}(t)\right] = \sum_{k=1}^{p} \mathbb{E}[{\varepsilon_i}^{(k)}] \phi^{(k)}(t) \qquad \forall t \in [a, b], \forall i \in 1, \ldots, n,$$

and hence $\mathbb{E}[\varepsilon_i^{(k)}] = 0$ for all $k = 1, ..., p$ and $i = 1, \ldots, n$. From (2) and the fact that $\{\phi^{(k)}(t)\}_{k=1}^p$ is a basis, we also have that for $k = 1, \ldots, p$, $\mathbb{E}[{\varepsilon_i^{(k)}}^2] < \infty$. Finally, the independence of the random functions $\varepsilon_i(t)$, $t \in [a, b]$, implies independence across units of the coefficients ${\varepsilon_i}^{(k)}$. Therefore, for fixed k, the error terms ${\varepsilon_i}^{(k)}$, $i = 1, \ldots, n$ are *i.i.d.* zero-mean random variables with finite variance. Note that we are not making assumptions instead on the auto-covariance structure of the $\varepsilon_i(t)$'s. Hence, for fixed i, the errors $\varepsilon_i^{(k)}$, $k = 1, \ldots, p$, are not assumed independent.

In practice, we often can not observe the complete response functions $y_i(t)$, $i = 1, \ldots, n$, and need to estimate them based on the finite number of observations. We refer to Ramsay and Silverman (2005) for a discussion about the choice of basis used to represent data and methods used to estimate the coefficients.

2.2 Model estimation

The ordinary least squares (OLS) estimators of the functional parameters $\beta_l(t)$, $l = 0, \ldots, L$, can be found by minimizing the sum over units of the L^2 distances between the functional data $y_i(t)$ and the quantity $\beta_0(t) + \sum_{l=1}^{L} \beta_l(t) x_{li}$ with respect to $\beta_l(t)$, $l = 0, \ldots, L$ (Ramsay and Silverman 2005):

$$\sum_{i=1}^{n} \int_a^b \left(y_i(t) - \beta_0(t) - \sum_{l=1}^{L} \beta_l(t) x_{li} \right)^2 \mathrm{d}t. \tag{5}$$

The minimization can be done separately for each coefficient of the basis expansion, even in presence of non-orthonormal basis components. Indeed, when using a basis expansion, (5) can be written as:

$$\sum_{i=1}^{n} \int_a^b \left[\sum_{k=1}^{p} \left(y_i^{(k)} - \boldsymbol{\beta}^{(k)\prime} \mathbf{x}_i \right) \phi^{(k)}(t) \right]^2 \mathrm{d}t, \tag{6}$$

where $\boldsymbol{\beta}^{(k)} = (\beta_0^{(k)}, \ldots, \beta_L^{(k)})'$ and $\mathbf{x}_i$ is the i-th row of the design matrix $X_n \in \mathbb{R}^{(n \times (L+1))}$ ($[X_n]_{i,1} = 1, \forall i = 1, \ldots, n$; $[X_n]_{i,j} = x_{j-1,i}$, $i = 1, \ldots, n$, $j = 2, \ldots, L+1$). Equation (6) is equivalent to:

$$\sum_{i=1}^{n} \sum_{k_1=1}^{p} \sum_{k_2=1}^{p} \left(y_i^{(k_1)} - \boldsymbol{\beta}^{(k_1)\prime} \mathbf{x}_i \right) \left(y_i^{(k_2)} - \boldsymbol{\beta}^{(k_2)\prime} \mathbf{x}_i \right) \int_a^b \phi^{(k_1)}(t) \phi^{(k_2)}(t) \mathrm{d}t,$$

which can be written using matrix notation as

$$\sum_{i=1}^{n} \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right)' W \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right), \tag{7}$$

where $\mathbf{y}_i = (y_i^{(1)}, \ldots, y_i^{(p)})' \in \mathbb{R}^p$, $\boldsymbol{\beta} \in \mathbb{R}^{((L+1) \times p)}$ is the matrix of coefficients, $[\boldsymbol{\beta}]_{l,k} = \beta_l^{(k)}$, and $W \in \mathbb{R}^{p \times p}$ is the matrix of inner products between basis functions $[W]_{k_1 k_2} = \int_a^b \phi(t)^{(k_1)} \phi(t)^{(k_2)} \mathrm{d}t$. As shown by Johnson and Wichern (2007), for any positive definite matrix W, we have that:

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum_{i=1}^{n} \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right)' W \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right) = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum_{i=1}^{n} \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right)' \left(\mathbf{y}_i - \boldsymbol{\beta}' \mathbf{x}_i \right)$$

that is, in the minimization, W can be replaced with the identity. Note that:

$$\sum_{i=1}^{n} \left(\mathbf{y}_i - \boldsymbol{\beta}'\mathbf{x}_i\right)'\left(\mathbf{y}_i - \boldsymbol{\beta}'\mathbf{x}_i\right) = \sum_{k=1}^{p}\sum_{i=1}^{n}\left(y_i^{(k)} - \boldsymbol{\beta}^{(k)'}\mathbf{x}_i\right)^2$$

and hence the minimization problem on the left hand side with respect to $\boldsymbol{\beta}$ is reduced to the family of p independent minimization problems, one for each component $k = 1, \ldots, p$. For each k, $\sum_{i=1}^{n}\left(y_i^{(k)} - \boldsymbol{\beta}^{(k)'}\mathbf{x}_i\right)^2$ is minimized by the OLS estimate $\hat{\boldsymbol{\beta}}^{(k)} = (\hat{\beta}_0^{(k)}, \ldots, \hat{\beta}_L^{(k)})$ of $\boldsymbol{\beta}^{(k)}$. Therefore $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}^{(1)}, \ldots, \hat{\boldsymbol{\beta}}^{(p)})$ is also the global OLS estimate minimizing (6). Hence, for each $l = 0, \ldots, L$, the estimate of the functional regression parameters $\beta_l(t)$ is

$$\hat{\beta}_l(t) = \sum_{k=1}^{p} \hat{\beta}_l^{(k)} \phi^{(k)}(t). \tag{8}$$

It is possible to establish asymptotic properties for the OLS estimates on each basis component k. Consider the following standard conditions:

C1 The matrix $X_m'X_m$ is non-singular for some $m \geq 1$ (implying that it is non-singular for all $n \geq m$), and the inverse $V = (X_n'X_n)^{-1}$ is s.t. the elements $[V]_{ij} \to 0$ as $n \to \infty$, for all $i, j = 1, ..., L+1$.

C2 For each $k = 1, ..., p$, the regression errors $\varepsilon_i^{(k)}$ satisfy:

$$\sup_{i=1,\ldots,n} \mathbb{E}\left[\varepsilon_i^{(k)^2}\right] < \infty.$$

Under conditions *C1*-*C2*, we have that for each $k = 1, \ldots, p$, the obtained OLS estimates $\hat{\beta}_0^{(k)}, \ldots, \hat{\beta}_L^{(k)}$ are strongly consistent estimates of $\beta_0^{(k)}, \ldots, \beta_L^{(k)}$ (Lai et al. 1979). Condition *C1* is a sufficient condition for finding an explicit expression of the OLS estimates, and guarantees convergence in probability. Condition *C2* assures almost sure convergence.

2.3 Model inference

One of the main challenges with inference for functional linear model (1) is performing valid tests of various hypotheses on the functional regression parameters. Analogously to the classical framework, we are, e.g., interested in testing if none of the covariates significantly affects the response, i.e., the functional version of classical F-test:

$$\begin{cases} H_{0,F} : \beta_l(t) = 0 & \forall l \in 1, \ldots, L,\ \forall t \in [a,b] \\ H_{1,F} : \beta_l(t) \neq 0 & \text{for some } l \in \{1, \ldots, L\} \text{ and } t \in [a,b] \end{cases} \tag{9}$$

together with tests of significance for any specific functional parameter $l \in \{0, \ldots, L\}$, i.e., the functional version of classical t-test:

$$\begin{cases} H_{0,l} : \beta_l(t) = 0 \quad \forall t \in [a, b] \\ H_{1,l} : \beta_l(t) \neq 0 \quad \text{for some } t \in [a, b]. \end{cases} \tag{10}$$

In the most general case we are interested in testing hypotheses on linear combinations of the functional parameters of the regression, specified by a combination matrix C. In detail, $C \in \mathbb{R}^{(q \times (L+1))}$ is any real-valued full rank matrix, where q denotes the number of hypotheses on the functional regression parameters to be jointly tested, with $1 \leq q \leq L + 1$. Moreover, let $\mathbf{c}_0(t) = (c_{01}(t), ..., c_{0q}(t))'$ be a vector of fixed functions from the space spanned by the basis functions $\{\phi^{(k)}(t)\}_{k=1,...,p}$, and let $\boldsymbol{\beta}(t) = (\beta_0(t), ..., \beta_L(t))'$ denote the vector of functional regression parameters. We are in general interested in testing hypotheses of the form:

$$\begin{cases} H_{0,C} : C\boldsymbol{\beta}(t) = \mathbf{c}_0(t) \quad \forall t \in [a, b] \\ H_{1,C} : C\boldsymbol{\beta}(t) \neq \mathbf{c}_0(t) \quad \text{for some } t \in [a, b]. \end{cases} \tag{11}$$

where the j-th element of vector $C\boldsymbol{\beta}(t)$ is a function obtained by means of a linear combination of the functional regression parameters $\beta_l(t)$ with weights $[C]_{jl}$: $[C\boldsymbol{\beta}(\boldsymbol{t})]_j = \sum_{l=0}^{L} [C]_{jl}\beta_l(t)$, $j = 1, ..., q$. There are two important special cases of the general functional linear hypotheses (11):

1. When $q = L$, $C = C_F = (\mathbf{0} | I_L) \in \mathbb{R}^{L \times (L+1)}$, and $\mathbf{c_0}(\mathbf{t}) = \mathbf{0} \in \mathbb{R}^L$, where I_L is the $L \times L$ identity matrix. Then, the hypotheses of (11) correspond to the hypotheses in (9);

2. For a fixed l, let $q = 1$, $C = C_l \in \mathbb{R}^{1 \times (L+1)}$ with $[C_l]_r = 1$ if $r = l$ and 0 otherwise, and $c(t) = 0$. Then, the hypotheses in (11) correspond to the hypotheses in (10).

By using the basis representation of $\boldsymbol{\beta}(t)$ and $\mathbf{c}_0(t)$, functional linear hypotheses (11) are translated into a family of p linear hypotheses pertaining the components of the basis expansion $k = 1, \ldots, p$:

$$\begin{cases} H_{0,C} = \bigcap_{k=1}^{p} H_{0,C}^{(k)}, \text{ with } H_{0,C}^{(k)} : C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)} \\ H_{1,C} = \bigcap_{k=1}^{p} H_{1,C}^{(k)}, \text{ with } H_{1,C}^{(k)} : C\boldsymbol{\beta}^{(k)} \neq \mathbf{c}_0^{(k)}, \end{cases} \tag{12}$$

where $\mathbf{c}_0^{(k)} \in \mathbb{R}^q$ is a vector composed by the k-th coefficients of the basis expansion of vector $\mathbf{c}_0(t)$ with the basis expansion performed for each element of the vector. Hence, as opposed to model estimation, the problem of inference for multiple components is not straightforward, as it involves a, possibly high-dimensional, family of dependent statistical tests.

The hypotheses on the functional parameters of equations (9)-(11) can thus be tested by performing multiple hypotheses tests on single basis components $\boldsymbol{\beta}^{(k)}$, testing:

$$H_{(0,C)}^{(k)} : C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)} \text{ versus } H_{(1,C)}^{(k)} : C\boldsymbol{\beta}^{(k)} \neq \mathbf{c}_0^{(k)}, \tag{13}$$

for various choices of C and $\mathbf{c}_0^{(k)}$. Based on the OLS estimates of the $\beta_l^{(k)}$'s it is rather straightforward to form natural test statistics such as F-tests and t-tests to test (families) of hypotheses (13), see Subsections 2.1-2.2. The challenge is to have control of the family-wise error rate arising from the many multiple (dependent) hypotheses tests.

In this paper we use the ITP to control on each interval the probability of falsely rejecting at least one true null hypothesis of the family, i.e., the interval-wise control of the FWER. The ITP is a three step procedure involving a basis expansion of functional data, the testing of each multiple-component hypothesis pertaining intervals of basis coefficients, and a multiplicity correction providing an interval-wise control of the FWER. In the following paragraphs we first give some details on the starting point of the ITP, that is the single-component testing, and then we describe how we construct the multiple-component tests and the corresponding p-value correction needed to keep control of the FWER..

2.3.1 Single-component testing

For hypotesis testing of (13) corresponding to a single component k, we use permutation tests based on the Freedman and Lane permutation scheme (Freedman and Lane 1983), which is briefly described in Appendix B. This permutation strategy is the most commonly used for linear models, and presents many advantages compared to other permutation techniques (Davison and Hinkley 1997; Anderson and Legendre 1999; Anderson and Robinson 2001; Zeng et al. 2011; Winkler et al. 2014). In particular, it can be shown empirically that its power is typically higher than the power of tests based on other permutation schemes (Anderson and Legendre 1999; Winkler et al. 2014). Permutation tests based on the Freedman and Lane scheme are based on permutations of the estimated residuals under the reduced model (i.e., the linear model under the null hypothesis of the test).

For the fixed component k, to perform the F-test with hypotheses

$$\begin{cases} H_{0,F}^{(k)} : \beta_l^{(k)} = 0 & \forall l \in 1, \dots, L \\ H_{1,F}^{(k)} : \beta_l^{(k)} \neq 0 & \text{for some } l \in \{1, \dots, L\} \end{cases} \tag{14}$$

we use the F-test statistic:

$$T_F^{(k)} = \frac{(n-L)\sum_{i=1}^{n}(\hat{y}_i^{(k)} - \bar{y}^{(k)})^2}{L\sum_{i=1}^{n}(y_i^{(k)} - \hat{y}_i^{(k)})^2}, \tag{15}$$

where $\hat{y}_i^{(k)} = \hat{\beta}_0^{(k)} + \sum_{l=1}^{L} \hat{\beta}_l^{(k)} x_{li}$, $\hat{\beta}_0^{(k)}, ..., \hat{\beta}_L^{(k)}$ being the OLS estimates when all L covariates are in the model, and $\bar{y}^{(k)} = \sum_{i=1}^{n} {y_i}^{(k)}/n$ is the sample mean of the response coefficients.

To perform the t-test for the l-th functional regression parameter, i.e., to test

$$\begin{cases} H_{0,l}^{(k)} : \beta_l^{(k)} = 0 \\ H_{1,l}^{(k)} : \beta_l^{(k)} \neq 0 \end{cases} \tag{16}$$

we use the absolute value of the t-test statistic:

$$T_{t,l}^{(k)} = \left| \frac{\hat{\beta}_l^{(k)}}{\mathrm{se}(\hat{\beta}_l^{(k)})} \right|, \tag{17}$$

where $\mathrm{se}(\hat{\beta}_l^{(k)})$ is the standard error of $\hat{\beta}_l^{(k)}$, being $\hat{\beta}_l^{(k)}$ the OLS estimate of $\beta_l^{(k)}$ when all covariates are in the model. As shown in Pesarin and Salmaso (2010), the test statistic (17) is permutationally equivalent to the squared partial correlation coefficient, commonly used in the literature of permutation tests for linear models (see, for instance, Anderson and Robinson 2001).

More generally, to perform single-component tests of linear hypotheses in (13), we can use the statistic:

$$T_C^{(k)} = \frac{1}{s^2} \left(C\hat{\boldsymbol{\beta}}^{(k)} - \mathbf{c}_0^{(k)} \right)' \left(C(X_n' X_n)^{-1} C' \right)^{-1} \left(C\hat{\boldsymbol{\beta}}^{(k)} - \mathbf{c}_0^{(k)} \right), \tag{18}$$

where $\hat{\boldsymbol{\beta}}^{(k)}$ is the OLS estimate of $\boldsymbol{\beta}^{(k)}$, and $s^2 = (\mathbf{y}^{(k)} - X_n\hat{\boldsymbol{\beta}})'(\mathbf{y}^{(k)} - X_n\hat{\boldsymbol{\beta}})/(n - L + 1)$ is the estimate of the variance of residuals at component k, with $\mathbf{y}^{(k)} = (y_1^{(k)}, \ldots, y_n^{(k)})'$.

2.3.2 Multiple tests and p-value correction

The p-values of single-component tests need to be adjusted to provide an interval-wise control of the FWER, according to the interval testing procedure (Pini and Vantini 2013). To perform this multiplicity correction, each multivariate hypothesis on intervals of components $\mathcal{I} = \{k_1, k_1 + 1, \ldots, k_2\}$ with $1 \leq k_1 < k_2 \leq p$ need to be tested. In detail, we need to test each hypothesis

$$H_{0,C}^{\mathcal{I}} = \bigcap_{k \in \mathcal{I}} H_{0,C}^{(k)}.$$

Such tests can be approached exploiting the NPC procedure (Pesarin and Salmaso 2010), which is briefly described in Appendix C. The NPC is a procedure that enables to build multivariate permutation tests by means of combining synchronized univariate permutation tests. The procedure applies in presence of dependence between univariate tests, which is the case in FDA.

Let $\lambda_C^{\mathcal{I}}$ denote the p-value corresponding to the multivariate test on hypothesis $H_{0,C}^{\mathcal{I}}$. The adjusted p-value $\lambda_{ITP,C}^{(k)}$ for the k-th component is then computed as the maximum between all p-values of univariate and multivariate tests containing that component, i.e.:

$$\lambda_{ITP,C}^{(k)} = \max_{\mathcal{I} \ni k} \lambda_C^{\mathcal{I}}.$$

The adjusted p-values can be used to select only the basis components leading to the rejection of the null hypothesis $H_{0,C}^{(k)}$, i.e., the ones with associated adjusted p-value lower than the desired significance level α.

It is important to point out that the ITP takes into account the dependence between the basis coefficients, which in the framework of a functional linear model means that it does not require to specify their covariance structure. Moreover, as tests are based on permutations, the procedure does not require the normality of residuals.

3 Theoretical results

In this section, we present theoretical properties of the inference on functional-on-scalar linear models performed along the line depicted in Section 2. All proofs are reported in Appendix A. The results are valid for the ITP, based on the NPC of tests using the Freedman and Lane scheme.

First, we prove that test of the family of linear hypotheses $\{H_{0,C}^{(k)}\}_{k=1,...,p}$ is provided with an asymptotic interval-wise control of the FWER. Pini and Vantini (2013) proved that, if all univariate and multivariate tests used to build the ITP are exact, the ITP based on the p components of any basis expansion is provided with an interval-wise control of the FWER. This result can be applied directly in the case of the F-test on the regression model, but has to be extended in the more general case of tests on linear hypotheses (including the t-tests on functional regression parameters), as in the latter the exactness of all tests is only asymptotical.

Theorem 3.1. *Under assumptions (C1-C2), the test of the family of linear hypotheses* (13) *based on the statistic* $T_C^{(k)}$ (18) *is provided with an asymptotic interval-wise control of the FWER. Formally, the ITP-adjusted p-values* $\lambda_{ITP,C}^{(k)}$*,* $\forall k = 1, ..., p$*, are s.t., for any interval* $\mathcal{I}$ *and any* $\alpha \in (0, 1]$*:*

$$\limsup_{n\to\infty} \mathbb{P}_{H_{0,C}^{\mathcal{I}}} \left[\exists k \in \mathcal{I} \text{ s.t. } \lambda_{ITP,C}^{(k)} \le \alpha \right] \le \alpha.$$

Since t-tests are specific cases of linear hypothesis, we obtain directly the following corollary.

Corollary 3.1. *Under assumptions (C1-C2), the test of the family of hypotheses* (16) *for the l-th functional regression parameter based on the statistic* $T_{t,l}^{(k)}$ (17) *is provided with an asymptotic interval-wise control of the FWER.*

Furthermore, the following proposition provides exact results for ITP-based F-test.

Proposition 3.1. *The test of the family of hypotheses* (14) *based on the statistic* $T_F^{(k)}$ (15) *is provided with an exact interval-wise control of the FWER. Formally, the ITP-adjusted p-values* $\lambda_{ITP,F}^{(k)}$, $\forall k = 1, ..., p$, *are s.t., for any interval* $\mathcal{I}$ *and any* $\alpha \in (0, 1]$*:*

$$\mathbb{P}_{H_{0,F}^{\mathcal{I}}}\left[\exists k \in \mathcal{I} \text{ s.t. } \lambda_{ITP,F}^{(k)} \leq \alpha\right] \leq \alpha.$$

Next, we focus on the property of consistency of the proposed tests. Let $\mathcal{A}$ denote the set of indexes associated to all components where $H_{0,C}^{(k)}$ is false, i.e., let $H_{1,C}^{\mathcal{A}} = \bigcap_{k\in\mathcal{A}} H_{1,C}^{(k)}$ hold. Then, the following theorem states that the probability of detecting every component of the set $\mathcal{A}$ converges to 1 as the sample size increases.

Theorem 3.2. *The test of the family of linear hypotheses* (13) *based on the test statistic* $T_C^{(k)}$ (18) *is consistent. Formally, for any set* $\mathcal{A} \subseteq \{1, ..., p\}$*, the ITP-adjusted p-values* $\lambda_{ITP,C}^{(k)}$ *are s.t.:*

$$\lim_{n\to\infty} \mathbb{P}_{H_{1,C}^{\mathcal{A}}}\left[\forall k \in \mathcal{A}, \lambda_{ITP,C}^{(k)} \leq \alpha\right] = 1.$$

As a consequence, we obtain also consistency results for ITP-based F-test and t-tests.

Corollary 3.2. *The test of the family of hypotheses* (14) *based on the* F*-test statistic* $T_F^{(k)}$ (15) *is consistent.*

Corollary 3.3. *The test of the family of hypotheses* (16) *for the* l*-th functional regression parameter based on the* t*-test statistic* $T_{t,l}^{(k)}$ (17) *is consistent.*

The ITP used in this paper provides the control of FWER on every interval. However, it is possible to consider for the multiplicity correction not only all possible intervals, but also the complementary sets of all intervals. In such way, the control can be extended over the interval complements as well. For details, we refer to Pini and Vantini (2013).

4 Analysis of knee kinematics

Anterior cruciate ligament injuries are common worldwide, and are typically treated either conservatively with physiotherapy ($\mathrm{ACL_{PT}}$) or with surgery together with physiotherapy ($\mathrm{ACL_R}$). We consider knee-joint kinematics data of knee flexion/extension on the sagital plane during a one-leg hop for distance in a total of $n = 95$ individuals. In detail, we compare individuals from the surgery and physiotherapy groups ($\mathrm{ACL_R}$ and $\mathrm{ACL_{PT}}$, respectively) with age and gender

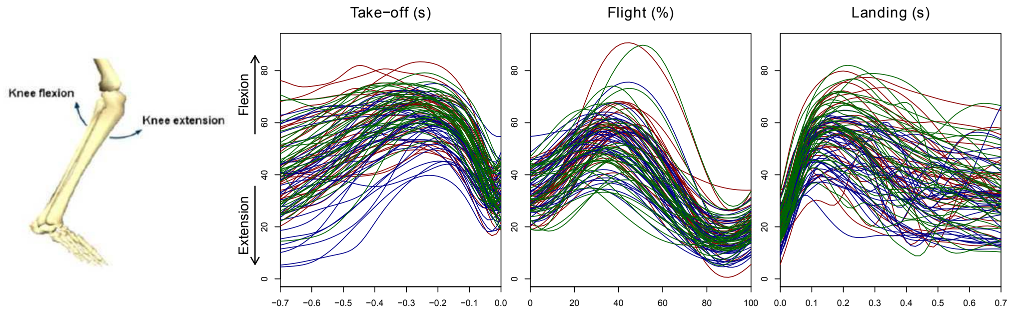

Figure 1: Definition of the knee angles in the sagittal plane (left) and flexion/extension curves of the physiotherapy (blue), surgery (red) and control (green) groups (right).

matched knee-healthy controls (CTRL). We analyse the functional data corresponding to the movement in the sagittal plane, i.e., knee flexion/extension (see, Figure 1). Traditional analysis of kinematic data typically reports results for landmarks of the curves, such as maximal knee angle in the sagittal plane during take-off. Previous results indicate less knee flexion among the individuals treated with physiotherapy (Tengman et al. 2014). A functional ANOVA comparing the three groups reported in Abramowicz et al. (2014) indicates, as well, that the ACL_{PT} group has different knee-joint kinematics during specific parts of the jump, detected especially in the flexion/extension angle during take-off and landing.

In this paper, we investigate if the identified difference from ANOVA analysis is only due to the treatment method, or if it can be explained by means individual-specific covariates, summarized in Table 1.

Variable	Surgery (31 subjects)	Physiotherapy (33 subjects)	Control (31 subjects)
Jump length (m)	1.13 (0.27)	1.00 (0.22)	1.08 (0.23)
BMI (kg/m^2)	27 (3)	28 (4)	25 (3)
Gender ($male/female$)	20/11	21/12	20/11
Age ($years$)	46 (5)	48 (6)	47 (5)

Table 1: Means and standard deviations for all variables considered as covariates in the presented analysis. For gender, frequencies are reported.

The analysis is performed individually on three different phases of the one-leg hop: take-off, flight, and landing. Take-off phase is a 0.7 seconds long time interval preceding the take-off instant, and landing phase is the 0.7 seconds long time interval succeeding the landing instant. Flight phase is the time interval between the take-off and landing instants. The time length of the flight phase differ between individuals, therefore, this phase is standardized to have the same take-off and landing instant for all individuals. For each phase, data are represented by piece-wise linear B-splines with 50 equally distributed knots. Note that in the case of a B-splines basis, each basis component has a compact support. Hence, by looking at the support of the basis components presenting a rejection of the null hypothesis, it is possible to select intervals of the domain presenting the rejection.

The model used to describe the functional data of knee kinematics includes the four covariates displayed in Table 1 and two indicator variables describing the group membership. In detail the model that we apply to describe functional data $y_i(t)$, $i = 1, \ldots, 95$ is the following:

$$\begin{aligned} y_i(t) = \beta_0(t) + \beta_J(t)x_{J,i} + \beta_{BMI}(t)x_{BMI,i} + \beta_A(t)x_{A,i} \\ + \beta_G(t)x_{G,i} + \beta_{CTRL}(t)x_{CTRL,i} + \beta_R(t)x_{R,i}, \end{aligned}$$

where covariates $x_{J,i}$, $x_{BMI,i}$, $x_{A,i}$, $x_{G,i}$ indicate the jump length, BMI, age and gender of each individual i, and $x_{CTRL,i}$, $x_{R,i}$ are the two indicators of CTRL and ACL_{R} groups, respectively. For each phase of the jump, we start with a full model. Then, we reduce the model by removing the covariates that are non-significant on all the domain. However, the group indicators are never excluded.

For all phases the final models include only the group indicators and jump length, and the results are presented in Figure 2. The first row presents the functional responses, together with the F-test results (14). The grey areas indicate the intervals where we detect significant effects of at least one covariate at 5% level. The second row present the results of t-tests of the effect of jump length $H_{0,J}^{(k)} : \beta_J^{(k)} = 0$, and rows three to six present the results of group comparisons, i.e., respectively:

$$H_{0,CTRL-PT}^{(k)} : \beta_{CTRL}^{(k)} = 0, \; H_{0,R-PT}^{(k)} : \beta_R^{(k)} = 0, \; H_{0,CTRL-R}^{(k)} : \beta_{CTRL}^{(k)} - \beta_R^{(k)} = 0.$$

The curves correspond to OLS estimates, and grey areas indicate the presence of a significant effect at 5% significance level. The three columns correspond to take-off, flight, and landing phases, respectively.

The F-test indicates the presence of at least one significant effect in the majority of all three phases. The jump length has a significant effect throughout all three phases, in a large part of the domains. The associated functional regression parameter is positive during take-off and landing while during flight it switches the sign. The sign is negative in the two intervals corresponding to the minimum of flexion and positive sign in an interval corresponding to the maximum of flexion. The ACL_{PT} group is significantly different from the other two groups during take-off and significantly different with respect to CTRL during landing, whereas the three groups do not differ significantly during flight. The regression parameters associated to the differences indicate a lower flexion in the ACL_{PT} group during these two phases with respect to individuals in ACL_{R} and CTRL. Even after having discounted for the jump length, ACL_{PT} group remains significantly different with respect to the other two groups, which is in line with the findings presented in Abramowicz et al. (2014).

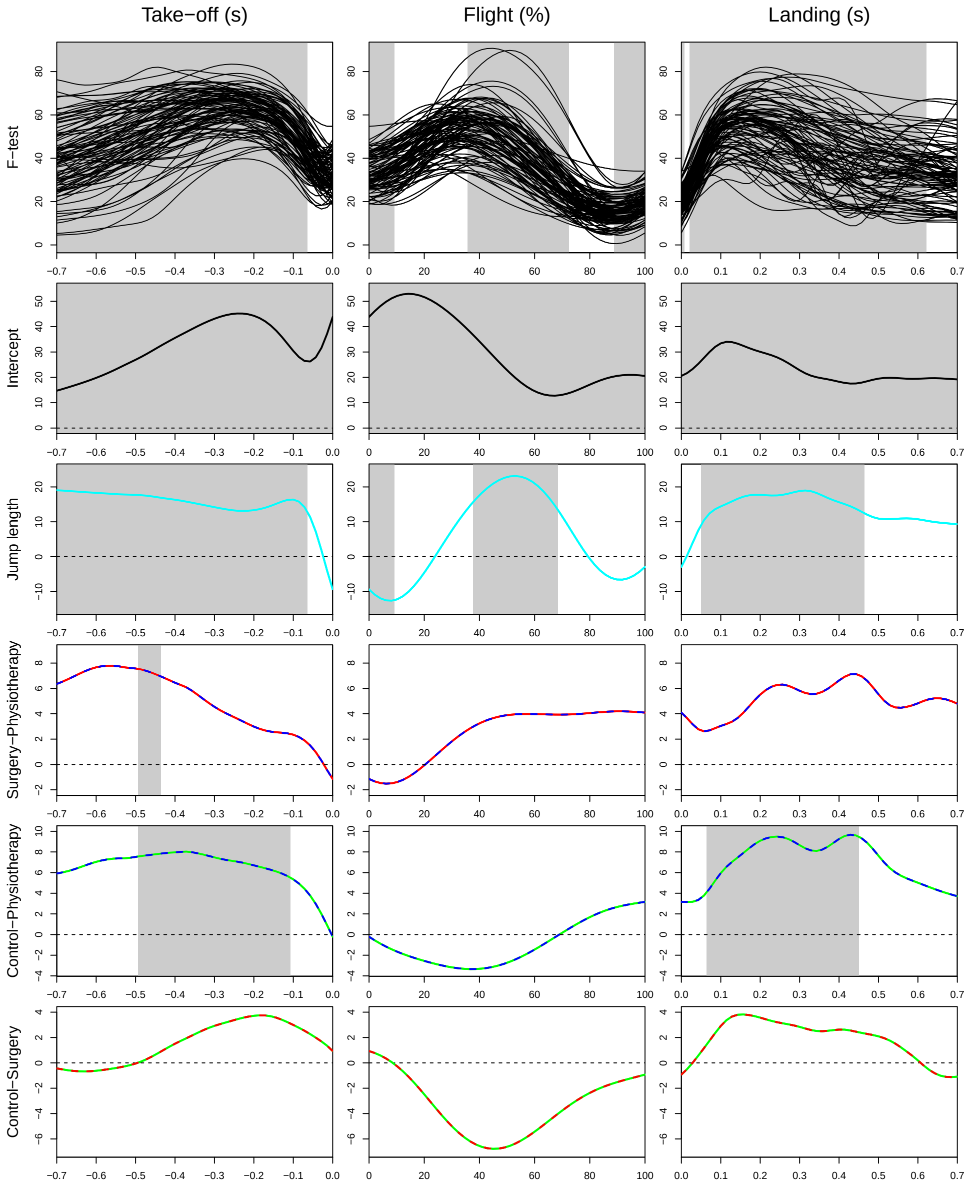

Figure 2: Results of the tests on the functional-on-scalar linear model on knee flexion angle for take-off (left), flight (middle), and landing (right). First row: functional responses and significant F-test intervals at 5% level (gray areas). Rows two to six: OLS estimates and results of functional t-tests and group comparisons (grey areas indicate the presence of a significant effect at 5% level). The three columns correspond to take-off, flight, and landing phases, respectively.

5 Discussion

In this work, we introduced a methodology to estimate and test a functional-on-scalar linear model, i.e., a linear model where the response variable is a function and the covariates are fixed scalar variables multiplied by fixed functional parameters. This type of model can be applied whenever functional data are described through a suitable basis expansion. We showed how the initial functional linear model can be decomposed in a family of dependent linear models, one for each component of the basis expansion.

We provided OLS estimates for the functional regression parameters, as well as tests on the model. Specifically, we provided: (i) a functional F-test for testing the regression model; and (ii) functional t-tests for testing the effects of single covariates. All tests are based on the Interval Testing Procedure (ITP), a non-parametric procedure for testing functional data. We provided theoretical properties for the ITP-based F-test on the regression model and the ITP-based t-tests on the functional regression parameters. In detail, we proved theoretically that the F-test on the regression model is provided with an interval-wise control of the Family Wise Error Rate, implying that the probability of falsely rejecting any interval of true null hypotheses pertaining basis components is controlled. Furthermore, we proved that the F-test on the regression model is consistent, in the sense that the probability of rejecting all false null hypothesis converges to one as the sample size increases. We proved that the t-test on functional regression parameters are provided with an asymptotic interval-wise control of the Family Wise Error Rate, and that they are consistent.

Data from a follow-up study after rehabilitation following anterior cruciate ligament injury are analyzed, applying the functional-on-scalar linear model previously described. Knee kinematics of individuals treated with physiotherapy or surgery and healthy controls were compared during a one-leg hop. The comparison between the three groups was carried out by taking into account individual-specific covariates, such as the jump length, BMI, age and gender. The analysis of these data showed that the effect of jump length on knee kinematics is significantly different from zero, while the effects of BMI and age are not. In line with previous findings, even after having discounted for the jump length, physiotherapy group remains significantly different with respect to the other two groups.

A Proofs

In this section, we prove the theoretical properties reported in Section 3. We first report the theoretical properties of single-components tests based on the Freedman and Lane scheme, i.e., the tests on each component of the basis expansion. Then, we report the theoretical properties of the corresponding multiple-component tests, i.e., the tests on intervals of basis components obtained by means of the NPC of single-components tests. Finally, we prove that the ITP-based tests of linear hypotheses on the functional-on-scalar linear model is provided with an asymptotic interval-wise control of the FWER and that they are consistent. Additionally we show that the ITP-based F-test on the regression model is provided with an exact interval-wise control of the FWER.

A.1 Single-component tests

As mentioned above, we first prove the theoretical properties of single-components tests, i.e., the tests on each component of the basis expansion. We start by showing asymptotic exactness of single-component tests on linear hypotheses.

Lemma A.1. *Under assumptions (C1-C2), and for each component $k = 1, ..., p$, the single-component test of linear hypotheses on the regression parameters* (13) *is asymptotically exact.*

Proof. Let $H_{0,C}^{(k)}$ hold, i.e., $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$. Under the null hypothesis, the model can be reduced by solving the linear system $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$. In detail, since C has full rank, $q \leq L+1$ regression parameters can be removed from the model. Let $\mathcal{Q}$ denote the set of indexes removed. The reduced model is then $y_i^{(k)} = \sum_{r \notin \mathcal{Q}} {\beta_r}^{(k)} a_r^{(k)} x_{ri} + {\varepsilon_i}^{(k)}$, where $x_{0i} = 1$, $a_r^{(k)}$ are fixed known coefficients (depending only on the solution of linear system $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$), and ${\varepsilon_i}^{(k)}$ are *i.i.d.* and zero-mean errors.

The Freedman and Lane permutation scheme is based on the permutations of the residuals $\hat{\varepsilon}_{i,C}^{(k)} = y_i^{(k)} - \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}^{(k)} a_r^{(k)} x_{ri}$, where $\hat{\beta}_{r,C}^{(k)}$, $r \notin \mathcal{Q}$ are the OLS estimate of parameters $\beta_r^{(k)}$ under the reduced model. Under conditions (*C1-C2*), we have strong consistency of the OLS parameter estimates, i.e., in our case: $\hat{\beta}_{r,C}^{(k)} \xrightarrow{a.s.} \beta_r^{(k)}$, $\forall r \notin \mathcal{Q}$. Hence, we also have the strong convergence of the residuals, i.e., $\hat{\varepsilon}_{i,C}^{(k)} \xrightarrow{a.s.} \varepsilon_i^{(k)}$, $\forall i = 1, ..., n$.

The errors ${\varepsilon_i}^{(k)}$ of the reduced linear model are exchangeable. Hence, the likelihood of every permutation is invariant, and equal to $1/n!$. Therefore, the test based on the permutations of the errors ${\varepsilon_i}^{(k)}$ is exact. As $\hat{\varepsilon}_{i,C}^{(k)} \xrightarrow{a.s.} \varepsilon_i^{(k)}$, the residuals are asymptotically exchangeable, i.e., the likelihood of every permutation is asymptotically invariant, and converges to $1/n!$. Hence, the test based on permutations of the residuals is asymptotically exact. □

Asymptotical exactness for the t-test (16) is a direct consequence of the above lemma. As an addition to asymptotic results for single component tests on any linear hypothesis, we prove exactness of single-component F-test.

Lemma A.2. *For each component $k = 1, \dots, p$, the single-component F-test of (14) is exact.*

Proof. Under $H_{0,F}^{(k)}$ we have ${y_i}^{(k)} = {\beta_0}^{(k)} + {\varepsilon_i}^{(k)}$. The estimated residuals of this model are $\hat{\varepsilon}_{i,0}^{(k)} = {\beta_0}^{(k)} + {\varepsilon_i}^{(k)} - \hat{\beta}_0^{(k)}$, where $\hat{\beta}_0^{(k)} = \bar{y}^{(k)}$ is the sample mean of the responses ${y_i}^{(k)}$. Note that the quantity ${\beta_0}^{(k)} + {\varepsilon_i}^{(k)} - \hat{\beta}_0^{(k)}$ is permutationally invariant. Hence, the independence between the errors implies the exchangeability of the residuals under $H_{0,F}^{(k)}$. Thus, the test is exact, as it is based on the permutation of exchangeable quantities (Pesarin and Salmaso 2010). □

In the next step, we verify the consistency of single-component tests on linear hypotheses.

Lemma A.3. *For each component $k = 1, \dots, p$, the single-component test of linear hypotheses on the regression parameters* (13) *based on the test statistic* $T_C^{(k)}$ (18) *is consistent.*

Proof. The statement follows directly from the fact that the test statistic $T_C^{(k)}$ is stochastically greater under $H_{1,F}^{(k)}$ than under $H_{0,F}^{(k)}$ (Pesarin and Salmaso 2010). □

As direct implication of Lemma A.3, we get the consistency of single-component F-test and t-tests based on test statistics $T_F^{(k)}$ and $T_{t,l}^{(k)}$, respectively.

A.2 Multiple-components tests

Next, we investigate the properties of multiple-component tests $H_{0,C}^{\mathcal{I}} = \bigcap_{k \in \mathcal{I}} H_{0,C}^{(k)}$, where $\mathcal{I} = \{k_1, ..., k_2\}$ and $1 \leq k_1 < k_2 \leq p$. To construct these tests from the results of joint single-component tests, we use the NPC methodology. We start by proving the asymptotic exactness of such tests on linear hypotheses.

Lemma A.4. *Under assumptions (C1-C2), for each interval of components $\mathcal{I}$, the multiple-component test of linear hypotheses on the regression parameters $H_{0,C}^{\mathcal{I}}$ is asymptotically exact.*

Proof. Let $H_{0,C}^{\mathcal{I}}$ hold, i.e., $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$, for any $k \in \mathcal{I}$. Under the null hypothesis, and for each $k \in \mathcal{I}$, the model can be reduced by solving the linear system $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$. In detail, since C has full rank, $q \leq L + 1$ regression parameters can be removed from the model. Let $\mathcal{Q}$ denote the set of indexes removed. The reduced model is then $y_i^{(k)} = \sum_{r \notin \mathcal{Q}} {\beta_r}^{(k)} a_r^{(k)} x_{ri} + {\varepsilon_i}^{(k)}$, where $x_{0i} = 1$, $a_r^{(k)}$ are fixed known coefficients (depending only on the solution of linear systems $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$), and ${\varepsilon_i}^{(k)}$ are *i.i.d.* and zero-mean errors.

The NPC applied to the Freedman and Lane permutation scheme is based on the joint permutations (the same for each k) of the residuals $\hat{\varepsilon}_{i,C}^{(k)} = y_i^{(k)} - \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}^{(k)} a_r^{(k)} x_{ri}$, where $\hat{\beta}_{r,C}^{(k)}$, $r \notin \mathcal{Q}$ are the OLS estimate of parameters $\beta_r^{(k)}$ under the reduced model. Under conditions (*C1-C2*), we have strong consistency of the OLS parameters estimates, i.e., in our case: $\hat{\beta}_{r,C}^{(k)} \xrightarrow{a.s.} \beta_r^{(k)}$, $\forall r \notin \mathcal{Q}$, and $\forall k \in \mathcal{I}$. Hence, we also have the strong convergence of the residuals, i.e., $\hat{\varepsilon}_{i,C}^{(k)} \xrightarrow{a.s.} \varepsilon_i^{(k)}$, $\forall i = 1, ..., n$ and $\forall k \in \mathcal{I}$.

The errors ${\varepsilon_i}^{(k)}$ of the linear model are jointly exchangeable. Hence, the likelihood of every joint permutation is invariant, and equal to $1/n!$. So, the test based on the joint permutations of the errors ${\varepsilon_i}^{(k)}$ is exact. As $\hat{\varepsilon}_{i,C}^{(k)} \xrightarrow{a.s.} \varepsilon_i^{(k)}$, $\forall k \in \mathcal{I}$, the residuals are jointly asymptotically exchangeable, i.e., the likelihood of every joint permutation is asymptotically invariant, and converges to $1/n!$. Hence, the test based on joint permutations of the residuals is asymptotically exact. □

Asymptotic exactness of multiple-component t-tests is a direct implication of the above lemma. As in single-component case, we can also show stronger result for multiple-component F-test.

Lemma A.5. *For each interval of components $\mathcal{I}$, the multiple-component F-test of the regression model is exact.*

Proof. Since all univariate tests are exact and consistent (Corollary A.2 and Lemma A.3), the combined test is also exact, due to results of Pesarin and Salmaso (2010). □

We proceed by proving consistency of multiple-component tests on linear hypotheses.

Lemma A.6. *For each interval of components $\mathcal{I}$, the multiple-component test of linear hypotheses on the regression parameters $H_{0,C}^{\mathcal{I}}$ is consistent.*

Proof. The consistency of the multiple-component test follows directly from the consistency of the corresponding single-component test (Lemma A.3) and results of Pesarin and Salmaso (2010). □

Once again, since F-test and t-tests are special cases of linear hypothesis tests, the consistency of the multiple-component F-test and t-tests follows immediately from Lemma A.6.

A.3 Properties of IPT-based tests

We start by proving Theorem 3.1, establishing asymptotic interval-wise control of ITP-based tests of linear hypotheses.

Proof of Theorem 3.1. Let $\mathcal{I}$ be an interval of components associated to only true null hypotheses. Consider a component k of the interval, $k \in \mathcal{I}$, and let $\mathcal{K}$ denote the set of all intervals containing the component k. The ITP-adjusted p-value associated to component k is $\lambda_{ITP,C}^{(k)} = \max_{\mathcal{J}\in\mathcal{K}} \lambda_C^{\mathcal{J}}$, where $\lambda_C^{\mathcal{J}}$ is the p-value of the permutation test on the interval $\mathcal{J}$. In particular, as $\mathcal{I} \in \mathcal{K}$, we have that $\lambda_{ITP,C}^{(k)} \geq \lambda_C^{\mathcal{I}}$, and $\mathbb{P}_{H_{0,C}^{\mathcal{I}}}[\lambda_{ITP,C}^{(k)} \leq \alpha] \leq \mathbb{P}_{H_{0,C}^{\mathcal{I}}}[\lambda_C^{\mathcal{I}} \leq \alpha]$. Since all tests are asymptotically exact (Lemmas A.1 and A.4), we have:

$$\lim_{n\to\infty} \mathbb{P}_{H_{0,C}^{\mathcal{I}}}\left[\lambda_C^{\mathcal{I}} \leq \alpha\right] = \alpha,$$

and therefore,

$$\limsup_{n\to\infty} \mathbb{P}_{H_{0,C}^{\mathcal{I}}}\left[\exists k \in \mathcal{I} : \lambda_{ITP,C}^{(k)} \leq \alpha\right] \leq \alpha.$$

□

Assertion of Proposition 3.1 follows directly from the results of Pini and Vantini (2013) and the fact that univariate and multivariate tests used to build the procedure are exact (Lemmas A.2 and A.5). We now prove Theorem 3.2, which guarantees consistency of ITP-based tests of linear hypothesis.

Proof of Theorem 3.2. Suppose that for a basis component k the alternative hypothesis $H_{1,C}^{(k)}$ is true, i.e. $C\boldsymbol{\beta}^{(k)} \neq \mathbf{c}_0^{(k)}$. Let $\mathcal{K}$ denote the set of every interval containing the component k. If $H_{1,C}^{(k)}$ is true, also each alternative hypothesis pertaining intervals in $\mathcal{K}$ is true. Since each test is consistent (Lemmas A.3, A.6), it follows that, for $n \to \infty$, for each $\mathcal{I} \in \mathcal{K}$, the p-value $\lambda_C^{\mathcal{I}}$ converges to zero almost surely. The ITP-adjusted p-value $\lambda_{ITP,C}^{(k)}$ is the maximum among all p-values of the tests containing k, i.e., $\lambda_{ITP,C}^{(k)} = \max_{\mathcal{I}\in\mathcal{K}} \lambda_C^{\mathcal{I}} \to 0$, almost surely. Then, for the ITP-adjusted p-value associated to the k-th component $\lambda_{ITP,C}^{(k)}$, we have:

$$\lim_{n\to\infty} \mathbb{P}_{H_{1,C}^{(k)}} \left[\lambda_{ITP,C}^{(k)} \leq \alpha\right] = 1.$$

The latter holds for any $k \in \mathcal{A}$, where $\mathcal{A}$ denotes the set of all false null hypotheses. Hence, we also have:

$$\lim_{n\to\infty} \mathbb{P}_{H_{1,C}^{\mathcal{A}}} \left[\forall k \in \mathcal{A}, \lambda_{ITP,C}^{(k)} \leq \alpha\right] = 1.$$

□

The Corollary 3.2 and Corollary 3.3 follow immediately from Theorem 3.2, as special cases.

B The Freedman and Lane permutation scheme

In this section, we give some details of the implementation of the Freedman and Lane permutation scheme for testing linear hypotheses on the regression model for each fixed component k (see eq. (4))

$$y_i^{(k)} = \sum_{l=0}^{L} \beta_l{}^{(k)} x_{li} + \varepsilon_i{}^{(k)}, \quad \forall i = 1, \ldots, n,$$

with $x_{i0} = 1$, $\forall i$. Further, we present the two specific cases: F-test on the regression model; and t-tests on the regression parameters.

The Freedman and Lane permutations are based on the following steps:

i the residuals of the reduced model (that is the linear model under the null hypothesis) are estimated;

ii the residuals of the reduced model are permuted;

iii the permuted responses are computed, through the reduced model and permuted residuals.

For more details about this method, we refer to Freedman and Lane (1983); Anderson and Legendre (1999).

B.1 Tests on linear hypotheses

Under the null hypothesis (13), the model (4) can be reduced by solving the linear system $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$. In detail, since C has full rank, q regression parameters can be removed from the model, by expressing them in terms of the others. Let $\mathcal{Q}$ denote the set of indexes of the removed regression parameters. The reduced model is then:

$$y_i^{(k)} = \sum_{r \notin \mathcal{Q}} \beta_r{}^{(k)} \tilde{x}_{ri} + \varepsilon_i{}^{(k)}, \tag{19}$$

i.e., the kth basis coefficients of the responses can be written in terms of a linear combination of modified covariates $\tilde{x}_{ri} = a_r^{(k)} x_{ri}$, where $a_r^{(k)}$ are fixed known coefficients, depending only on the solution of linear system $C\boldsymbol{\beta}^{(k)} = \mathbf{c}_0^{(k)}$, and $\varepsilon_i{}^{(k)}$ are *i.i.d.* and zero-mean errors.

The residuals of the reduced model can then be estimated as $\hat{\varepsilon}_{i,C}^{(k)} = y_i^{(k)} - \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}^{(k)} \tilde{x}_{ri}$, where $\hat{\beta}_{r,C}^{(k)}$ are the OLS estimates of parameters $\beta_r^{(k)}$, $r \notin \mathcal{Q}$, of model (19). Then, the residuals $\hat{\varepsilon}_{i,C}^{(k)}$ are permuted, and the permuted responses are evaluated using the permuted residuals $\hat{\varepsilon}_{i,C}^{(k)^*}$ in the reduced model (19):

$$y_i^{(k)^*} = \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}^{(k)} \tilde{x}_{ri} + \hat{\varepsilon}_{i,C}^{(k)^*}. \tag{20}$$

B.2 F-test for the regression model

In the case of the F-test (14), under the null hypothesis all regression parameters except the intercept are null. So, the reduced model is:

$$y_i{}^{(k)} = \beta_0{}^{(k)} + \varepsilon_i{}^{(k)}.$$

The estimated residuals of the reduced model are $\hat{\varepsilon}_{i,F}^{(k)} = y_i{}^{(k)} - \bar{y}^{(k)}$, where $\bar{y}^{(k)}$ is the sample mean of the responses $y_i{}^{(k)}$. Therefore, using the permuted residuals $\hat{\varepsilon}_{i,F}^{(k)^*}$, we get:

$$y_i{}^{(k)^*} = \bar{y}^{(k)} + \hat{\varepsilon}_{i,F}^{(k)^*}.$$

Note that in this case permuting the residuals $\hat{\varepsilon}_{i,F}^{(k)}$ is equivalent to permuting the responses $y_i{}^{(k)}$.

B.3 t-tests on regression parameters

In the case of t-tests, the model under null hypothesis (16) reduces to:

$$y_i^{(k)} = \beta_0{}^{(k)} + \sum_{r \neq l} \beta_r{}^{(k)} x_{ri} + \varepsilon_i{}^{(k)}.$$

The estimated residuals of such model are $\hat{\varepsilon}_{i,l}^{(k)} = y_i^{(k)} - \hat{\beta}_{0,l}^{(k)} + \sum_{r \neq l} \hat{\beta}_{r,l}^{(k)} x_{li}$, where $\hat{\beta}_{r,l}^{(k)}$ are the OLS estimates of the parameters of the reduced model. Then, the permuted responses are:

$$y_i^{(k)^*} = \hat{\beta}_{0,l}^{(k)} + \sum_{r \neq l} \hat{\beta}_{r,l}^{(k)} + \hat{\varepsilon}_{i,l}^{(k)^*}, \tag{21}$$

where $\hat{\varepsilon}_{i,l}^{(k)^*}$ are the permuted residuals.

C The Non Parametric Combination procedure

The NPC methodology (Pesarin and Salmaso 2010) allows to build multivariate permutation tests starting from the results of a family of joint univariate permutation tests.

Consider a family of null hypotheses $\{H_0^{(m)}\}_{m \in \mathcal{M}}$, with $\mathcal{M} = \{m_1, \ldots, m_d\}$. Each hypothesis $H_0^{(m)}$ is tested against $H_1^{(m)}$, $m \in \mathcal{M}$, by means of a suitable permutation test, with test statistic $T^{(m)}$. Let $\lambda^{(m)}$ denote the resulting p-value. We want to test the multivariate hypothesis $H_0^{(\mathcal{M})} = \bigcap_{m \in \mathcal{M}} H_0^{(m)}$ against the alternative $H_1^{(\mathcal{M})} = \bigcup_{m \in \mathcal{M}} H_1^{(m)}$, using the results of the univariate tests of $H_0^{(m)}$ versus $H_1^{(m)}$, with $m \in \mathcal{M}$. The test statistic for such test is the combination of univariate p-values: $\psi(\lambda^{(m_1)}, ..., \lambda^{(m_d)})$, where ψ is any valid combining function, i.e., any function $\psi : [0,1]^d \mapsto \mathbb{R}$ satisfying:

(P1) ψ is non-increasing in each argument;

(P2) ψ is invariant with respect to rearrangements of its arguments:

$$\psi(\lambda^{(m_1)}, ..., \lambda^{(m_d)}) = \psi(\lambda^{(m_1^*)}, ..., \lambda^{(m_d^*)}),$$

where $(\lambda^{(m_1^*)}, ..., \lambda^{(m_d^*)})$ is any rearrangement of $(\lambda^{(m_1)}, ..., \lambda^{(m_d)})$;

(P3) ψ attains its supremum value $\bar{\psi}$ (possibly not finite) even when only one argument attains zero;

(P4) for $\alpha \in (0,1]$, let ψ_α denote the critical value of the test statistic, i.e., $\psi_\alpha = F_\psi^{-1}(\alpha)$, where F_ψ is the cdf of the test statistic ψ. Then, for a valid combining function, ψ_α is finite and strictly smaller than $\bar{\psi}$.

The following theorem, reported in Pesarin and Salmaso (2010) shows the properties of combined tests:

Theorem C.1. *If permutation tests for respectively $H_0^{(m)}$ against $H_1^{(m)}$, $m \in \mathcal{M}$ are exact and consistent, then the NPC test based on a combining function ψ satisfying (P1) to (P4) is an exact and consistent test for $H_0^{\mathcal{M}}$ against $H_1^{\mathcal{M}}$.*

References

Abramovich, F. and Angelini, C. (2006), Testing in mixed-effects FANOVA models, *J. Statist. Plann. Inference* **136**(12), 4326–4348.

Abramowicz, K., Häger, C., Hébert-Losier, K., Pini, A., Schelin, L., Strandberg, J. and Vantini, S. (2014), An inferential framework for domain selection in functional ANOVA, *in* E. G. Bongiorno, E. Salinelli, A. Goia and P. Vieu, eds, 'Contributions in infinite-dimensional statistics and related topics', Società Editrice Esculapio, pp. 13–18.

Anderson, M. J. and Legendre, P. (1999), An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model, *J. Stat. Comput. Sim.* **62**(3), 271–303.

Anderson, M. J. and Robinson, J. (2001), Permutation tests for linear models, *Aust. NZ J. Stat.* **43**(1), 75–88.

Antoniadis, A. and Sapatinas, T. (2007), Estimation and inference in functional mixed-effects models, *Comput. Statist. Data Anal.* **51**(10), 4793–4813.

Cardot, H., Prchal, L. and Sarda, P. (2007), No effect and lack-of-fit permutation tests for functional regression, *Comput. Statist.* **22**(3), 371–390.

Cuesta-Albertos, J. A. and Febrero-Bande, M. (2010), A simple multiway ANOVA for functional data, *TEST* **19**(3), 537–557.

Cuevas, A., Febrero, M. and Fraiman, R. (2004), An ANOVA test for functional data, *Comput. Statist. Data Anal.* **47**(1), 111–122.

Davison, A. C. and Hinkley, D. V. (1997), *Bootstrap methods and their application*, Cambridge university press.

Fan, J. and Zhang, J. T. (2000), Two-step estimation of functional linear models with applications to longitudinal data, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62**(2), 303–322.

Freedman, D. and Lane, D. (1983), A nonstochastic interpretation of reported significance levels, *J. Bus. Econ. Stat.* **1**(4), 292–298.

Gertheiss, J., Goldsmith, J., Crainiceanu, C. and Greven, S. (2013), Longitudinal scalar-on-functions regression with application to tractography data, *Biostatistics* **14**(3), 447–461.

Holm, S. (1979), A simple sequentially rejective multiple test procedure, *Scand. J. Stat.* **6**, 65–70.

Johnson, R. A. and Wichern, D. W. (2007), *Applied multivariate statistical analysis*, 6th edn, Pearson Education Inc.

Lai, T. L., Robbins, H. and Wei, C. Z. (1979), Strong consistency of least squares estimates in multiple regression, *J. Multivariate Anal.* **9**(3), 343–361.

Pesarin, F. and Salmaso, L. (2010), *Permutation tests for complex data: theory, applications and software*, John Wiley & Sons Inc.

Pini, A. and Vantini, S. (2013), The interval testing procedure: Inference for functional data controlling the family wise error rate on intervals., Technical Report 13/2013, MOX, Politecnico di Milano.

R Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*

Ramsay, J. O. and Silverman, B. W. (2005), *Functional data analysis*, Springer, New York.

Reiss, P. T., Huang, L. and Mennes, M. (2010), Fast function-on-scalar regression with penalized basis expansions, *Int. J. Biostat.* **6**(1).

Schott, J. R. (2007), Some high-dimensional tests for a one-way MANOVA, *J. Multivariate Anal.* **98**(9), 1825–1839.

Spitzner, D. J., Marron, J. S. and Essick, G. K. (2003), Mixed-model functional ANOVA for studying human tactile perception, *J. Amer. Statist. Assoc.* **98**(462), 263–272.

Tengman, E., Brax Olofsson, L., Nilsson, K. G., Tegner, Y., Lundgren, L. and Häger, C. (2014), Anterior cruciate ligament injury after more than 20 years: I. physical activity level and knee function, *Scand. J. Med. Sci. Sport.* .

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. and Nichols, T. E. (2014), Permutation inference for the general linear model, *NeuroImage* **92**, 381–397.

Zeng, C., Pan, Z., MaWhinney, S., Barón, A. E. and Zerbe, G. O. (2011), Permutation and F distribution of tests in the multivariate general linear model, *Am. Stat.* **65**(1), 31–36.

Zhang, J. and Liang, X. (2014), One-way ANOVA for functional data via globalizing the pointwise F-test, *Scand. J. Stat.* **41**(1), 51–71.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

65/2015 Abramowicz, K.; de Luna, S.; Hger, C.; Pini, A.; Schelin, L.; Vantini, S.
Distribution-Free Interval-Wise Inference for Functional-on-Scalar Linear Models

64/2015 Menafoglio, A.; Petris, G.
Kriging for Hilbert-space valued random fields: the Operatorial point of view

63/2015 Pini, A.; Stamm, A.; Vantini, S.
Hotelling s T^2 in functional Hilbert spaces

62/2014 Andr, C.; Parolini, N.; Verani, M.
Using gambling simulators to foster awareness about gambling risks

61/2014 Taffetani, M.; Ciarletta, P.
Elasto-capillarity controls the formation and the morphology of beads-on-string structures in solid fibres

59/2014 Manzoni, A.; Pagani, S.; Lassila, T.
Accurate solution of Bayesian inverse uncertainty quantification problems using model and error reduction methods

58/2014 Menafoglio, A.; Secchi, P.; Guadagnini, A.
A Class-Kriging predictor for Functional Compositions with Application to Particle-Size Curves in Heterogeneous Aquifers

57/2014 Giverso, C.; Verani, M.; Ciarletta P.;
Branching instability in expanding bacterial colonies

60/2014 Signorini, M.; Micheletti, S.; Perotto, S.
CMFWI: Coupled Multiscenario Full Waveform Inversion for seismic inversion

55/2014 Antonietti, P. F.; Houston P.; Sarti, M.; Verani, M.
Multigrid algorithms for hp-version Interior Penalty Discontinuous Galerkin methods on polygonal and polyhedral meshes