Testing Phase and Amplitude Variability in Functional Data Analysis: a Hierarchical Permutation Test Approach

Alessia Pini and Simone Vantini

Key words: Functional Data Analysis, Permutation Test, Phase Variability, Amplitude Variability, NPC test

1 Introduction

In many actual fields of research, data are curves observed in a continuous domain. The area in statistics that deals with such data is Functional Data Analysis. For this type of data, many classical inferential tools becomes nearly useless, as they require the number of sample units to be greater than the dimension of the space in which inference has to be carried out. In this work, we propose a novel non parametric hierarchical procedure based on nested permutation tests that enables inference of functional data, when testing for differences between two populations. Moreover, this new approach allows to investigate, in case of rejection of the null hypothesis, the aspects that lead to the rejection, both in terms of phase and amplitude. In this work we focus on periodic functional data, although the proposed framework is very general, and could be applied to any functional data set represented by means of a suitable basis (see [3] for details).

2 Hierarchical testing procedure

Let $\{\mathbf{y}_1, ..., \mathbf{y}_n\}$ be a collection of $n = n_a + n_b$ curves. Assume that the first n_a curves represent a random sample from a first population $\mathbf{Y}_a$, and the remaining ones a

Alessia Pini and Simone Vantini
MOX, Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, - 20133 Milan, Italy e-mail: alessia.pini@mail.polimi.it; simone.vantini@polimi.it

sample from a second population $\mathbf{Y}_b$. We aim at testing the null hypothesis $\mathbf{Y}_a \overset{d}{=} \mathbf{Y}_b$ against the alternative $\mathbf{Y}_a \overset{d}{\neq} \mathbf{Y}_b$, in both the uncoupled and the coupled scenario.

In this work, we focus on periodic functional data observed on a grid $t_1, ..., t_J$ during a period T. This type of data set can be represented in the frequency domain by means of a truncated Fourier expansion:

$$y_i(t) = \mu_{0i} + \sum_{k=1}^{(J-1)/2} A_{ki} \cos\left(\frac{2\pi}{T} kt + \phi_{ki}\right) \tag{1}$$

Thus, we can represent the ith unit by means of the $(J-1)/2$ coefficients of phase ϕ_{ik} and of the $(J-1)/2$ coefficients of amplitude A_{ik} associated with the first $(J-1)/2$ frequencies.[1] Coherently, for the "0th" frequency, we can define the phase and amplitude coefficients as $A_{0i} = |\mu_{0i}|$ and $\phi_{0i} = \pi[1 - \text{sign}(\mu_{0i})]/2$.

In order to provide tests on differences between the two populations, we propose a hierarchical combination of the $p = J + 1$ univariate permutation tests associated to the parameters of the Fourier expansion (1), with single combinations being based on NPC tests [2]. The test is defined by a family of transformations of the data set which preserves the likelihood under H_0, suitable test statistics for the univariate tests, and a suitable hierarchy of combination functions. The family of transformations of the original data set depends on the type of test we want to perform:

- in the **uncoupled** case the family of transformations is composed by any permutation over the sample units of the observed values. So $\frac{n!}{n_a! n_b!}$ different rearrangements of the data set are obtained.
- in the **coupled** case the family of transformations in composed by within-couple permutations of the observed values. So, $2^{n_a} = 2^{n_b}$ different rearrangements of the data set are obtained.[2]

The univariate test statistic used for each parameter at frequency k depends on the variable we intend to test and on the type of test. In particular:

- for the kth **amplitude** test $H_{amp}^{(k)}$ (i.e., $H_{0_{amp}}^{(k)} : A_{k_a} \overset{d}{=} A_{k_b}$ vs $H_{1_{amp}}^{(k)} : A_{k_a} \overset{d}{\neq} A_{k_b}$), we use in the uncoupled case, the absolute value of the logarithm ratio between geometric sample means: $T_{amp}(\mathbf{A}_k^*) = \left|\log\left(\frac{\left(\prod_{i=1}^{n_a} A_{ki}^*\right)^{1/n_a}}{\left(\prod_{i=n_a+1}^{n} A_{ki}^*\right)^{1/n_b}}\right)\right|$. In the coupled case we use the absolute value of the logarithm of the sample mean of ratio between coupled data: $T_{amp}(\mathbf{A}_k^*) = \left|\log\left(\prod_{i=1}^{n_a} \frac{A_{ki_a}^*}{A_{ki_b}^*}\right)^{1/n_a}\right|$.
- for the kth **phase** test $H_{ph}^{(k)}$ (i.e., $H_{0_{ph}}^{(k)} : \phi_{k_a} \overset{d}{=} \phi_{k_b}$ vs $H_{1_{ph}}^{(k)} : \phi_{k_a} \overset{d}{\neq} \phi_{k_b}$), we use the absolute value of the geodesic distance between the two geodesic sample means of the phases: $T_{ph}(\phi_k^*) = \arccos\langle \mathbf{r}_{k_a}^*, \mathbf{r}_{k_b}^* \rangle$, where $\mathbf{r}_{k_a}^*, \mathbf{r}_{k_b}^*$ denotes the

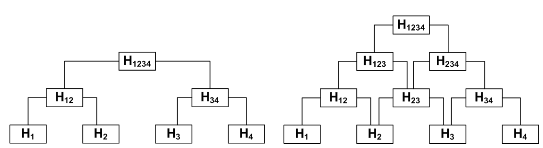
[1] Here, we suppose the number of grid points J to be odd. In the general case, we can only calculate $\hat{J} = J - 1$ coefficients.

[2] In the coupled scenario the number of sample units of the two groups of course coincide, thus we have necessarily $n_a = n_b$.

unit vectors associated to the two geodesic sample means. In the coupled case we use the geodesic sample mean of geodesic distances between coupled data: $T_{ph}(\phi_k^*) = m_{geo}[d_{geo}(\mathbf{q}^*_{ki_a}, \mathbf{q}^*_{ki_b})]$, where $\mathbf{q}^*_{ki_a}, \mathbf{q}^*_{ki_b}$ denotes the unit vectors of direction $\phi^*_{ki_a}, \phi^*_{ki_b}$.

The results of univariate tests $H_{amp}^{(k)}$ and $H_{ph}^{(k)}$ have to be corrected, taking into account multiplicity, in order to control the global error. For this purpose, we could use the closed testing procedure described in [1]. The idea is to test at the same α level each marginal hypothesis and multivariate intersection of them, and to reject any hypothesis $H_{amp}^{(k)}$ (or $H_{ph}^{(k)}$) when the test of every intersection containing $H_{amp}^{(k)}$ ($H_{ph}^{(k)}$) is statistically significant at level α. This procedure provides a strong control of the Family-Wise Error (FWE), although the actual number of tests to be performed is $\mathscr{O}(2^p)$, which grows exponentially with the dimension of the sample space. By consequence, this procedure becomes unfeasible in the framework of functional or high-dimensional data.

Thus, we propose a variation of the described procedure that focuses only on suitable hypothesis intersections. In particular, we chose to combine amplitude and phase results separately, and we focus our attention only on adjacent hypothesis intersections, i.e., only intersections of the type $\{H_{amp}^{(j)} \cap H_{amp}^{(j+1)} \cap \ldots \cap H_{amp}^{(j+h)}\}$, or $\{H_{ph}^{(j)} \cap H_{ph}^{(j+1)} \cap \ldots \cap H_{ph}^{(j+h)}\}$, with $j, j+h \in \{0, \ldots, (J-1)/2\}$, obtained in a hierarchical way. In fact, due to the intrinsic regularity of functional data, we expect false hypotheses not to be spread at isolated points along the frequency domain but aggregated in intervals of frequencies. We construct two types of trees, with the structures shown on Figure 1. In the first case (Figure 1.a) we only consider $2p$ intersections, whereas in the second case (Figure 1.b) the number of intersections is $\frac{p(p-1)}{2}$. The number of tests to be performed in the two cases is thus $\mathscr{O}(p)$ and $\mathscr{O}(p^2)$, respectively, making them useful for dealing with functional data. With the procedure we proposed, we only provide a partial control of the FWE, which is intermediate between weak control (i.e., control of the significance level of the global test) and strong control (i.e., a control of significance level of all multivariate tests constituted by hypothesis subsets). By means of MC simulations - not reported here for brevity - we show that, in scenarios with aggregated false hypotheses, the power of the hierarchical testing procedure is greater than the one obtained through the complete closed testing procedure, supporting the heuristic idea that the hierarchical testing procedure is able, in this framework, to jointly reduce the computational effort while improving the statistical power. A sound theoretical study of this property characterizing the hierarchical testing procedure is currently under investigation.



(a) First combining tree (b) Second combining tree

Fig. 1 Trees of combined adjacent hypotheses with the two proposed strategies of hierarchical combination, in the example $p = 4$

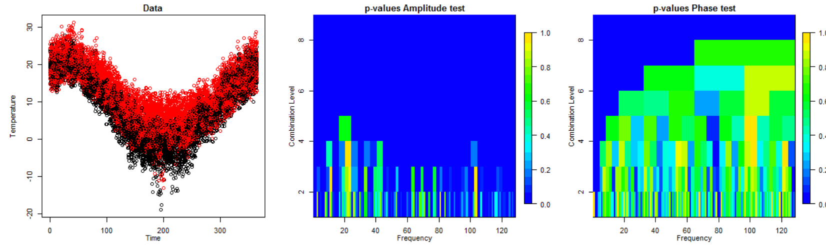


Fig. 2 NASA daily temperatures data, amplitude and phase tests results using the first tree

3 Application to NASA daily temperature data

We present an application of the procedure on a case study. Data are daily temperatures in Milan and Paris registered from July 1983 to June 2005 and stored in the database NASA *Earth Surface Meteorology for Solar Energy*. In the application reported, we identified the 22 years as sample units and the 365 records available for each year as 365 point-wise evaluations of the functional data. We tested for differences between Milan (black dots) and Paris (red dots) temperature profiles in the coupled scenario. Data are displayed on the left panel of Figure 2.

We perform phase and amplitude tests through the hierarchical testing procedure proposed in Section 2, using the Fisher combination function and the first-tree aggregation strategy. In the central and right panels of Figure 2 we represent the result of each test included in the hierarchical procedure (i.e., the p-value dendrogram). In particular, the horizontal axis is associated to frequencies and the vertical one to levels of combination. Colors represent the p-values of each corresponding test. Frequencies to be rejected are those who produce a significant test for all investigated intersections, i.e., frequencies associated to a dark blue vertical line on the graph. In this case, significant phase differences are detected for the sinusoids of period one-year and half-year: these differences are related to a longer Spring and a shorter Fall in Paris with respect to Milan. Significant amplitude differences are instead detected for the constant term and for the sinusoid of period one-year: these differences are related to an average colder weather in Milan, due to colder Winters. Interestingly, at higher frequencies significant differences are found in terms of amplitude but not of phase, showing a much higher short-term variability of the daily temperature in Paris.

References

1. Marcus, R., Eric, P., Gabriel, K.R.: On Closed Testing Procedures with Special Reference to Ordered Analysis of Variance. Biometrika **73** No.3, 655–660 (1976).
2. Pesarin, F., Salmaso, L.: Permutation Tests for Complex Data. John Wiley & Sons (2010).
3. Pini, A., Vantini, S.: Hierarchical Permutation Tests for Functional Data. Technical Report MOX, Dip. di Matematica 'F. Brioschi', Politecnico di Milano (2012).