NONPARAMETRIC LOCAL INFERENCE FOR FUNCTIONAL DATA DEFINED ON MANIFOLD DOMAINS

Niels Lundtorp Olsen[1], Alessia Pini [2] and Simone Vantini[3]

[1] Department of Applied Mathematics and Computer Science, Technical University of Denmark , (e-mail: nalo@dtu.dk)

[2] Department of Statistical Sciences, Università Cattolica del Sacro Cuore, (e-mail: alessia.pini@unicatt.it)

[3] MOX - Department of Mathematics, Politecnico di Milano, (e-mail: simone.vantini@polimi.it)

ABSTRACT: We propose a method to test locally functional data whose domain is a Riemaniann manifold. The procedure is based on testing hypotheses on a suitably defined family of balls of the domain, and can be applied to a vast variety of different functional tests. The final result is an adjusted p-value function defined on the same domain as functional data, and controlling the ball-wise error rate.

KEYWORDS: functional data, manifolds, permutation tests, adjusted p-value.

1 Introduction

In functional data analysis (FDA), the object of statistical analysis are typically functions modeled as random elements of a Hilbert space. Inference on functional data is particularly challenging since it deals with elements of infinite dimensional spaces. A currently popular topic in FDA is local inference, i.e., the continuous statistical testing of a null hypothesis along the domain of data. The principal issue in this topic is the infinite amount of tested hypotheses, which can be seen as an extreme case of multiple testing. Local inferential techniques are either based on simultaneous confidence bands (Liebl & Reimherr, 2023), or on the definition of a p-value function, that is a function assigning a p-value at each point of the domain. Methods based on such a p-value function typically adjust p-values for guaranteeing a control of a quantity related with the error rate on the whole domain, that could either be related to the family-wise error rate (e.g., Pini & Vantini, 2017, Abramowicz *et al.*, 2022) or to the false discovery rate (e.g., Lundtorp Olsen *et al.*, 2021). In particular, Pini & Vantini, 2017 introduced the interval-wise testing procedure which performs local inference for functional data defined on an interval domain, where the output is an adjusted p-value function that controls for type I errors on intervals. The interval-wise testing procedure provides a control of the interval-wise error rate, that is the probability that, if on an interval the null hypothesis is true, at least one part of it is detected as significant.

Most of the current literature focuses on functional data whose domain is an interval of $\mathbb{R}$. The few exceptions considering more complex domains are based on the false discovery rate control (Lundtorp Olsen *et al.*, 2021), or on an asymptotic control of the family-wise error rate (Abramowicz *et al.*, 2022). In this work, instead, we extend the method proposed by Pini & Vantini, 2017 to functional data defined on manifold domains. The resulting method will provide a finite sample control of the ball-wise error rate, which is an extension of the interval-wise error rate to the multidimensional setting.

We extend this idea to a general setting where domain is a Riemannian manifolds. This requires new methodology such as how to define adjustment sets on product manifolds and how to approximate the test statistic when the domain has non-zero curvature. The resulting method will provide a finite sample control of the ball-wise error rate, which is an extension of the interval-wise error rate to the multidimensional setting. This extended abstract describes an overview of the proposed statistical method. More details on the method, its theoretical properties, a simulation and an application to real data can be found in Lundtorp Olsen *et al.*, 2023.

2 Methods

We will assume that the domain of our functional data are *Riemannian manifolds*. In the following, we give a definition of the manifold, as well as the one of ball, that will be of particular importance to define the error control provided by the method.

Definition 1 *A manifold M of finite dimension is a smooth manifold together with a smoothly varying 2-tensor field g on M which is an inner product at each point. The inner product g defines a metric d and a measure μ on M, which we will refer to as the* Riemannian metric *and the* Riemannian measure*, respectively.*

Definition 2 *For a given manifold M with metric d, define the ball of radius* ε *and center x as*

$$B(x,\varepsilon) = \{y \in M | d(x,y) < \varepsilon\}, \quad x \in M, \varepsilon > 0.$$

Let M be a manifold with metric d. We assume that we have observed n smooth functional data $\xi_1, \ldots, \xi_n$: $M \mapsto \mathbb{R}$. For simplicity of notation, here, functional data are assumed to be observed on a single manifold domain. We refer to Lundtorp Olsen *et al.*, 2023 for a more general version, where the domain can be as well a product of a finite number of manifolds.

Assume that we would like to test at every point $x \in M$, a pointwise null hypothesis $H_0(x)$, against an alternative hypothesis $H_1(x)$. We further assume that hypotheses can be tested by means of a pointwise test statistic $T(x)$, which is stochastically greater under $H_1(x)$ than under $H_0(x)$. Finally, let $p(x)$ denote the unadjusted p-value of the test at point x.

The procedure to define an adjusted p-value function on this setting is based on testing the null and alternative hypothesis on every ball of M of size $\varepsilon \leq r$, with a fixed r (ball-wise testing), and then adjusting the p-values in order to obtain a desired multiplicity control.

Ball-wise testing. Let $B = B(y, \varepsilon)$ be a fixed ball in M. We define the null and alternative hypotheses on the ball as

$$H_0^B : \cap_{x \in B(y,\varepsilon)} H_0(x); \quad H_1^B : \cup_{t \in B(y,\varepsilon)} H_1(x). \tag{1}$$

The hypotheses 1 can be tested with the integral test statistic

$$T^B = \int_B T(x) \mathrm{d}\mu(x) \tag{2}$$

Let p^B be the p-value of the obtained test on ball B. In the ball-wise testing phase, the null and alternative hypotheses H_0^B and H_1^B are tested on every ball $B \in M$ with radius $\varepsilon \leq r$, with a fixed r. The constant r is a parameter of the procedure, and will affect the power and error control of the obtained procedure. We refer to Lundtorp Olsen *et al.*, 2023 for a discussion on the effect of the parameter in the test results.

We here give the general definition of ball-wise hypotheses and p-values. Note that the tests can be performed with any procedure, given that the obtained p-values are exact. In particular, in Lundtorp Olsen *et al.*, 2023 we propose to use permutation tests for testing pointwise and ball-wise hypotheses.

Adjustment. Let $\mathcal{B}$ denote the set of all balls $B \in M$ with radius $\varepsilon \leq r$. The adjusted p-value at point $x \in M$ is defined as

$$\tilde{p}(x) = \sup_{B \in \mathcal{B} : x \in B} p^B. \tag{3}$$

In particular, the null hypothesis $H_0(x)$ is rejected by $\tilde{p}(x)$ at level α only if all null hypotheses on balls $B \in \mathcal{B}$ that contain the point x are also rejected at the same level. This is sufficient to guarantee that the procedure controls the ball-wise error rate Lundtorp Olsen *et al.*, 2023, that is, $\forall \alpha \in (0,1)$:

$$\forall B \in \mathcal{B} : H_0^B \text{ is true}, \mathbb{P}(\exists x \in B : \tilde{p}(x) \leq \alpha) \leq \alpha. \tag{4}$$

References

ABRAMOWICZ, K., PINI, A., SCHELIN, L., SJÖSTEDT DE LUNA, S., STAMM, A., & VANTINI, S. 2022. Domain selection and familywise error rate for functional data: A unified framework. *Biometrics*.

LIEBL, D., & REIMHERR, M. 2023. Fast and fair simultaneous confidence bands for functional parameters. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 04.

LUNDTORP OLSEN, N., PINI, A., & VANTINI, S. 2021. False discovery rate for functional data. *Test*, **30**(3), 784–809.

LUNDTORP OLSEN, N., PINI, A., & VANTINI, S. 2023. *Local inference for functional data on manifold domains using permutation tests*. Tech. rept. arXiv.

PINI, A., & VANTINI, S. 2017. Interval-wise testing for functional data. *Journal of Nonparametric Statistics*, **29**(2), 407–424.