A note on repeated measures analysis for functional data

Łukasz Smaga¹

Received: 3 June 2018 / Accepted: 21 December 2018 / Published online: 4 January 2019
© The Author(s) 2019

Abstract

In this paper, the repeated measures analysis for functional data is considered. The known testing procedures for this problem are based on test statistic being the integral of the difference between sample mean functions, which takes into account only “between group variability”. We modify this test statistic to use also information about “within group variability”. More precisely, we construct the new test statistics being integral and supremum of pointwise test statistic obtained by adapting the classical paired t -test statistic to functional data framework. The testing procedures are based on different methods of approximating the null distribution of the test statistics, namely the Box-type approximation, nonparametric and parametric bootstrap and permutation approaches. These approximations do not perform equally well under finite samples, which is established in simulation experiments, indicating the best new tests. The simulations and an application to mortality data suggest that some of the new procedures outperform the known tests in terms of size control and power.

Keywords Bootstrap · Box-type approximation · Functional data analysis · Hypothesis testing · Permutation · Repeated measures analysis

1 Introduction

In the recent past two decades, the functional data analysis (FDA) dealing with samples of random curves or functions has increased its popularity. The first reason for this is the technological progress allowing to continuously monitor and record data for many processes in biostatistics, chemometrics, econometrics, geophysics, medicine, meteorology etc. Moreover, the increasing computer power allows to conduct an effective analysis of such (high-dimensional) data in a reasonable time. Secondly, the monographs by Ramsay and Silverman (1997, 2002, 2005) offering a broad range of the FDA methods and presenting their applications to real data problems have helped in

✉ Łukasz Smaga
ls@amu.edu.pl

¹ Faculty of Mathematics and Computer Science, Adam Mickiewicz University, ul. Umultowska 87, 61-614 Poznań, Poland

learning the basics of the FDA in an orderly and accessible way for a wide audience. The nonparametric perspective of the FDA is proposed by Ferraty and Vieu (2006). The books by Horváth and Kokoszka (2012) and Zhang (2013) complement the previous monographs on new theoretical and practical aspects of the FDA.

Testing statistical hypothesis about functions or curves is one of the most commonly investigated problems in the FDA. The testing procedures for functional analysis of variance problem, where one wants to check the equality of mean functions in a number of groups of independent functional observations, are very popular. In Cuevas et al. (2004), Górecki and Smaga (2015), Ramsay and Silverman (2005), Zhang (2013), Zhang et al. (2018) and Zhang and Liang (2014), the parametric functional ANOVA was studied by a broad range of different methods. On the other hand, Delicado (2007) considered nonparametric functional ANOVA, while Górecki and Smaga (2017) investigated the multivariate analysis of variance for functional data. Some other hypothesis testing problems are as follows: Gabrys and Kokoszka (2007) proposed a simple portmanteau test of independence for functional observations based on their principal component expansion. In functional regression analysis, the statistical significance of (fixed or functional) coefficients was investigated, for example, by Collazos et al. (2016), Kokoszka et al. (2008) and Maslova et al. (2010). Testing the linearity of a scalar-on-function regression relationship is considered by McLean et al. (2015). Kosiorowski et al. (2017) proposed a novel statistic for detecting a structural change in functional time series based on a local Wilcoxon statistic. The interval testing procedure for functional data in different frameworks (e.g., one or two-population frameworks) was proposed by Pini and Vantini (2017). Their procedure is based on means of different basis expansions. In the case of spatial functional data, Giraldo et al. (2018) proposed the test based on the Mantel statistic for checking spatial autocorrelation, which is an important step in the statistical analysis of spatial data.

In this paper, we consider the repeated measures analysis for functional data, where we have the repeated functional data for the same subjects probably submitted to different conditions. Martínez-Camblor and Corral (2011) proposed first testing procedure for this problem. Their tests are based on test statistic being the integral of the difference between sample mean functions and bootstrap and permutation approaches. Smaga (2017) proposed the test based on the same test statistic, but its distribution was approximated by the Box-type approximation, which resulted in less time-consuming procedure than resampling and permutation methods. Nevertheless, all these tests perform equally well in terms of size control and power under finite samples. Note that since the tests considered in these papers are constructed mainly on asymptotic distribution of the test statistic, it is based only on “between group variability”. In this work, we show how to modify this test statistic to take into account also “within group variability”. Namely, we adapt the usual test statistic of the classical paired two-sample test to functional data framework obtaining an appropriate pointwise test statistic. We construct two test statistics being integral and supremum of the pointwise test statistic. Their distributions are approximated by parametric methods based on derived asymptotic distributions as well as nonparametric bootstrap and permutation approaches. Simulation results indicate that the new testing procedures may not perform equally well under finite samples (in contrast to methods proposed by Martínez-Camblor and Corral 2011, and Smaga 2017), and some of them are not recommended when the

sample size is small. However, the best new permutation tests usually work better than the tests by Martínez-Cambor and Corral (2011) and Smaga (2017), which is seen in simulations and real data application.

The rest of the paper is organized as follows: In Sect. 2, we first formulate the repeated measures analysis for functional data. Then, we propose new test statistics for this problem and consider different methods of approximating their null distribution to construct testing procedures. Section 3 is devoted to Monte Carlo simulation studies providing an idea of finite sample behavior of the new tests and comparing them with the known tests. In Sect. 4, a real data example based on the mortality data is presented. Section 5 concludes the paper. In Appendix, the technical assumptions and the proof of theoretical result are given.

2 Repeated measures analysis for functional data

In this section, we present new testing procedures for the repeated measures analysis for functional data. They can be seen as modifications of the tests proposed in Martínez-Cambor and Corral (2011) and Smaga (2017) by taking into account also “within group variability”.

We formulate the repeated measures analysis for functional data using the notation of Martínez-Cambor and Corral (2011). In this setting, we are interested in checking the equality of two functions obtained from the same subject (probably submitted to different conditions). Assume that $X_1(t), \dots, X_n(t)$, $t \in [0, 2]$ is a functional sample consisting of independent stochastic processes. Since $t \in [0, 2]$, we ignore the (possible) period in which the subjects are not monitored. Additionally, we suppose that the following additivity assumption holds:

$$X_i(t) = m(t) + e_i(t), \quad (1)$$

where $i = 1, \dots, n$, $t \in [0, 2]$, $e_i(t)$ is random function with zero mean and covariance function $\mathbb{C}(s, t)$, $s, t \in [0, 2]$. By (1), the null hypothesis is given by the following formula:

$$H_0 : m(t) = m(t + 1) \quad \text{for all } t \in [0, 1], \quad (2)$$

while the alternative hypothesis is $H_1 : m(t) \neq m(t + 1)$ for some $t \in [0, 1]$.

2.1 Test statistics

In classical paired t -test for sample of random variables $\mathbf{X}_i = (X_{i,1}, X_{i,2})^\top$, $i = 1, \dots, n$, the test statistic is constructed based on the statistic of the form

$$\frac{\sqrt{n}\bar{D}_n}{V_n}, \quad (3)$$

where $\bar{D}_n = n^{-1} \sum_{i=1}^n D_i = \bar{X}_1 - \bar{X}_2$ is the difference of the means, $D_i = X_{i,1} - X_{i,2}$ are the differences of the pairs for $i = 1, \dots, n$ and $V_n^2 = (n-1)^{-1} \sum_{i=1}^n (D_i - \bar{D}_n)^2$

is the sample variance of the D_i 's. Recently, Konietzschke and Pauly (2014) studied various bootstrap and permutation methods to approximate the null distribution of the paired t -test type statistic obtaining better testing procedures than the standard paired t -test.

For paired sample problem for functional data, Martínez-Camblor and Corral (2011) and Smaga (2017) constructed tests by adapting the numerator of (3) ("between group variability") to functional data framework, i.e., they considered the following test statistic

$$C_n = n \int_0^1 (\bar{X}(t) - \bar{X}(t+1))^2 dt,$$

where $\bar{X}(t) = n^{-1} \sum_{i=1}^n X_i(t)$, $t \in [0, 2]$. The tests based on C_n are quite powerful. However, there is a question of adapting the "whole" test statistic (3) ("between and within group variabilities") to repeated functional data to obtain even more powerful testing procedure. The answer seems to be positive. The natural pointwise version of V_n^2 is as follows

$$\begin{aligned} V_n^2(t) &= \frac{1}{n-1} \sum_{i=1}^n (X_i(t) - X_i(t+1) - \bar{X}(t) + \bar{X}(t+1))^2 \\ &= \hat{C}(t, t) - 2\hat{C}(t, t+1) + \hat{C}(t+1, t+1) \\ &= \hat{K}(t, t), \end{aligned}$$

where

$$\hat{C}(s, t) = \frac{1}{n-1} \sum_{i=1}^n (X_i(s) - \bar{X}(s))(X_i(t) - \bar{X}(t)) \quad (4)$$

is the unbiased estimator of the covariance function $\mathbb{C}(s, t)$, $s, t \in [0, 2]$ and

$$\hat{K}(s, t) = \hat{C}(s, t) - \hat{C}(s, t+1) - \hat{C}(s+1, t) + \hat{C}(s+1, t+1) \quad (5)$$

$s, t \in [0, 1]$ is the unbiased estimator of the asymptotic covariance function

$$\mathbb{K}(s, t) = \mathbb{C}(s, t) - \mathbb{C}(s, t+1) - \mathbb{C}(s+1, t) + \mathbb{C}(s+1, t+1)$$

of the process $n^{1/2}(\bar{X}(t) - \bar{X}(t+1))$ under the null hypothesis (2) (see the proof of Theorem 1 in Martínez-Camblor and Corral 2011). Therefore, the pointwise adaptation of the test statistic (3) to the repeated measures analysis for functional data can be of the form

$$\frac{n(\bar{X}(t) - \bar{X}(t+1))^2}{\hat{K}(t, t)}, \quad t \in [0, 1]. \quad (6)$$

To obtain test statistic from (6), we can simply use the integration as in C_n , namely the first new test statistic is

$$\mathcal{D}_n = n \int_0^1 \frac{(\bar{X}(t) - \bar{X}(t+1))^2}{\hat{\mathbb{K}}(t, t)} dt.$$

This test statistic is similar to that of the globalizing pointwise F -test for analysis of variance problem for functional data considered by Zhang and Liang (2014). Although integration seems to be natural operation on (6), some other use of the pointwise test statistic may result in construction of more powerful test. Zhang et al. (2018) based their testing procedure on the supremum of pointwise test statistic for one-way functional analysis of variance obtaining better test than the one of Zhang and Liang (2014) in case of high and moderate correlation between observations at any two different time points. Following this idea, the second new test statistic, which we propose, is of the form

$$\mathcal{E}_n = \sup_{t \in [0, 1]} \left\{ \frac{n(\bar{X}(t) - \bar{X}(t+1))^2}{\hat{\mathbb{K}}(t, t)} \right\}.$$

The tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ can be seen as adaptations of the globalizing pointwise F -test of Zhang and Liang (2014) and the $F_{\max}$ -test of Zhang et al. (2018) to the repeated measures analysis considered, respectively. Namely, the “between group variability” $n(\bar{X}(t) - \bar{X}(t+1))^2$ corresponds to the pointwise between-subject variation $\text{SSR}_n(t)$, while the “within group variability” $\hat{\mathbb{K}}(t, t)$ corresponds to the within-subject variation $\text{SSE}_n(t)$.

In fact, the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ can use more information from the data than the test statistic $\mathcal{C}_n$, which may result in more powerful testing procedures for (2), as we will see in simulation experiments and real data example (Sects. 3 and 4).

2.2 Approximating the null distribution

To construct the tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$, we use different methods to approximate their null distributions. Similarly to Martínez-Camblor and Corral (2011), we start with the use of the asymptotic null distributions of the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$. They are presented in the following result. Its proof and technical assumptions are given and discussed in Appendix. Throughout the paper, $\xrightarrow{d}$ denotes convergence in distribution (Laha and Rohatgi 1979, p. 474), and $X \stackrel{d}{=} Y$ denotes that the random variables X and Y have the same distribution.

Theorem 1 *Under model (1) and the null hypothesis (2):*

1. *If Assumptions A1–A5 given in Appendix are satisfied, then*

$$\mathcal{D}_n \xrightarrow{d} \int_0^1 \frac{(z(t) - z(t+1))^2}{\mathbb{K}(t, t)} dt \stackrel{d}{=} \sum_{k=1}^{\infty} \lambda_k^* A_k, \quad (7)$$

as $n \rightarrow \infty$, where $z(t)$, $t \in [0, 2]$ is a Gaussian process with zero mean and covariance function $\mathbb{C}(s, t)$, $s, t \in [0, 2]$, $\lambda_1^*, \lambda_2^*, \dots$ are the decreasing-ordered

eigenvalues of $\mathbb{K}_*(s, t) = \mathbb{K}(s, t)/(\mathbb{K}(s, s)\mathbb{K}(t, t))^{1/2}$, and $A_1, A_2, \dots$ are the independent random variables following a central chi-squared distribution with one degree of freedom.

2. If Assumptions A1–A6 given in Appendix are satisfied, then

$$\mathcal{E}_n \xrightarrow{d} \sup_{t \in [0, 1]} \left\{ \frac{(z(t) - z(t+1))^2}{\mathbb{K}(t, t)} \right\}, \quad (8)$$

as $n \rightarrow \infty$, where $z(t)$, $t \in [0, 2]$ is the same Gaussian process as in 1.

The asymptotic null distributions of the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ given in (7) and (8) containing the Gaussian process $z(t)$, $t \in [0, 2]$ are not known in full, since they depend on the unknown covariance function $\mathbb{K}(s, t)$, $s, t \in [0, 1]$. However, they can be used to develop a simple Monte Carlo method to estimate the p -values of the tests. For this purpose, we have to generate the Gaussian processes $z_b(t)$, $t \in [0, 2]$, $b = 1, \dots, B$ with zero mean and covariance function $\hat{\mathbb{C}}(s, t)$, $s, t \in [0, 2]$ given by (4), where B is a large number. The p -values of the testing procedures based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ are then as follows

$$\begin{aligned} & \frac{1}{B} \sum_{b=1}^B I \left(\int_0^1 \frac{(z_b(t) - z_b(t+1))^2}{\hat{\mathbb{K}}(t, t)} dt > \mathcal{D}_n \right), \\ & \frac{1}{B} \sum_{b=1}^B I \left(\sup_{t \in [0, 1]} \left\{ \frac{(z_b(t) - z_b(t+1))^2}{\hat{\mathbb{K}}(t, t)} \right\} > \mathcal{E}_n \right) \end{aligned}$$

respectively, where $I(A)$ denotes the indicator function on the set A (takes value 1 if A is true and 0 otherwise) and $\hat{\mathbb{K}}(s, t)$, $s, t \in [0, 1]$ is given by (5). This method is similar to the methods used by Cuevas et al. (2004) and Martínez-Camblor and Corral (2011). We will refer to it as parametric bootstrap method.

By (7), the asymptotic null distribution of $\mathcal{D}_n$ is a central χ^2 -type mixture. The distribution of this mixture can be approximated by the Box-type approximation (Box 1954). Recently, Smaga (2017) used this method to approximate the asymptotic null distribution of $\mathcal{C}_n$. We can easily adapt his results to our purpose, since the only difference between the asymptotic null distributions of $\mathcal{C}_n$ and $\mathcal{D}_n$ is that the eigenvalues are of different covariance functions $\mathbb{K}(s, t)$ and $\mathbb{K}_*(s, t)$, $s, t \in [0, 1]$, respectively. Namely, we approximate the asymptotic null distribution of $\mathcal{D}_n$ by that of $\beta \chi_d^2$, where

$$\beta = \text{tr}(\mathbb{K}_*^{\otimes 2}), \quad d = \frac{1}{\text{tr}(\mathbb{K}_*^{\otimes 2})}, \quad (9)$$

and $\mathbb{K}_*^{\otimes 2}(s, t) = \int_0^1 \mathbb{K}_*(s, u)\mathbb{K}_*(u, t)du$, $s, t \in [0, 1]$, as $\text{tr}(\mathbb{K}_*) = 1$. We obtain the plug-in estimators $\hat{\beta}$ and $\hat{d}$ of β and d using the estimate (5) of the covariance function $\mathbb{K}(s, t)$, $s, t \in [0, 1]$. These estimators are consistent under Assumptions presented in Appendix, which can be proved similarly as Theorem 2.1 in Smaga (2017). Then, the estimated p -value is of the form

$$P\left(\chi_d^2 > \frac{\mathcal{D}_n}{\hat{\beta}}\right).$$

The parametric bootstrap method is much more time-consuming and more liberal than the test based on Box-type approximation (see Sect. 3).

The distribution of a χ^2 -type mixture can be also approximated by exact methods as, for example, the Imhof's algorithm (Imhof 1961). However, the simulation results (not shown) indicate that such algorithms work similarly to the Box-type approximation, but are usually more time-consuming. Thus, we do not consider the exact methods in the paper.

Since the tests mentioned above are constructed based on the asymptotic null distributions of the test statistics, they may need larger number of observations to perform satisfactorily, as we will see in simulations. Thus, we also investigate nonparametric approaches, namely nonparametric bootstrap and permutation procedures, which were considered by Martínez-Camblor and Corral (2011) and Konietzschke and Pauly (2014). These procedures do not require the assumptions as the tests based on the asymptotic null distributions of the test statistics.

In nonparametric bootstrap approach of Martínez-Camblor and Corral (2011), we first select independent bootstrap samples $X_1^{*,k}(t), \dots, X_n^{*,k}(t)$, $t \in [0, 2]$, $k = 1, \dots, K_1$ drawn with replacement from the original sample $X_1(t), \dots, X_n(t)$, $t \in [0, 2]$, where K_1 is a large number. Then the estimated p -values of the tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ are

$$\begin{aligned} & \frac{1}{K_1} \sum_{k=1}^{K_1} I\left(n \int_0^1 \frac{(\bar{X}^{*,k}(t) - \bar{X}(t) + \bar{X}(t+1) - \bar{X}^{*,k}(t+1))^2}{\hat{\mathbb{K}}^{*,k}(t, t)} dt > \mathcal{D}_n\right), \\ & \frac{1}{K_1} \sum_{k=1}^{K_1} I\left(\sup_{t \in [0, 1]} \left\{ \frac{n(\bar{X}^{*,k}(t) - \bar{X}(t) + \bar{X}(t+1) - \bar{X}^{*,k}(t+1))^2}{\hat{\mathbb{K}}^{*,k}(t, t)} \right\} > \mathcal{E}_n\right), \end{aligned}$$

where $\bar{X}^{*,k}(t)$, $t \in [0, 2]$ and $\hat{\mathbb{K}}^{*,k}(s, t)$, $s, t \in [0, 1]$ are the sample mean and the estimator of covariance function $\mathbb{K}(s, t)$ computed on the bootstrap sample. We refer to this method as the nonparametric bootstrap method (I) or NB(I) for short.

Following Konietzschke and Pauly (2014), we also consider other nonparametric bootstrap method, which will be referred to as the nonparametric bootstrap method (II) or NB(II) for short. In this method, we first center the sample, i.e., $X_{1,c}(t) = X_1(t) - \bar{X}(t)$, $\dots$, $X_{n,c}(t) = X_n(t) - \bar{X}(t)$, $t \in [0, 2]$. Then the independent bootstrap samples $X_{1,c}^{*,k}(t), \dots, X_{n,c}^{*,k}(t)$, $t \in [0, 2]$, $k = 1, \dots, K_2$ (K_2 is a large number) are selected in such a way that $X_{1,c}^{*,k}(t), \dots, X_{n,c}^{*,k}(t)$, $t \in [0, 1]$ are randomly drawn with replacement from $X_{1,c}(t), \dots, X_{n,c}(t)$, $t \in [0, 1]$, while independently, $X_{1,c}^{*,k}(t), \dots, X_{n,c}^{*,k}(t)$, $t \in [1, 2]$ are randomly drawn with replacement from $X_{1,c}(t), \dots, X_{n,c}(t)$, $t \in [1, 2]$. Finally, the estimated p -values are of the form

$$\frac{1}{K_2} \sum_{k=1}^{K_2} I\left(n \int_0^1 \frac{(\bar{X}_c^{*,k}(t) - \bar{X}_c^{*,k}(t+1))^2}{\hat{\mathbb{K}}_c^{*,k}(t, t)} dt > \mathcal{D}_n\right),$$

$$\frac{1}{K_2} \sum_{k=1}^{K_2} I \left(\sup_{t \in [0,1]} \left\{ \frac{n(\bar{X}_c^{*,k}(t) - \bar{X}_c^{*,k}(t+1))^2}{\hat{\mathbb{K}}_c^{*,k}(t,t)} \right\} > \mathcal{E}_n \right),$$

where $\bar{X}_c^{*,k}(t)$, $t \in [0, 2]$ and $\hat{\mathbb{K}}_c^{*,k}(s, t)$, $s, t \in [0, 1]$ are the sample mean and the estimator of covariance function $\mathbb{K}(s, t)$ computed on the bootstrap sample.

Note that the nonparametric bootstrap tests take into account the null hypothesis at the moment of computing the value of the test statistics instead of at the resampling moment, as it is suggested, for example, in Martínez-Camblor and Corral (2011, 2012).

On the other hand, the null hypothesis is involved at the moment to draw the permutation sample. Namely, the permutation method is based on selected independent permutation samples $X_1^{*,l}(t), \dots, X_n^{*,l}(t)$, $t \in [0, 2]$, $l = 1, \dots, L$, where L is a large number and $X_i^{*,l}(t) = X_i(t)$ for $t \in [0, 2]$ (with probability $1/2$) or $X_i^{*,l}(t) = X_i(t+1)$ and $X_i^{*,l}(t+1) = X_i(t)$ for $t \in [0, 1]$, $i = 1, \dots, n$, and then the p -values of the permutation tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ are given by

$$\begin{aligned} & \frac{1}{L} \sum_{l=1}^L I \left(n \int_0^1 \frac{(\bar{X}^{*,l}(t) - \bar{X}^{*,l}(t+1))^2}{\hat{\mathbb{K}}^{*,l}(t,t)} dt > \mathcal{D}_n \right), \\ & \frac{1}{L} \sum_{l=1}^L I \left(\sup_{t \in [0,1]} \left\{ \frac{n(\bar{X}^{*,l}(t) - \bar{X}^{*,l}(t+1))^2}{\hat{\mathbb{K}}^{*,l}(t,t)} \right\} > \mathcal{E}_n \right), \end{aligned}$$

where $\bar{X}^{*,l}(t)$, $t \in [0, 2]$ and $\hat{\mathbb{K}}^{*,l}(s, t)$, $s, t \in [0, 1]$ are the sample mean and the estimator of covariance function $\mathbb{K}(s, t)$ computed on the permutation sample.

Simulation and real data example results in Martínez-Camblor and Corral (2011) and Smaga (2017) indicate that the all tests (namely, using parametric and nonparametric bootstrap, permutation approach and Box-type approximation) based on the test statistic $\mathcal{C}_n$ perform very similarly under finite samples in terms of size control and power. However, this is not true for the testing procedures based on the new test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ as we will see in the next sections.

3 Simulation studies

In this section, the simulation studies are performed to investigate the behavior of the tests proposed in Sect. 2 with regard to maintaining the pre-assigned type-I error level under the null hypothesis, and the power under alternatives. Since all testing procedures based on the test statistic $\mathcal{C}_n$ behave very similarly under finite samples, we considered only the least time-consuming BT test by Smaga (2017), as a competitor to the new tests.

3.1 Simulation setup

The paired functional samples are generated according to the formulas: $X_i(t) = m_1(t) + e_{i1}(t)$ and $X_i(t+1) = m_2(t) + e_{i2}(t)$ for $t \in [0, 1]$, $i = 1, \dots, n$. The number of observations n is set to 15, 25, 35.

We consider the following three models of particular forms of mean functions:

$$\begin{aligned} \text{M1 } m_1(t) &= m_2(t) = (\sin(2\pi t^2))^5 I_{[0,1]}(t), \\ \text{M2 } m_1(t) &= (\sin(2\pi t^2))^5 I_{[0,1]}(t), m_2(t) = (\sin(2\pi t^2))^7 I_{[0,1]}(t), \\ \text{M3 } m_1(t) &= (\sin(2\pi t^2))^5 I_{[0,1]}(t), m_2(t) = (\sin(2\pi t^2))^3 I_{[0,1]}(t). \end{aligned}$$

The null (resp. alternative) hypothesis holds under Model M1 (resp. Models M2 and M3). Models M1, M2 and M3 are very similar to simulation models M4, M6 and M5 of Martínez-Camblor and Corral (2011). Figure 3 of that paper shows the shapes of the above mean functions. We also considered other models of Martínez-Camblor and Corral (2011), but the results were very similar or uninteresting (e.g., the empirical powers were always equal to 100%), so they are omitted to save space.

The following three settings of different error types were considered:

$$\begin{aligned} \text{Normal setting: } e_{i1}(t) &= 0.5B_{i1}(t), e_{i2}(t) = \rho e_{i1}(t) + 0.5\sqrt{1-\rho^2}B_{i2}(t), \\ \text{Lognormal setting: } e_{i1}(t) &= \exp(0.5B_{i1}(t)), e_{i2}(t) = \exp(\rho e_{i1}(t) + 0.5\sqrt{1-\rho^2}B_{i2}(t)), \\ \text{Mixed setting: } e_{i1}(t) &= 0.5B_{i1}(t), e_{i2}(t) = \exp(\rho e_{i1}(t) + 0.5\sqrt{1-\rho^2}B_{i2}(t)), \end{aligned}$$

where $i = 1, \dots, n$, $t \in [0, 1]$, B_{i1} and B_{i2} are two independent standard Brownian Bridges, $\rho = 0, 0.25, 0.5, 0.75$. The error functions $\exp(e_{ij}(t))$, $i = 1, \dots, n$, $j = 1, 2$ are adequately centered.

Since the functional data are not usually continuously observed in practice, the trajectories of $X_1(t), \dots, X_n(t)$, $t \in [0, 2]$ are discretized at design time points $t_1, \dots, t_I, t_1 + 1, \dots, t_I + 1$, where t_i , $i = 1, \dots, I$ are equispaced in $[0, 1]$. We consider $I = 26, 101$.

The empirical sizes and powers of the tests were estimated based on 1000 simulation samples. The p -values of the bootstrap and permutation testing procedures were computed from 1000 replications. For simplicity, the nominal significance level $\alpha = 5\%$ is chosen. The simulation experiments as well as the real data example of Sect. 4 were conducted in the R program (R Core Team 2017).

3.2 Simulation results

The empirical sizes of the tests obtained under Model M1 are given in Table 1. First of all, we observe that the parametric bootstrap test based on the test statistic $\mathcal{E}_n$ tends to be highly liberal in all cases. The tests based on the test statistic $\mathcal{D}_n$ and the Box-type approximation and parametric bootstrap approach are also too liberal especially for greater number of design time points ($I = 101$) or small number of observations ($n = 15$). Thus, the tests based on the asymptotic null distributions of the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ may require a larger number of observations to maintain the pre-assigned type-I error rate. On the other hand, the nonparametric bootstrap testing procedures usually have conservative character, which is especially evident in case

Table 1 Empirical sizes (as percentages) of the tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ proposed in Section 2 and of the BT test based on $\mathcal{C}_n$ by Smaga (2017) obtained in Model M1

S	n	ρ	$I = 26$										$I = 101$																					
			$\mathcal{C}_n$					$\mathcal{D}_n$					$\mathcal{E}_n$					$\mathcal{C}_n$					$\mathcal{D}_n$					$\mathcal{E}_n$						
			BT		PB		NB(I)	NB(II)		P	BT		PB		NB(I)	NB(II)		P	BT		PB		NB(I)	NB(II)		P	BT		PB		NB(I)	NB(II)		P
			BT	PB	BT	PB		BT	PB		BT	PB	BT	PB		BT	PB		BT	PB	BT	PB		BT	PB		BT	PB	BT	PB		BT	PB	
N	15	0.00	5.6	7.2	7.0	3.2	3.6	4.7	14.0	2.0	2.7	4.5	6.5	8.6	8.4	2.8	3.8	4.6	20.4	1.2	3.0	4.6	4.6	20.4	1.2	3.0	4.6	4.6	20.4	1.2	3.0	4.6	4.6	
		0.25	5.7	7.5	7.5	3.0	3.7	4.7	14.6	1.6	3.3	4.6	6.6	8.7	8.6	2.1	3.7	5.1	19.1	1.5	3.7	5.1	19.1	1.5	3.7	5.1	19.1	1.5	3.7	5.1	3.7	4.6		
		0.50	6.5	7.8	7.9	2.8	3.8	5.0	14.5	1.3	3.9	4.9	6.5	8.6	8.8	2.6	3.6	5.1	18.9	1.2	3.7	5.1	18.9	1.2	3.7	5.1	18.9	1.2	3.7	5.1	4.8	4.8		
		0.75	6.4	8.0	7.9	3.2	4.8	5.4	15.2	1.4	4.8	5.2	6.2	8.0	7.9	2.7	3.7	4.9	20.3	1.0	3.7	4.9	20.3	1.0	3.7	4.9	20.3	1.0	3.7	4.9	4.2	4.2		
	25	0.00	5.1	5.9	6.1	3.3	3.6	4.7	9.9	3.1	3.9	5.2	5.0	6.8	6.4	3.3	3.9	4.1	11.0	3.0	3.3	3.9	4.1	11.0	3.0	3.3	3.9	4.1	11.0	3.0	3.3	4.7	4.7	
		0.25	4.2	5.4	5.5	3.6	3.8	4.2	9.2	3.2	3.7	4.6	5.2	6.0	5.7	3.3	3.6	4.4	10.2	2.6	3.5	4.4	10.2	2.6	3.5	4.4	10.2	2.6	3.5	4.4	4.5	4.5		
		0.50	4.4	5.1	5.3	3.7	3.7	4.2	8.7	2.8	3.9	4.6	5.3	5.8	5.7	3.5	3.7	4.5	9.7	2.6	3.9	4.5	9.7	2.6	3.9	4.5	9.7	2.6	3.9	4.5	4.4	4.4		
		0.75	5.2	5.8	6.0	3.5	4.0	4.2	9.8	2.5	3.9	4.8	4.6	5.7	5.8	3.1	3.8	4.2	10.9	3.0	4.3	4.2	10.9	3.0	4.3	4.2	10.9	3.0	4.3	4.2	5.2	5.2		
	35	0.00	5.3	5.9	6.0	4.3	4.5	4.9	8.8	3.4	3.9	4.5	5.3	5.9	6.3	4.4	4.9	4.9	8.0	2.7	3.5	4.9	8.0	2.7	3.5	4.9	8.0	2.7	3.5	4.9	4.5	4.5		
		0.25	4.7	5.7	5.8	4.5	4.3	4.9	8.2	3.8	4.4	4.9	4.9	5.8	6.0	4.2	4.8	5.1	8.5	2.8	4.2	4.8	5.1	8.5	2.8	4.2	4.8	5.1	8.5	2.8	4.2	4.1	4.1	
		0.50	4.5	5.6	5.9	3.9	4.1	4.3	8.7	3.7	4.4	5.3	5.1	5.9	6.4	4.1	4.6	5.0	8.6	3.1	4.1	4.6	5.0	8.6	3.1	4.1	4.6	5.0	8.6	3.1	4.1	4.3	4.3	
		0.75	4.8	5.6	5.7	3.8	4.7	4.6	9.4	3.9	5.5	5.6	4.9	5.9	6.5	3.7	4.5	4.8	8.2	2.4	3.9	4.5	4.8	8.2	2.4	3.9	4.5	4.8	8.2	2.4	3.9	4.4	4.4	
L	15	0.00	3.4	7.0	6.9	1.1	2.0	4.3	16.7	2.1	2.9	4.9	4.6	8.0	8.0	2.6	3.7	5.0	21.2	1.2	2.5	4.7	5.0	21.2	1.2	2.5	4.7	5.0	21.2	1.2	2.5	4.7	4.7	
		0.25	3.4	6.6	7.4	1.3	2.6	3.9	16.7	2.0	3.4	5.2	4.5	8.6	8.6	2.4	3.4	4.8	22.3	1.0	3.4	4.8	4.8	22.3	1.0	3.4	4.8	4.8	22.3	1.0	3.4	5.3	5.3	
		0.50	3.5	7.0	7.2	1.2	2.9	4.5	17.3	1.8	3.8	5.1	4.4	8.7	8.5	2.4	3.4	5.4	22.6	1.3	3.8	5.4	5.4	22.6	1.3	3.8	5.4	5.4	22.6	1.3	3.8	5.4	5.4	
		0.75	4.5	8.3	7.8	1.2	3.3	5.0	15.7	1.8	4.4	5.8	5.0	8.3	8.7	2.0	3.5	5.4	21.2	1.4	4.5	5.4	5.4	21.2	1.4	4.5	5.4	5.4	21.2	1.4	4.5	5.8	5.8	
	25	0.00	3.5	5.2	5.0	2.9	2.8	4.0	10.6	2.1	2.7	4.2	5.2	6.4	6.7	4.1	4.3	5.2	12.0	2.0	3.1	4.3	5.2	12.0	2.0	3.1	4.3	5.2	12.0	2.0	3.1	4.0	4.0	
		0.25	3.4	5.8	5.7	2.8	2.9	4.0	10.4	2.6	3.4	4.6	4.9	6.0	6.0	3.5	4.1	5.2	12.0	1.7	2.8	4.1	5.2	12.0	1.7	2.8	4.1	5.2	12.0	1.7	2.8	4.3	4.3	
		0.50	3.4	4.9	4.6	2.5	2.9	4.1	10.4	2.4	3.7	4.8	4.4	6.1	5.8	3.0	3.5	4.4	11.4	1.9	3.7	3.5	4.4	11.4	1.9	3.7	3.5	4.4	11.4	1.9	3.7	4.1	4.1	
		0.75	3.9	5.7	6.3	2.9	3.3	4.5	9.6	2.1	4.2	4.9	4.6	5.9	5.8	3.5	3.6	4.5	11.7	2.1	3.5	3.6	4.5	11.7	2.1	3.5	3.6	4.5	11.7	2.1	3.5	4.2	4.2	

Table 1 continued

S	n	ρ	$I = 26$												$I = 101$																																																																																												
			C_n						$\mathcal{D}_n$						C_n						$\mathcal{D}_n$																																																																																						
			$\mathcal{E}_n$			$\mathcal{D}_n$			$\mathcal{E}_n$			$\mathcal{D}_n$			$\mathcal{E}_n$			$\mathcal{D}_n$			$\mathcal{E}_n$			$\mathcal{D}_n$																																																																																			
			BT	PB	NB(II)	P	PB	NB(I)	NB(II)	P	BT	PB	NB(I)	NB(II)	P	BT	PB	NB(I)	NB(II)	P	BT	PB	NB(I)	NB(II)	P	BT	PB	NB(I)	NB(II)	P																																																																													
35	0.00	4.1	5.2	5.0	2.5	2.9	4.2	8.5	3.2	3.9	4.5	5.1	6.4	6.2	4.4	4.6	5.6	10.1	3.4	4.2	5.2	M	15	0.00	4.6	7.6	7.5	1.8	2.9	4.2	15.1	1.7	2.7	4.6	5.4	8.7	9.1	2.7	4.0	5.7	21.5	0.9	3.1	5.9	0.25	4.9	8.3	7.6	2.5	3.9	4.7	15.3	1.6	3.2	5.0	5.8	9.0	9.3	2.7	4.5	5.6	20.9	1.1	3.0	5.9	0.50	4.3	5.6	5.5	3.3	3.6	4.8	9.3	2.8	3.5	4.4	4.9	6.0	5.6	4.1	4.7	5.0	8.9	2.9	3.4	3.9	0.75	4.4	5.8	5.9	3.2	3.8	4.4	8.6	2.9	3.3	4.1	4.5	5.5	5.6	4.0	4.4	4.8	9.9	3.1	4.5	4.4
		4.1	5.2	5.0	2.5	2.9	4.2	8.5	3.2	3.9	4.5	5.1	6.4	6.2	4.4	4.6	5.6	10.1	3.4	4.2	5.2				4.6	7.6	7.5	1.8	2.9	4.2	15.1	1.7	2.7	4.6	5.4	8.7	9.1	2.7	4.0	5.7	21.5	0.9	3.1	5.9		4.9	8.3	7.6	2.5	3.9	4.7	15.3	1.6	3.2	5.0	5.8	9.0	9.3	2.7	4.5	5.6	20.9	1.1	3.0	5.9		4.3	5.6	5.5	3.3	3.6	4.8	9.3	2.8	3.5	4.4	4.9	6.0	5.6	4.1	4.7	5.0	8.9	2.9	3.4	3.9		4.4	5.8	5.9	3.2	3.8	4.4	8.6	2.9	3.3	4.1	4.5	5.5	5.6	4.0	4.4	4.8	9.9	3.1	4.5	4.4
		4.1	5.2	5.0	2.5	2.9	4.2	8.5	3.2	3.9	4.5	5.1	6.4	6.2	4.4	4.6	5.6	10.1	3.4	4.2	5.2				4.6	7.6	7.5	1.8	2.9	4.2	15.1	1.7	2.7	4.6	5.4	8.7	9.1	2.7	4.0	5.7	21.5	0.9	3.1	5.9		4.9	8.3	7.6	2.5	3.9	4.7	15.3	1.6	3.2	5.0	5.8	9.0	9.3	2.7	4.5	5.6	20.9	1.1	3.0	5.9		4.3	5.6	5.5	3.3	3.6	4.8	9.3	2.8	3.5	4.4	4.9	6.0	5.6	4.1	4.7	5.0	8.9	2.9	3.4	3.9		4.4	5.8	5.9	3.2	3.8	4.4	8.6	2.9	3.3	4.1	4.5	5.5	5.6	4.0	4.4	4.8	9.9	3.1	4.5	4.4
		4.1	5.2	5.0	2.5	2.9	4.2	8.5	3.2	3.9	4.5	5.1	6.4	6.2	4.4	4.6	5.6	10.1	3.4	4.2	5.2				4.6	7.6	7.5	1.8	2.9	4.2	15.1	1.7	2.7	4.6	5.4	8.7	9.1	2.7	4.0	5.7	21.5	0.9	3.1	5.9		4.9	8.3	7.6	2.5	3.9	4.7	15.3	1.6	3.2	5.0	5.8	9.0	9.3	2.7	4.5	5.6	20.9	1.1	3.0	5.9		4.3	5.6	5.5	3.3	3.6	4.8	9.3	2.8	3.5	4.4	4.9	6.0	5.6	4.1	4.7	5.0	8.9	2.9	3.4	3.9		4.4	5.8	5.9	3.2	3.8	4.4	8.6	2.9	3.3	4.1	4.5	5.5	5.6	4.0	4.4	4.8	9.9	3.1	4.5	4.4
		4.1	5.2	5.0	2.5	2.9	4.2	8.5	3.2	3.9	4.5	5.1	6.4	6.2	4.4	4.6	5.6	10.1	3.4	4.2	5.2				4.6	7.6	7.5	1.8	2.9	4.2	15.1	1.7	2.7	4.6	5.4	8.7	9.1	2.7	4.0	5.7	21.5	0.9	3.1	5.9		4.9	8.3	7.6	2.5	3.9	4.7	15.3	1.6	3.2	5.0	5.8	9.0	9.3	2.7	4.5	5.6	20.9	1.1	3.0	5.9		4.3	5.6	5.5	3.3	3.6	4.8	9.3	2.8	3.5	4.4	4.9	6.0	5.6	4.1	4.7	5.0	8.9	2.9	3.4	3.9		4.4	5.8	5.9	3.2	3.8	4.4	8.6	2.9	3.3	4.1	4.5	5.5	5.6	4.0	4.4	4.8	9.9	3.1	4.5	4.4
25	0.00	4.9	7.2	6.6	3.4	4.3	5.2	11.6	3.5	4.3	5.8	7.3	9.7	9.5	5.5	6.2	5.7	14.9	3.9	4.5	5.9	35	0.00	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3	0.25	5.0	6.9	7.0	3.6	5.3	5.0	10.0	3.6	3.8	5.2	6.6	8.9	9.0	5.6	8.0	6.2	14.4	4.4	5.0	6.0	0.50	5.1	6.5	6.9	3.4	8.1	4.6	9.9	3.4	3.0	5.6	7.0	9.0	9.0	5.5	10.4	6.3	13.9	5.4	4.1	6.0	0.75	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3	
		4.9	7.2	6.6	3.4	4.3	5.2	11.6	3.5	4.3	5.8	7.3	9.7	9.5	5.5	6.2	5.7	14.9	3.9	4.5	5.9			5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3		5.0	6.9	7.0	3.6	5.3	5.0	10.0	3.6	3.8	5.2	6.6	8.9	9.0	5.6	8.0	6.2	14.4	4.4	5.0	6.0		0.75	5.1	6.5	6.9	3.4	8.1	4.6	9.9	3.4	3.0	5.6	7.0	9.0	9.0	5.5	10.4	6.3	13.9	5.4	4.1		6.0	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3
		4.9	7.2	6.6	3.4	4.3	5.2	11.6	3.5	4.3	5.8	7.3	9.7	9.5	5.5	6.2	5.7	14.9	3.9	4.5	5.9			5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3		5.0	6.9	7.0	3.6	5.3	5.0	10.0	3.6	3.8	5.2	6.6	8.9	9.0	5.6	8.0	6.2	14.4	4.4	5.0	6.0		0.75	5.1	6.5	6.9	3.4	8.1	4.6	9.9	3.4	3.0	5.6	7.0	9.0	9.0	5.5	10.4	6.3	13.9	5.4	4.1		6.0	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3
		4.9	7.2	6.6	3.4	4.3	5.2	11.6	3.5	4.3	5.8	7.3	9.7	9.5	5.5	6.2	5.7	14.9	3.9	4.5	5.9			5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3		5.0	6.9	7.0	3.6	5.3	5.0	10.0	3.6	3.8	5.2	6.6	8.9	9.0	5.6	8.0	6.2	14.4	4.4	5.0	6.0		0.75	5.1	6.5	6.9	3.4	8.1	4.6	9.9	3.4	3.0	5.6	7.0	9.0	9.0	5.5	10.4	6.3	13.9	5.4	4.1		6.0	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3
		4.9	7.2	6.6	3.4	4.3	5.2	11.6	3.5	4.3	5.8	7.3	9.7	9.5	5.5	6.2	5.7	14.9	3.9	4.5	5.9			5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3		5.0	6.9	7.0	3.6	5.3	5.0	10.0	3.6	3.8	5.2	6.6	8.9	9.0	5.6	8.0	6.2	14.4	4.4	5.0	6.0		0.75	5.1	6.5	6.9	3.4	8.1	4.6	9.9	3.4	3.0	5.6	7.0	9.0	9.0	5.5	10.4	6.3	13.9	5.4	4.1		6.0	5.1	6.6	6.9	3.7	4.6	5.2	10.6	3.7	4.3	5.3	6.8	9.0	8.8	5.3	6.8	6.1	13.6	4.1	4.8	6.3
35	0.00	6.3	8.0	7.5	5.3	5.9	5.9	12.6	4.7	5.3	6.1	6.3	7.9	8.3	4.9	5.4	6.2	12.9	4.4	4.8	5.8	0.25	6.2	7.5	7.9	5.1	6.7	6.2	11.8	4.6	5.1	5.8	6.2	7.5	7.2	4.6	6.4	5.9	12.5	4.8	4.9	6.1	0.50	6.4	7.5	7.9	5.5	7.5	6.1	11.2	5.9	5.2	5.6	6.0	6.9	6.9	4.5	6.9	5.5	12.2	4.3	4.2	6.1	0.75	6.7	7.5	7.8	5.9	9.1	6.4	11.4	5.9	4.7	6.3	5.9	6.1	6.4	4.6	9.0	5.5	11.8	5.2	3.2	6.1																							
		6.3	8.0	7.5	5.3	5.9	5.9	12.6	4.7	5.3	6.1	6.3	7.9	8.3	4.9	5.4	6.2	12.9	4.4	4.8	5.8		6.2	7.5	7.9	5.1	6.7	6.2	11.8	4.6	5.1	5.8	6.2	7.5	7.2	4.6	6.4	5.9	12.5	4.8	4.9	6.1		6.4	7.5	7.9	5.5	7.5	6.1	11.2	5.9	5.2	5.6	6.0	6.9	6.9	4.5	6.9	5.5	12.2	4.3	4.2	6.1		6.7	7.5	7.8	5.9	9.1	6.4	11.4	5.9	4.7	6.3	5.9	6.1	6.4	4.6	9.0	5.5	11.8	5.2	3.2	6.1																							
		6.3	8.0	7.5	5.3	5.9	5.9	12.6	4.7	5.3	6.1	6.3	7.9	8.3	4.9	5.4	6.2	12.9	4.4	4.8	5.8		6.2	7.5	7.9	5.1	6.7	6.2	11.8	4.6	5.1	5.8	6.2	7.5	7.2	4.6	6.4	5.9	12.5	4.8	4.9	6.1		6.4	7.5	7.9	5.5	7.5	6.1	11.2	5.9	5.2	5.6	6.0	6.9	6.9	4.5	6.9	5.5	12.2	4.3	4.2	6.1		6.7	7.5	7.8	5.9	9.1	6.4	11.4	5.9	4.7	6.3	5.9	6.1	6.4	4.6	9.0	5.5	11.8	5.2	3.2	6.1																							
		6.3	8.0	7.5	5.3	5.9	5.9	12.6	4.7	5.3	6.1	6.3	7.9	8.3	4.9	5.4	6.2	12.9	4.4	4.8	5.8		6.2	7.5	7.9	5.1	6.7	6.2	11.8	4.6	5.1	5.8	6.2	7.5	7.2	4.6	6.4	5.9	12.5	4.8	4.9	6.1		6.4	7.5	7.9	5.5	7.5	6.1	11.2	5.9	5.2	5.6	6.0	6.9	6.9	4.5	6.9	5.5	12.2	4.3	4.2	6.1		6.7	7.5	7.8	5.9	9.1	6.4	11.4	5.9	4.7	6.3	5.9	6.1	6.4	4.6	9.0	5.5	11.8	5.2	3.2	6.1																							
		6.3	8.0	7.5	5.3	5.9	5.9	12.6	4.7	5.3	6.1	6.3	7.9	8.3	4.9	5.4	6.2	12.9	4.4	4.8	5.8		6.2	7.5	7.9	5.1	6.7	6.2	11.8	4.6	5.1	5.8	6.2	7.5	7.2	4.6	6.4	5.9	12.5	4.8	4.9	6.1		6.4	7.5	7.9	5.5	7.5	6.1	11.2	5.9	5.2	5.6	6.0	6.9	6.9	4.5	6.9	5.5	12.2	4.3	4.2	6.1		6.7	7.5	7.8	5.9	9.1	6.4	11.4	5.9	4.7	6.3	5.9	6.1	6.4	4.6	9.0	5.5	11.8	5.2	3.2	6.1																							

The column "S" refers to settings of error types (N, normal; L, lognormal; M, mixed). Moreover, BT, PB, NB(I), NB(II) and P refer to Box-type approximation, parametric bootstrap, nonparametric bootstrap and permutation approach, respectively

of small sample size ($n = 15$). In most cases, the nonparametric bootstrap method (II) gives better results than the nonparametric bootstrap method (I), i.e., NB(II) is less conservative than NB(I). However, we also observe that the test based on the test statistic $\mathcal{D}_n$ and the nonparametric bootstrap method (II) may have too liberal character under mixed setting, that is, under non-exchangeable distributions. The BT test and the permutation tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ control the nominal type-I error level, but the BT test may tend to highly over-reject the null hypothesis when the number of observations is small or the distributions in the groups are different (mixed setting).

Now, we investigate the empirical power of tests, which do not have too liberal character in most cases, namely the BT test and the nonparametric bootstrap and permutation tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$. (Note, however, that the power of the BT test may be overstated under normal and mixed settings. The same is true for the test based on the test statistics $\mathcal{D}_n$ and the nonparametric bootstrap method (II) under mixed setting.) The empirical power of the other tests is not comparable, as they do not maintain the type-I error rate. The empirical powers obtained under Models M2 and M3 are presented in Tables 2 and 3.

In general, the empirical power increases when the number of observations or the correlation between curves also increases. The estimated statistical powers are often comparable for different numbers of design time points. However, it may happen that they increase with I , which is indicated by simulation results not shown. Unfortunately, they may also decrease with I for the tests based on the test statistic $\mathcal{E}_n$ (especially nonparametric bootstrap method (I)) when the number of observations is small.

The empirical powers of the nonparametric bootstrap tests are usually smaller than these of the permutation procedures, which follows from conservativity of the first tests. From permutation tests, the test based on $\mathcal{E}_n$ outperforms that based on $\mathcal{D}_n$, but sometimes their powers are comparable. In the group of nonparametric bootstrap methods, the NB(II) approach usually has greater empirical power, as it is less conservative than the NB(I) method. Under normal and mixed settings, the permutation test based on $\mathcal{D}_n$ works better than the BT test, except under mixed setting and $\rho = 0.75$, where the empirical powers of these tests are comparable. Under the same settings, the nonparametric bootstrap tests based on $\mathcal{E}_n$ usually have greater power than the permutation procedure based on $\mathcal{D}_n$. Under lognormal setting, the empirical powers of the BT test, permutation test based on $\mathcal{D}_n$ and nonparametric bootstrap tests based on $\mathcal{E}_n$ are comparable for $n = 25, 35$, but when $n = 15$, the last three tests perform usually better, worse than the others, respectively. The conservativity of the nonparametric bootstrap tests based on $\mathcal{D}_n$ implies that they often have smaller power than the BT test.

To summarize, the permutation test based on $\mathcal{E}_n$ usually outperforms the other tests, but the permutation procedure based on $\mathcal{D}_n$ also works quite well and has larger power than the BT test in most cases. The nonparametric bootstrap tests have conservative character especially in case of small sample size, which may result in the loss of power. On the other hand, the asymptotic tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ are not recommended because of their liberality.

The finite sample behavior of the tests under the null and alternative hypotheses discussed here will be seen in the real data application of the next section.

Table 2 Empirical powers (as percentages) of the nonparametric bootstrap and permutation tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ proposed in Sect. 2 and of the BT test based on $\mathcal{C}_n$ by Smaga (2017) obtained in Model M2

S	n	ρ	$I = 26$										$I = 101$									
			C_n					$\mathcal{E}_n$					C_n					$\mathcal{E}_n$				
			BT	NB(I)	NB(II)	P	P	NB(I)	NB(II)	P	P	BT	NB(I)	NB(II)	P	P	NB(I)	NB(II)	P	P		
N	15	0.00	14.5	8.4	10.9	14.3	17.3	24.7	30.3	16.9	9.7	12.4	16.7	11.6	21.6	30.1						
		0.25	19.0	11.3	15.0	21.1	24.0	35.4	41.9	21.4	13.4	15.9	21.9	18.5	32.3	39.3						
		0.50	28.3	20.9	27.7	35.9	40.9	58.4	62.5	30.8	21.3	27.5	36.1	29.9	51.0	56.6						
		0.75	62.0	55.9	67.4	75.4	83.9	93.1	94.4	63.2	57.1	69.1	74.9	73.4	90.1	91.7						
	25	0.00	21.6	20.8	22.7	29.0	52.4	56.7	60.7	21.8	20.2	21.9	28.0	47.5	53.6	57.9						
		0.25	32.2	32.9	36.0	43.2	71.6	75.7	79.4	31.7	32.6	35.4	41.9	65.6	71.2	74.7						
		0.50	52.9	57.2	62.8	68.0	92.3	94.6	95.3	53.8	58.4	63.3	69.2	87.9	91.7	92.2						
		0.75	94.3	97.3	98.9	98.9	100	100	100	94.0	96.2	97.6	98.1	100	100	100						
	35	0.00	30.3	36.5	37.9	43.3	78.6	79.2	81.0	31.7	35.8	37.9	42.7	74.2	76.9	79.2						
		0.25	46.1	56.1	58.0	63.2	91.7	92.8	93.4	46.2	55.7	57.2	64.0	90.2	91.9	93.3						
		0.50	74.2	84.1	87.2	88.6	98.9	99.3	99.3	73.9	85.3	87.4	89.6	99.0	99.4	99.3						
		0.75	99.8	99.8	99.9	100	100	100	100	99.5	100	100	100	100	100	100						
L	15	0.00	32.6	21.7	28.6	37.4	23.0	31.4	41.6	35.3	22.7	27.2	36.4	16.4	30.5	41.7						
		0.25	45.3	31.0	40.8	49.5	32.2	45.9	55.8	46.8	32.1	41.6	50.5	25.5	44.0	55.4						
		0.50	67.5	53.1	63.9	72.4	52.4	69.8	77.3	69.3	52.5	64.2	73.0	44.5	68.3	76.7						
		0.75	96.1	90.4	95.1	97.3	89.5	97.2	98.0	97.0	90.7	95.9	97.7	86.0	97.0	98.1						
	25	0.00	60.8	54.4	56.1	65.0	64.2	68.2	72.0	65.6	57.7	61.7	69.4	64.6	70.5	76.2						
		0.25	78.5	72.7	75.2	80.5	78.8	83.8	86.7	81.1	74.9	78.0	82.5	80.5	85.2	88.4						
		0.50	94.3	91.4	93.5	95.3	95.3	97.2	97.9	95.8	92.3	94.4	96.4	94.6	97.1	97.8						
		0.75	100	99.9	100	100	99.9	100	100	100	99.9	99.9	99.9	100	100	100						

Table 2 continued

S	n	ρ	$I = 26$						$I = 101$					
			C_n			$\mathcal{E}_n$			C_n			$\mathcal{E}_n$		
			BT	NB(I)	P	NB(II)	P	NB(I)	BT	NB(I)	P	NB(II)	P	NB(I)
35	0.00	0.00	81.4	78.2	81.3	81.3	83.9	87.1	88.7	90.2	27.3	81.0	84.6	87.3
			93.5	90.6	92.8	92.8	94.1	95.9	96.8	97.5	27.3	93.5	94.8	95.5
			99.5	99.4	99.5	99.5	99.6	99.6	99.7	99.6	27.3	99.6	99.7	99.9
			100	100	100	100	100	100	100	100	27.3	100	100	100
			100	100	100	100	100	100	100	100	27.3	100	100	100
M	15	0.00	16.6	9.1	13.2	13.2	20.7	13.4	19.4	27.3	27.3	15.5	21.1	8.1
			20.0	11.9	17.6	17.6	23.7	17.1	24.7	32.4	32.4	18.8	23.5	13.3
			23.8	14.7	24.1	24.1	26.2	23.5	29.4	39.3	39.3	24.0	27.5	18.6
			27.4	17.6	32.9	32.9	28.2	39.3	40.4	56.7	56.7	35.0	30.7	36.5
			33.3	30.2	33.0	33.0	38.7	44.6	48.8	54.0	54.0	30.8	38.4	41.6
25	0.00	0.25	41.1	35.1	39.9	39.9	44.6	55.3	58.7	64.0	64.0	40.8	45.2	53.1
			49.0	40.7	52.2	52.2	50.6	69.5	70.3	75.8	75.8	55.1	53.8	71.3
			56.9	46.8	69.0	69.0	57.5	86.6	83.3	91.4	91.4	73.8	58.8	89.2
			49.1	49.4	51.0	51.0	56.0	71.0	72.7	75.7	75.7	51.0	56.4	68.9
			59.0	56.9	61.3	61.3	64.4	79.8	81.3	84.4	84.4	62.9	64.1	79.8
35	0.25	0.50	70.6	65.6	76.6	76.6	72.9	92.3	91.3	93.2	93.2	77.1	72.8	89.9
			80.7	75.0	91.9	91.9	81.4	98.5	97.7	99.1	99.1	92.1	80.4	98.8

The column “S” refers to settings of error types (N, normal; L, lognormal; M, mixed). Moreover, BT, NB(I), NB(II) and P refer to Box-type approximation, nonparametric bootstrap and permutation approach, respectively

Table 3 Empirical powers (as percentages) of the nonparametric bootstrap and permutation tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ proposed in Sect. 2 and of the BT test based on $\mathcal{C}_n$ by Smaga (2017) obtained in Model M3

S	n	ρ	$I = 26$					$I = 101$								
			C_n		$\mathcal{D}_n$		$\mathcal{E}_n$		C_n		$\mathcal{D}_n$		$\mathcal{E}_n$			
			BT	NB(I)	NB(II)	P	P	BT	NB(I)	NB(II)	P	P	NB(I)	NB(II)	P	
N	15	0.00	40.1	34.0	42.1	52.6	52.5	65.5	71.6	40.7	37.3	45.7	57.5	49.4	67.7	76.8
		0.25	55.6	52.1	61.9	71.7	71.3	82.8	86.2	59.4	55.9	64.9	75.7	68.7	83.6	89.8
		0.50	82.4	79.8	89.3	93.2	90.2	96.6	98.0	84.9	82.0	89.9	93.3	90.1	97.1	98.7
		0.75	99.5	99.3	100	99.9	99.7	100	100	99.8	99.7	100	100	99.8	100	100
	25	0.00	68.0	80.2	83.9	87.9	96.3	97.3	97.9	71.7	82.1	84.4	88.6	97.5	98.1	98.5
		0.25	88.5	94.0	96.1	97.5	99.4	99.6	99.6	87.6	94.9	96.6	97.5	99.5	99.7	99.7
		0.50	98.9	99.6	99.9	99.9	99.9	100	100	99.4	99.9	100	100	100	100	100
		0.75	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	35	0.00	93.1	97.8	98.2	99.1	99.9	99.8	99.9	91.9	97.6	98.5	98.9	100	100	100
		0.25	99.0	99.9	100	100	100	100	100	98.8	100	100	100	100	100	100
		0.50	100	100	100	100	100	100	100	100	100	100	100	100	100	100
		0.75	100	100	100	100	100	100	100	100	100	100	100	100	100	100
L	15	0.00	86.3	72.6	80.5	87.8	66.7	79.2	87.0	89.5	76.0	83.1	90.3	59.9	79.8	88.3
		0.25	95.4	88.5	92.6	95.7	82.3	91.7	94.6	96.6	89.8	93.8	96.3	77.9	91.5	95.8
		0.50	99.4	97.2	99.4	99.7	96.1	99.1	99.6	99.8	98.3	99.3	99.7	93.0	98.4	99.5
		0.75	100	100	100	100	100	100	100	100	100	100	100	99.9	100	100
	25	0.00	99.5	98.8	99.1	99.8	99.3	99.9	100	99.8	99.2	99.7	99.6	98.6	99.1	99.6
		0.25	100	100	100	100	100	100	100	100	99.9	100	100	99.7	100	100
		0.50	100	100	100	100	100	100	100	100	100	100	100	100	100	100
		0.75	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table 3 continued

S	n	ρ	$I = 26$						$I = 101$					
			$\mathcal{D}_n$			$\mathcal{E}_n$			$\mathcal{C}_n$			$\mathcal{D}_n$		
			BT	NB(I)	P	NB(II)	P	NB(I)	BT	NB(I)	P	NB(II)	P	NB(II)
35		0.00	100	100	100	100	100	100	100	100	100	100	100	100
		0.25	100	100	100	100	100	100	100	100	100	100	100	100
		0.50	100	100	100	100	100	100	100	100	100	100	100	100
		0.75	100	100	100	100	100	100	100	100	100	100	100	100
		0.00	71.9	58.2	66.5	77.8	78.9	56.0	74.3	59.9	68.4	78.8	48.8	80.0
M	15	0.25	81.9	68.7	77.8	85.3	85.3	69.0	84.3	69.3	78.0	86.1	59.5	88.9
		0.50	90.4	76.7	88.2	91.2	91.2	84.2	91.7	78.6	90.3	92.3	77.2	96.8
		0.75	96.7	85.2	96.3	95.6	95.6	95.9	97.3	86.3	97.3	97.0	95.5	99.6
		0.00	96.7	95.5	96.2	98.3	98.3	97.9	97.0	96.4	97.8	98.9	98.3	99.4
		0.25	99.0	98.5	99.0	99.4	99.4	99.6	99.6	99.2	99.8	100	99.8	100
	25	0.50	99.8	99.4	99.9	99.9	99.8	100	100	99.9	100	100	100	100
		0.75	100	99.8	100	100	100	100	100	100	100	100	100	100
		0.00	99.9	100	100	100	100	100	99.8	99.9	99.9	100	100	100
		0.25	100	100	100	100	100	100	100	100	100	100	100	100
		0.50	100	100	100	100	100	100	100	100	100	100	100	100
	35	0.75	100	100	100	100	100	100	100	100	100	100	100	100

The column “S” refers to settings of error types (N, normal; L, lognormal; M, mixed). Moreover, BT, NB(I), NB(II) and P refer to Box-type approximation, nonparametric bootstrap and permutation approach, respectively

4 Real data example

In this section, we present a real data example to illustrate the practical use of the tests proposed in Sect. 2 and of the tests of Martínez-Camblor and Corral (2011) and Smaga (2017). For this purpose, we consider the human mortality data, which are often used to quantify health state of people living in different countries and to appraise biological limits of longevity (Oeppen and Vaupel 2002; Vaupel et al. 1998). The data were obtained from the Human Mortality Database (Wilmoth et al. 2017; <http://www.mortality.org>).

For illustrative purposes, we take into account the mortality rates for two decades 1980–1989 and 1990–1999 for the following 32 countries: Australia, Austria, Belarus, Belgium, Bulgaria, Canada, Czech Republic, Denmark, Estonia, Finland, France, Hungary, Iceland, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Russia, Slovakia, Spain, Sweden, Switzerland, United Kingdom, Ukraine, USA. We focus on the mortality rates of older individuals, i.e., in age from 60 to 100 years. The death rates for these countries were also studied by using the principal component analysis for repeated functional observations in Chen and Müller (2012), but for longer time period.

For a given decade, the observed period mortality rates can be treated as realizations of some random process being a function of age ($t \in [60, 100]$). Moreover, the observations of the mortality rates for two decades can be considered as two samples of repeated functional data, since they are measured repeatedly for different countries, which are the subjects in this example. Thus, we have $n = 32$ repeated functional observations measured at $I = 41$ design time points in each of two conditions (decades). These data as well as the corresponding sample mean functions are shown in the first row of Fig. 1.

We are interested in testing the equality of the mean curves of mortality rates in two decades, which is a repeated measures analysis for functional data. To solve it, we apply the tests proposed in Sect. 2 and those by Martínez-Camblor and Corral (2011) and Smaga (2017). The obtained p -values are presented in Table 4 ($n = 32$). All new testing procedures reject the null hypothesis and their p -values are almost all equal to zero. On the other hand, the p -values of the tests based on $\mathcal{C}_n$ are slightly greater than standard significance level $\alpha = 5\%$, so these tests do not reject the null hypothesis and their decision is not sure. Since the tests based on $\mathcal{D}_n$ and $\mathcal{E}_n$ are strongly against the null hypothesis and the tests based on $\mathcal{C}_n$ are on boundary of rejecting and accepting this hypothesis, the mean functions of mortality rates for two decades 1980–1989 and 1990–1999 seem to be unlikely the same.

The sample mean functions for mortality data indicate that the death rates decreased in decade 1990–1999 compared to decade 1980–1989 for most ages. However, for ages about 100 years, the reverse seems to be true. This may be due to the fact that for five countries, namely Belarus, Bulgaria, Lithuania, Russia and Ukraine, the mortality rates increased significantly over time, especially for people in age from 90 to 100 years (see the second row of Fig. 1). On the other hand, the other countries share a declining trend of death rates with increasing decades (see the third row of Fig. 1). This was also observed in analysis presented in Chen and Müller (2012). Moreover, most of the tests based on the test statistics $\mathcal{C}_n$, $\mathcal{D}_n$ and

Fig. 1 The curves of mortality rates for two decades 1980–1989 and 1990–1999 and the all $n = 32$ countries considered (the first row), the $n = 5$ Eastern European countries (the second row), the $n = 27$ other countries (the third row) and the sample mean functions (solid line for decade 1980–1989; dashed line for decade 1990–1999)

$\mathcal{E}_n$ applied to these two groups of countries (the $n = 5$ Eastern European countries and the $n = 27$ other countries) reject the null hypothesis and indicate the significance of the above observations (Table 4). Note, however, that for the Eastern European countries, the tests based on the nonparametric bootstrap method (I) and on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ fail to reject the null hypothesis, which is an effect

Table 4 *P*-values of the tests based on the test statistics $\mathcal{D}_n$ and $\mathcal{E}_n$ proposed in Sect. 2 and of the tests based on $\mathcal{C}_n$ by Martínez-Camblor and Corral (2011) and Smaga (2017) obtained for the mortality data

<i>n</i>	$\mathcal{C}_n$			$\mathcal{D}_n$			$\mathcal{E}_n$		
	BT	PB	NB(I)	P	BT	PB	NB(I)	NB(II)	P
32	0.065	0.061	0.076	0.052	0.000	0.000	0.001	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.120	0.009	0.000
27	0.001	0.001	0.005	0.000	0.000	0.000	0.000	0.000	0.000

of small sample size ($n = 5$) and conservativity of these tests indicated in simulations of Sect. 3. On the other hand, the tests based on the nonparametric bootstrap method (II) have much smaller p -values and reject the null hypothesis, which confirms their less conservative character than that of the first nonparametric bootstrap approach.

5 Conclusions

In this paper, we have studied the testing procedures for the repeated measures analysis for functional data, which are based on test statistics taking into account both “between and within group variabilities”, in contrast to tests by Martínez-Camblor and Corral (2011) and Smaga (2017). By simulation studies and real data example, different proposed methods of approximating the null distribution of the test statistics do not work equally well under finite samples. The Box-type approximation and parametric bootstrap are too liberal when the number of observations is small, and therefore they are not recommended in such situation. On the other hand, the nonparametric bootstrap method may result in conservative tests and the loss of power may be noted. In terms of size control and power, the permutation test based on the test statistic being the supremum of the pointwise test statistic seems to be the best testing procedure, which also outperforms the tests of Martínez-Camblor and Corral (2011) and Smaga (2017). We also observed that the procedures proposed in these papers may be too anticonservative for very small sample size ($n \leq 15$).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

In this Appendix, we present and discuss the technical assumptions used in Theorem 1 and give its proof.

Let $L^2([a, b])$ (resp. $C^\beta([a, b])$) denote the set of all square-integrable functions (resp. Hölder continuous functions with exponent $\beta \in (0, 1]$) over $[a, b]$, $a < b$, and let $\text{tr}(\gamma) = \int_a^b \gamma(t, t) dt$ denote the trace of the covariance function $\gamma(s, t)$, $s, t \in [a, b]$. The following assumptions used in Theorem 1 are quite common in hypothesis testing problems for functional data.

Assumptions:

- A1. The mean function $m(t) \in L^2([0, 2])$ and $\text{tr}(\mathbb{C}) < \infty$.
- A2. The subject-effect function $v_1(t) = X_1(t) - m(t)$ satisfies

$$\mathbb{E}\|v_1\|^4 = \mathbb{E}\left(\int_0^2 v_1^2(t) dt\right)^2 < \infty.$$

- A3. For any $t \in [0, 2]$, $\mathbb{C}(t, t) > 0$, and $\max_{t \in [0, 2]} \mathbb{C}(t, t) < \infty$.
 A4. For any $(s, t) \in [0, 2]^2$, $E(v_1^2(s)v_1^2(t)) < C < \infty$, where C is certain constant independent of any $(s, t) \in [0, 2]^2$.
 A5. For any $t \in [0, 1]$, $\mathbb{K}(t, t) > 0$.
 A6. The covariance function $\mathbb{K}(s, t) \in C^{2\beta}([0, 1]^2)$, $2\beta \in (0, 1]$.

The regular Assumption A1 guarantees that the sample mean function $\bar{X}(t)$, $t \in [0, 2]$ will converge to certain Gaussian process weakly, as $n \rightarrow \infty$. The other assumptions are imposed to obtain well-defined pointwise test statistic (6) and to guarantee that the sample covariance functions $\hat{\mathbb{C}}(s, t)$ and $\hat{\mathbb{K}}(s, t)$ converge to the covariance functions $\mathbb{C}(s, t)$ and $\mathbb{K}(s, t)$ uniformly over $[0, 2] \times [0, 2]$ and $[0, 1] \times [0, 1]$, respectively, as $n \rightarrow \infty$. Let us devote a little more attention to comment on Assumption A5. For any fixed $t \in [0, 1]$, we have

$$\begin{aligned} \mathbb{K}(t, t) &= \mathbb{C}(t, t) - 2\mathbb{C}(t, t+1) + \mathbb{C}(t+1, t+1) \\ &\geq \mathbb{C}(t, t) - 2\sqrt{\mathbb{C}(t, t)\mathbb{C}(t+1, t+1)} + \mathbb{C}(t+1, t+1) \end{aligned} \quad (10)$$

$$\begin{aligned} &= \left(\sqrt{\mathbb{C}(t, t)} - \sqrt{\mathbb{C}(t+1, t+1)} \right)^2 \\ &\geq 0. \end{aligned} \quad (11)$$

In (10) (resp. (11)), the equality holds if and only if the random variables are linearly dependent (resp. $\text{Var}(X(t)) = \text{Var}(X(t+1))$). This indicates that Assumption A5 seems to be usually satisfied.

Proof of Theorem 1 By Lemma 2.1 in Smaga (2017), $\hat{\mathbb{K}}(t, t) \xrightarrow{P} \mathbb{K}(t, t)$ uniformly over $[0, 1]$, as $n \rightarrow \infty$, where $\xrightarrow{P}$ denotes convergence in probability. From the proof of Theorem 1 in Martínez-Camblor and Corral (2011), it follows that $n^{1/2}(\bar{X}(t) - \bar{X}(t+1)) \xrightarrow{d} z(t) - z(t+1)$ for all $t \in [0, 1]$, as $n \rightarrow \infty$. Therefore, we have

$$\frac{\sqrt{n}(\bar{X}(t) - \bar{X}(t+1))}{\sqrt{\hat{\mathbb{K}}(t, t)}} \xrightarrow{d} \frac{z(t) - z(t+1)}{\sqrt{\mathbb{K}(t, t)}},$$

as $n \rightarrow \infty$ (Slutsky's theorem). Hence, by the continuous mapping theorem, we conclude that $\mathcal{D}_n \xrightarrow{d} \int_0^1 (z(t) - z(t+1))^2 / \mathbb{K}(t, t) dt$, as $n \rightarrow \infty$. Since $(z(t) - z(t+1)) / (\mathbb{K}(t, t))^{1/2}$ is a Gaussian process with zero mean and covariance function $\mathbb{K}_*(s, t)$, $s, t \in [0, 1]$ and $\text{tr}(\mathbb{K}_*) = 1$, Theorem 4.2 in Zhang (2013) implies

$$\int_0^1 \frac{(z(t) - z(t+1))^2}{\mathbb{K}(t, t)} dt = \left\| \frac{z(t) - z(t+1)}{\sqrt{\mathbb{K}(t, t)}} \right\|^2 \stackrel{d}{=} \sum_{k=1}^{\infty} \lambda_k^* A_k,$$

which completes the proof of (7). To prove (8), we observe that

$$\frac{n(\bar{X}(t) - \bar{X}(t+1))^2}{\hat{\mathbb{K}}(t, t)} \xrightarrow{P} \frac{(z(t) - z(t+1))^2}{\mathbb{K}(t, t)},$$

for all $t \in [0, 1]$, as $n \rightarrow \infty$, and we continue the proof in much the same way as in the proof of Proposition 2.1 in Zhang et al. (2018). $\square$

References

- Box, G.E.P.: Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Ann. Math. Stat.* **25**, 290–302 (1954)
- Chen, K., Müller, H.G.: Modeling repeated functional observations. *J. Am. Stat. Assoc.* **107**, 1599–1609 (2012)
- Collazos, J.A.A., Dias, R., Zambom, A.Z.: Consistent variable selection for functional regression models. *J. Multivariate Anal.* **146**, 63–71 (2016)
- Cuevas, A., Febrero, M., Fraiman, R.: An ANOVA test for functional data. *Comput. Stat. Data Anal.* **47**, 111–122 (2004)
- Delicado, P.: Functional k -sample problem when data are density functions. *Comput. Stat.* **22**, 391–410 (2007)
- Ferraty, F., Vieu, P.: *Nonparametric Functional Data Analysis: Theory and Practice*. Springer, New York (2006)
- Gabrys, R., Kokoszka, P.: Portmanteau test of independence for functional observations. *J. Am. Stat. Assoc.* **102**, 1338–1348 (2007)
- Giraldo, R., Caballero, W., Camacho-Tamayo, J.: Mantel test for spatial functional data. *AStA Adv. Stat. Anal.* **102**, 21–39 (2018)
- Górecki, T., Smaga, Ł.: A comparison of tests for the one-way ANOVA problem for functional data. *Comput. Stat.* **30**, 987–1010 (2015)
- Górecki, T., Smaga, Ł.: Multivariate analysis of variance for functional data. *J. Appl. Stat.* **44**, 2172–2189 (2017)
- Horváth, L., Kokoszka, P.: *Inference for Functional Data with Applications*. Springer, New York (2012)
- Imhof, J.P.: Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419–426 (1961)
- Kokoszka, P., Maslova, I., Sojka, J., Zhu, L.: Testing for lack of dependence in the functional linear model. *Can. J. Stat.* **36**, 207–222 (2008)
- Konietschke, F., Pauly, M.: Bootstrapping and permuting paired t -test type statistics. *Stat. Comput.* **24**, 283–296 (2014)
- Kosiorowski, D., Rydlewski, J.P., Snarska, M.: Detecting a structural change in functional time series using local Wilcoxon statistic. *Stat. Papers* (2017). <https://doi.org/10.1007/s00362-017-0891-y>
- Laha, R.G., Rohatgi, V.K.: *Probability Theory*. Wiley, New York (1979)
- Martínez-Camblor, P., Corral, N.: Repeated measures analysis for functional data. *Comput. Stat. Data Anal.* **55**, 3244–3256 (2011)
- Martínez-Camblor, P., Corral, N.: A general bootstrap algorithm for hypothesis testing. *J. Stat. Plann. Inference* **142**, 589–600 (2012)
- Maslova, I., Kokoszka, P., Sojka, J., Zhu, L.: Statistical significance testing for the association of magnetometer records at high-, mid- and low latitudes during substorm days. *Planet. Space Sci.* **58**, 437–445 (2010)
- McLean, M.W., Hooker, G., Ruppert, D.: Restricted likelihood ratio tests for linearity in scalar-on-function regression. *Stat. Comput.* **25**, 997–1008 (2015)
- Oeppen, J., Vaupel, J.W.: Broken limits to life expectancy. *Science* **296**, 1029–1031 (2002)
- Pini, A., Vantini, S.: Interval-wise testing for functional data. *J. Nonparametric Stat.* **29**, 407–424 (2017)
- Ramsay, J.O., Silverman, B.W.: *Functional Data Analysis*. Springer, New York (1997)
- Ramsay, J.O., Silverman, B.W.: *Applied Functional Data Analysis. Methods and Case Studies*. Springer, New York (2002)
- Ramsay, J.O., Silverman, B.W.: *Functional Data Analysis*, 2nd edn. Springer, New York (2005)
- R Core Team: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (2017). Accessed 11 Apr 2018
- Smaga, Ł.: Repeated measures analysis for functional data using Box-type approximation—with applications. *REVSTAT*. <https://www.ine.pt/revstat/pdf/REPEATEDMEASURESANALYSISFORFUNC.pdf> (2017)

- Vaupel, J.W., Carey, J.R., Christensen, K., Johnson, T.E., Yashin, A.I., Holm, N.V., Iachine, I.A., Kannisto, V., Khazaeli, A.A., Liedo, P., Longo, V.D., Zeng, Y., Manton, K.G., Curtsinger, J.W.: Biodemographic trajectories of longevity. *Science* **280**, 855–860 (1998)
- Wilmoth, J.R., Andreev, K., Jdanov, D., Glei, D.A.: Methods protocol for the human mortality database. <http://www.mortality.org/Public/Docs/MethodsProtocol.pdf> (2017). Accessed 7 Mar 2018
- Zhang, J.T.: *Analysis of Variance for Functional Data*. Chapman & Hall, London (2013)
- Zhang, J.T., Cheng, M.Y., Wu, H.T., Zhou, B.: A new test for functional one-way ANOVA with applications to ischemic heart screening. *Comput. Stat. Data Anal.* (2018). <https://doi.org/10.1016/j.csda.2018.05.004>
- Zhang, J.T., Liang, X.: One-way ANOVA for functional data via globalizing the pointwise F -test. *Scand. J. Stat.* **41**, 51–71 (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.