Chapter 19
Two-sample tests for multivariate functional data

Qing Jiang, Simos G. Meintanis and Lixing Zhu

Abstract We consider two–sample tests for functional data with observations which may be uni– or multi–dimensional. The new methods are formulated as L2–type criteria based on empirical characteristic functions and are convenient from the computational point of view.
Keywords: Functional data, Empirical characteristic function, Two–sample problem

19.1 Introduction

Suppose that we observe data X_{1ij} and X_{2ij} arising from two different groups. For each fixed i, X_{1ij}, is viewed as realization of a curve $x_{1i}(t)$ observed at distinct time points t_{1ij}, $j = 1,...,m_{1i}$, and we index the curves by $i = 1,...,n_1$, for the first group. Likewise suppose that X_{2ij} is realization of a curve $x_{2i}(t)$, observed at times t_{2ij}, $j = 1,...,m_{2i}$, and indexed by $i = 1,...,n_2$, for the second group. The observation times t_{1ij}, t_{2ij} are assumed to belong to some closed bounded interval $\mathbb{T}$, and we often take $\mathbb{T} = [0,1]$. Although we work under the assumption of independence between groups, we allow for noise in the observations. Specifically we consider the model

$$X_{1ij} = x_{1i}(t_{1ij}) + \varepsilon_{1ij}, \;\; X_{2ij} = x_{2i}(t_{2ij}) + \varepsilon_{2ij}, \tag{19.1}$$

Qing Jiang
School of Statistics, Beijing Normal University e-mail: 201631011006@mail.bnu.edu.cn

Simos G. Meintanis (✉)
Department of Economics, National and Kapodistrian University of Athens, Athens, Greece and Unit for Business Mathematics and Informatics, North–West University, Potchefstroom, South Africa e-mail: simosmei@econ.uoa.gr

Lixing Zhu
School of Statistics, Beijing Normal University and Department of Mathematics, Hong Kong Baptist University, Hong Kong e-mail: lzhu@hkbu.edu.hk

G. Aneiros et al. (eds.), *Functional Statistics and Related Fields*,
Contributions to Statistics, DOI 10.1007/978-3-319-55846-2_19

where $x_{11}(t),...,x_{1n_1}(t)$, are independent and identically distributed as $x_1(t)$, and independent of the the errors $\varepsilon_1(t)$, for $t \in \mathbb{T}$, and likewise $x_{21}(t),...,x_{2n_2}(t)$ are iid as $x_2(t)$, and independent of $\varepsilon_2(t)$. The errors are also assumed to be mutually independent with zero means. We wish to test the null hypothesis

$$\mathscr{H}_0 : x_1(t) \stackrel{d}{=} x_2(t), \text{ for each } t \in \mathbb{T}, \tag{19.2}$$

where $\stackrel{d}{=}$ stands for equality in law.

Earlier works for the two–sample problem with functional data include testing for common location ([6, 5, 16]), and for common covariance matrix ([11, 8]), while [1] considers the more general problem of testing for common principal components. The framework of the current paper though is much in the spirit of [4] where the two–sample problem was first studied in its full generality of the null hypothesis (19.2); see also [12]. However we deviate from this paper by proposing procedures which instead of the empirical distribution function, utilize the empirical characteristic function (ECF). Apart from other favorable features which will become apparent along the paper, note that ECF–based procedures for scalar data are readily extended to multidimensional observations which is not always true if one employs classical procedures based on the empirical distribution function.

19.2 Test Statistics

19.2.1 Univariate case

Our approach for testing the null hypothesis $\mathscr{H}_0$ in (19.2) will be based on the fact that $\mathscr{H}_0$ is tantamount to the identity

$$\varphi_{x_1(t)}(u) = \varphi_{x_2(t)}(u), \ \forall \ u \in \mathbb{R}, \text{ and each } t \in \mathbb{T}, \tag{19.3}$$

and vice versa. Here, as well as elsewhere below, $\varphi_{z(t)}(u) := \mathbb{E}(e^{\imath u z(t)}), \ (\imath = \sqrt{-1})$, will denote the characteristic function (CF) of the stochastic quantity $z(t)$. Based on this fact, [10] and [9] develop two–sample testing procedures for multivariate data. Here we follow this approach and in line with [4], we assume that the curves $x_{1i}(t)$ and $x_{2i}(t)$ may be recovered following non–parametric techniques and write $\widehat{x}_{1i}(t)$ and $\widehat{x}_{2i}(t)$ for the resulting curve estimators. Consider the corresponding ECFs

$$\widehat{\varphi}_{1t}(u) = \frac{1}{n_1}\sum_{i=1}^{n_1} e^{\imath u \widehat{x}_{1i}(t)}, \quad \widehat{\varphi}_{2t}(u) = \frac{1}{n_2}\sum_{i=1}^{n_2} e^{\imath u \widehat{x}_{2i}(t)}, \tag{19.4}$$

computed from $\widehat{x}_{11}(t),...,\widehat{x}_{1n_1}(t)$ and $\widehat{x}_{21}(t),...,\widehat{x}_{2n_2}(t)$, respectively. Then in view of (19.3) we suggest the test statistic

$$D_w = \int_{\mathbb{T}} \int_{\mathbb{R}} \delta_t(u) w(u) du dt, \tag{19.5}$$

where

$$\delta_t(u) = |\widehat{\varphi}_{1t}(u) - \widehat{\varphi}_{2t}(u)|^2, \tag{19.6}$$

and $w > 0$ denotes a weight function satisfying $\int_{\mathbb{R}} w(u)du < \infty$.

19.2.2 Multivariate case

The latent curves $x_k(t) = (\chi_{k1}(t), ..., \chi_{kp}(t))'$, $k = 1, 2$, may also be multidimensional. This is a new area where functional data are observed over time t, but realizations are complex geometrical structures in dimension $p > 1$; see [3], [7], and [2], for recent contributions on statistical techniques for multivariate functional data. Following the lines of the previous section for testing the null hypothesis (19.2) we will consider a criterion analogous to that in (19.5). However, in order to avoid nonparametric estimation which is problematic in high dimension we will have to modify our assumptions regarding model (19.1). Specifically we adopt the model

$$X_{1i}(t) = x_{1i}(t) + \varepsilon_{1i}(t), \;\; X_{2i}(t) = x_{2i}(t) + \varepsilon_{2i}(t), \tag{19.7}$$

and we assume that observations are collected over time $t := t_j$, $j = 1, ..., m$, for both groups, with m being large, i.e., we have a common sampling design between the two groups which is dense. Moreover we assume that sampling noise is equidistributed between the two groups, $\varepsilon_1(t) \stackrel{d}{=} \varepsilon_2(t)$, with a common CF that never vanishes. Under these assumptions and using the Fourier identities $\varphi_{X_k(t)}(u) = \varphi_{x_k(t)}(u)\varphi_{\varepsilon_k(t)}(u)$, $k = 1, 2$, resulting from model (19.7), we conclude that the null hypothesis $\mathscr{H}_0$ in (19.2) holds if and only if

$$\varphi_{X_1(t)}(u) = \varphi_{X_2(t)}(u), \; \forall \, u \in \mathbb{R}^p, \text{ and each } t \in \mathbb{T}. \tag{19.8}$$

In view of this fact we propose the test statistic

$$\Delta_W = \int_{\mathbb{R}^p} \delta(u) W(u) du, \tag{19.9}$$

with $W : \mathbb{R}^p \mapsto (0, \infty)$ and satisfying $\int_{\mathbb{R}^p} W(u)du < \infty$, where

$$\delta(u) = \frac{1}{m} \sum_{j=1}^{m} \left|\phi_{1j}(u) - \phi_{2j}(u)\right|^2, \tag{19.10}$$

with

$$\phi_{1j}(u) = \frac{1}{n_1} \sum_{i=1}^{n_1} e^{\imath u' X_{1ij}}, \quad \phi_{2j}(u) = \frac{1}{n_2} \sum_{i=1}^{n_2} e^{\imath u' X_{2ij}}, \tag{19.11}$$

being the ECFs computed directly from the observed data $X_{k1j},...,X_{kn_kj}$, and which correspond to the CFs $\varphi_{X_1}(u)$ and $\varphi_{X_2}(u)$, respectively, considered at fixed time points t_j, for each $j = 1,...,m$.

19.3 Computations and Interpretations

19.3.1 Univariate case

Our procedures enjoy the advantage of computational simplicity. To see this we first proceed from (19.6) and by using simple algebra and trigonometric identities we get

$$\delta_t(u) = \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_1}\cos(u(\widehat{x}_{1i}(t)-\widehat{x}_{1\ell}(t))) + \frac{1}{n_2^2}\sum_{i,\ell=1}^{n_2}\cos(u(\widehat{x}_{2i}(t)-\widehat{x}_{2\ell}(t))) - \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{\ell=1}^{n_2}\cos(u(\widehat{x}_{1i}(t)-\widehat{x}_{2\ell}(t))) \quad (19.12)$$

Then by making use of the previous equation in (19.5) we conclude that the test statistic can be written as

$$D_w = \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_1} I_{w,\mathbb{T}}(\widehat{x}_{1i},\widehat{x}_{1\ell}) + \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_2} I_{w,\mathbb{T}}(\widehat{x}_{2i},\widehat{x}_{2\ell}) - \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{\ell=1}^{n_2} I_{w,\mathbb{T}}(\widehat{x}_{1i},\widehat{x}_{2\ell}) \quad (19.13)$$

where

$$I_{w,\mathbb{T}}(z_1,z_2) = \int_{\mathbb{T}}\int_{\mathbb{R}} \cos(u(z_1(t)-z_2(t)))w(u)dudt. \quad (19.14)$$

The weight function $w(\cdot)$ in (19.14) may be chosen in a way that avoids numerical integration in the inner integral $\int\cos(u(z))w(u)du$ but for further details on this we refer to the next subsection. Then again having computed $\int\cos(u(z))w(u)du := g(z(t))$, say, one also has to compute the outer integral $\int g(z(t))dt$, over $\mathbb{T}$. However even in the simplest case of local linear smoothers[4], the closed form obtained for $\widehat{x}_{ki}(t)$ is quite complicated and therefore one needs to resort to numerical integration. Despite this, integration in closed bounded domains is a well studied numerical problem and there exist several routines available for this purpose. Hence we do not expect any complications to be associated with this part of our procedure. For simplicity we take $\mathbb{T} = [0,1]$

The choice for the weight function $w(\cdot)$ is usually based upon computational considerations. In fact if $w(\cdot)$ integrates to one (perhaps after some scaling) and satisfies $w(-u) = w(u)$ then the inner integral in (19.14) can be interpreted as the CF of a symmetric around zero random variable having density $w(\cdot)$. In this connection $w(\cdot)$ can be chosen as the density of any such distribution. Typically we consider a fixed family of weight functions, say $w := w_\gamma$ indexed by a parameter $\gamma > 0$. For instance a weight function $w_\gamma(u)$ which is proportional to $e^{-\gamma u^2}$, corresponds to the

Gaussian density, but for computational purposes any density with a simple CF will do. In fact, one might wonder whether there is a weight function which is optimal in some sense. The issue is still open but based on earlier results it appears that the issue of the choice of the weight function is similar to the corresponding problem of choosing a kernel and a bandwidth in nonparametric estimation: the asymptotics of the test are qualitatively invariant with respect to w_γ. Moreover most weight functions (kernels) render similar finite–sample behavior of the resulting test statistic, which is very competitive compared to classical procedures based on the empirical distribution function. Nevertheless there is some sensitivity of the ECF–tests with respect to the "bandwidth" parameter γ; see [10] and [9]. This is a highly technical problem that has been tackled only under the restrictive scenario of testing goodness–of–fit for a given parametric distribution, and even then a good choice of γ depends on the direction away from the null hypothesis; see [15]. Thus in our context the approach to the weight function is in some sense pragmatic: we use the Gaussian weight function which has become a standard, and investigate the behavior of the criterion over a grid of values of the weight parameter γ. However in our Monte Carlo study, alternative weight functions will also be tried.

19.3.2 Multivariate case

We proceed from (19.10) and by using (19.12) we obtain

$$\delta(u) = \frac{1}{m}\sum_{j=1}^{m}\delta_j(u), \tag{19.15}$$

where

$$\begin{aligned}\delta_j(u) = {} & \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_1}\cos(u'(X_{1ij}-X_{1\ell j})) + \frac{1}{n_2^2}\sum_{i,\ell=1}^{n_2}\cos(u'(X_{2ij}-X_{2\ell j})) \\ & -\frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{\ell=1}^{n_2}\cos(u'(X_{1ij}-X_{2\ell j}))\end{aligned} \tag{19.16}$$

Consequently the test statistic can be written as

$$\begin{aligned}\Delta_W = {} & \frac{1}{m}\sum_{j=1}^{m}\Big(\frac{1}{n_1^2}\sum_{i,\ell=1}^{n_1}I_W(X_{1ij}-X_{1\ell j}) + \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_2}I_W(X_{2ij}-X_{2\ell j}) \\ & -\frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{\ell=1}^{n_2}I_W(X_{1ij}-X_{2\ell j})\Big)\end{aligned} \tag{19.17}$$

where

$$I_W(x) = \int_{\mathbb{R}^p}\cos(u'x)W(u)du. \tag{19.18}$$

As already mentioned, the weight function $W(\cdot)$ in (19.18) may be chosen in a way that avoids numerical integration, which is problematic in higher dimension. To see this recall that the CF of any spherical random variable Z is given by $\varphi_Z(u) = \Psi(\|u\|)$, for some, family specific, scalar function $\Psi(\cdot)$, where $\|u\|$ denotes the usual Euclidean norm. Hence if $f_Z(z)$ denotes the density corresponding to $\varphi_Z(u)$ we have

$$\int_{\mathbb{R}^p} \cos(u'z) f_Z(z) dz = \Psi(\|u\|).$$

The last equation implies that if $f_Z(\cdot)$ is used as weight function $W(\cdot)$ in (19.18), then the resulting test statistic, say Δ_Ψ, reduces to

$$\Delta_\Psi = \frac{1}{m} \sum_{j=1}^{m} \Big(\frac{1}{n_1^2} \sum_{i,\ell=1}^{n_1} \Psi(\|X_{1ij} - X_{1\ell j}\|) + \frac{1}{n_1^2} \sum_{i,\ell=1}^{n_2} \Psi(\|X_{2ij} - X_{2\ell j}\|) - \frac{2}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{\ell=1}^{n_2} \Psi(\|X_{1ij} - X_{2\ell j}\|) \Big). \tag{19.19}$$

The test criterion in (19.19) is further advanced by considering specific families of spherically symmetric distributions with a simple CF. Such a family of distributions is the family of spherical stable distributions with $\Psi(u) = e^{-u^\alpha}$, where $0 < \alpha \le 2$, stands for the characteristic exponent. Interesting special cases of spherical stable distributions are the multivariate Cauchy and normal distributions corresponding to $\alpha = 1$ and $\alpha = 2$, respectively. Other convenient choices are the multivariate Laplace distribution with $\Psi(u) = (1+u^2)^{-1}$ and some special cases of the family of multivariate Kotz–type distributions.

We will elaborate here on the case of the spherical stable distribution as weight function in (19.18). Note that if this function is used in (19.19), it yields the test criterion

$$\Delta_\alpha = \frac{1}{m} \sum_{j=1}^{m} \Big(\frac{1}{n_1^2} \sum_{i,\ell=1}^{n_1} e^{-\|X_{1ij} - X_{1\ell j}\|^\alpha} + \frac{1}{n_2^2} \sum_{i,\ell=1}^{n_2} e^{-\|X_{2ij} - X_{2\ell j}\|^\alpha} - \frac{2}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{\ell=1}^{n_2} e^{-\|X_{1ij} - X_{2\ell j}\|^\alpha} \Big). \tag{19.20}$$

Interestingly there is a connection between (19.20) and another two–sample test statistic in the literature. To see this let Z follow a spherical stable distribution with characteristic exponent α and choose the density of the random variable $Z/\gamma^{1/\alpha}$ as weight function in (19.17)–(19.18), for some $\gamma > 0$. To get a formula for the resulting criterion recall that the CF of the last random variable is given by $e^{-\frac{\|u\|^\alpha}{\gamma}}$ which yields a test statistic, say $\widetilde{\Delta}_{\alpha,\gamma}$, analogous to the criterion in (19.20) but with $\|\cdot\|^\alpha$ being replace by $\|\cdot\|^\alpha/\gamma$ throughout eqn. (19.20). Now if we take a two–term expansion $e^{-\|x\|^\alpha/\gamma} = 1 - \|x\|^\alpha/\gamma + o(\gamma^{-1})$, $\gamma \to \infty$, in the new test statistic $\widetilde{\Delta}_{\alpha,\gamma}$, this will lead after some algebra to

$$\lim_{\gamma\to\infty}\gamma\widetilde{\Delta}_{\alpha,\gamma} = \frac{1}{m}\sum_{j=1}^{m}\Big(\frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{\ell=1}^{n_2}\|X_{1ij}-X_{2\ell j}\|^{\alpha} - \frac{1}{n_1^2}\sum_{i,\ell=1}^{n_1}\|X_{1ij}-X_{1\ell j}\|^{\alpha}$$
$$-\frac{1}{n_2^2}\sum_{i,\ell=1}^{n_2}\|X_{2ij}-X_{2\ell j}\|^{\alpha}\Big). \tag{19.21}$$

The criterion in the r.h.s. of (19.21) is the so–called *energy statistic* of [13] adapted to the functional context. We mention in this connection that energy statistics have gained considerable popularity lately as they have been employed not just for two–sample testing but also for testing for independence as well as in nonparametric analysis of variance. The reader is referred to the review of [14] for more information on energy statistics.

19.4 Resampling procedures

The null distribution of the test statistics considered depends, among other things, on the underlying stochastic properties of the random variables $x_1(t)$ and $x_2(t)$ involved. In order to deal with this issue we apply appropriate resampling procedures for computing critical points and actually carrying out the tests. To this end, let $\mathscr{D} = \mathscr{D}(\xi_1,...,\xi_n)$ be a generic notation for a test statistic which depends on a sample of size n of observations $\xi_j,\ 1\le j\le n$. Clearly in our case $n = n_1+n_2$. We will apply the permutation procedure whereby we randomly generate a permutation $b=\{b_1,...,b_n\}$ of $\{1,...,n\}$, and compute the test statistic $\mathscr{D}_b = \mathscr{D}(\xi_{b_1},...,\xi_{b_n})$. The procedure is repeated a number of times $b = 1,...,\mathrm{B}$, and the critical point of the test of size α is determined as the corresponding $(1-\alpha)$ quantile $D_{((1-\alpha)\mathrm{B})}$ of the values $\mathscr{D}_b$, $b=1,...,\mathrm{B}$. The null hypothesis is then rejected if $\mathscr{D} > \mathscr{D}_{((1-\alpha)\mathrm{B})}$.

Suppose that data $X_{k1j},\dots,X_{kn_kj}$ are observed at fixed time points t_j, for each $j = 1,\dots,m$. For univariate data, the critical point of the test statistic in (19.5) is computed as in [4], i.e., by permuting $\{\widehat{x}_{11},...,\widehat{x}_{1n_1},\widehat{x}_{21},...,\widehat{x}_{2n_2}\}$. In turn with multivariate data, permutations for the criterion in (19.9) are performed on $\{X_{11j},\ ...,\ X_{1n_1j},\ X_{21j},\ ...,\ X_{2n_2j}\}$, for each $j=1,...,m$.

In the univariate case, we generate data $\{(t_{kij},x_{ki}(t_{kij})) : j\}_{i=1}^{n_k}, k=1,2$, mainly as in [4]. For completeness we describe the data as follows: the sampling design for the curves is assumed balanced $(m_{1i}=m_{2i}=m), \forall i$, and regular. Specifically, suppose that $t_{ki\cdot}, k=1,2, i=1,\dots,n_k$ are discrete uniform fixed time points on $[0,1]$. It is assumed that $x_{1i}(t)=\sum_{k=1}^{15} e^{-k/2}N_{k1i}\psi_k(t)$ and

$$x_{2i}(t)=\sum_{k=1}^{15} e^{-k/2}N_{k21i}\psi_k(t)+\delta\sum_{k=1}^{15}k^{-2}N_{k22i}\psi_k^{*}(t)$$

where $N_{k1i},N_{k21i},N_{k22i}$ are i.i.d. standard normal variables, $\delta\ge 0$ controls the deviation from the null hypothesis ($\delta=0$ under H_0). Here $\psi_1(t)\equiv 1,\forall t$ and $\psi_k(t)=\sqrt{2}\sin\{(k-1)\pi t\}$ are orthonormal basis functions. Also

$$\psi_k^*(t) = \begin{cases} 1 & \text{if } k = 1 \\ \sqrt{2}\sin\{(k-1)\pi(2t-1)\} & \text{if } k \text{ is odd and } k > 1 \\ \sqrt{2}\cos\{(k-1)\pi(2t-1)\} & \text{if } k \text{ is even} \end{cases}$$

are orhonormal basis functions. Two scenarios are considered: *i*) $m = 20$ points per curve, and *ii*) $m = 100$ points per curve. Figure 19.1 (without sampling noise) illustrates the ECF test results for significance level $a = 0.05$. The simulation results are based on 500 samples, and the critical values of the test are obtained from 1000 permutation samples.

Figure 19.1 illustrates that the level is well respected under the null hypothesis and the power increases for larger values of m, n and δ in two conditions. The number of observations per curve, m, has limited impact on the power and the conclusion is consistent with [4]. However, compared with their results, the empirical power of the ECF test increases at a faster rate than the CVM test of [4]. This should not be surprising, as we do not need to estimate all basis functions by smoothing the data when observations are without noise or sampling noise is equidistributed. When observations are without noise, we can directly estimate the test statistic Δ_α in (19.20). Also the Fourier identities makes it consistent to transform the null hypothesis (19.2) to equation (19.8) when sampling noise is equidistributed.

19.5 Conclusion

We suggest a new procedure for testing the two–sample null hypothesis with functional data. The procedure is an adaptation to the functional– data set up of earlier methods for the same problem with perfectly observed i.i.d. data. Here we present only the main ideas of the new methods and a small Monte Carlo study. A detailed study of the asymptotic as well as the finite–sample behavior of the methods is currently under investigation and will be reported elsewhere.

References

[1] Benko, M., Härdle, W., Kneip, A.: Common functional principal components. Ann. Stat. **37**, 1–34 (2009)
[2] Berrendero, J.R., Justel, A., Svarc, M.: Principal components for multivariate functional data. Comput. Stat. Data An. **55**, 2619–2634 (2011)
[3] Chiou, J.M., Müller, H.G.: Linear manifold modeling of multivariate functional data. J. Roy. Stat. Soc. B. **76**, 605–626 (2014)
[4] Hall, P., Van Keilegom, I.: Two–sample tests in functional data analysis starting from discrete data. Stat. Sinica. **17**, 1511–1531 (2007)
[5] Horváth, L., Kokoszka, P.: Inference for Functional Data with Applications. Springer Series in Statistics, Springer, New York (2012)

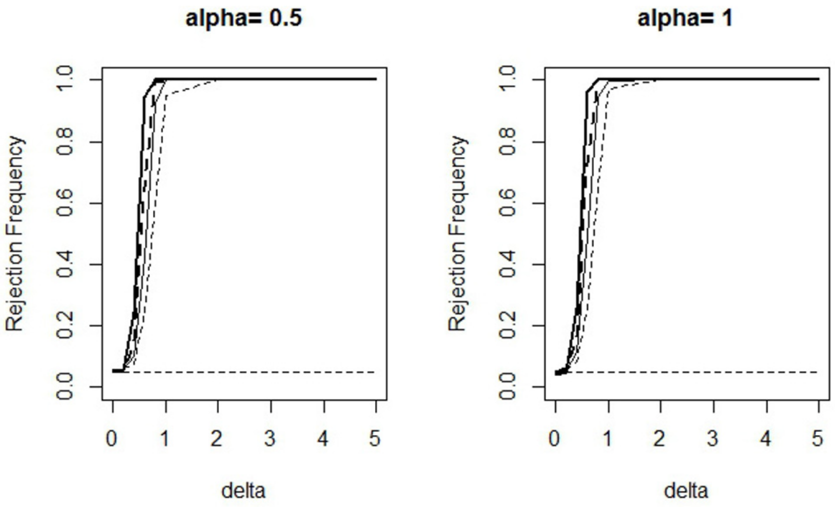

Fig. 19.1: Rejection frequencies of the ECF test for data without noise when $n = 15$ (dashed line) and $n = 25$ (solid line) corresponding to level of significance is $a = 5\%$ for the test statistic in (20) with characteristic exponent α. The thin lines correspond to $m = 20$, the thick lines to $m = 100$.

[6] Horváth, L., Kokoszka, P., Reeder, R.: Estimation of the mean of functional time series and a two sample problem. J. Roy. Stat. Soc. B. **75**, 103–122 (2013)
[7] Jacques, J., Preda, C.: Model–based clustering for multivariate functional data. Comput. Stat. Data An. **71**, 92–106 (2014)
[8] Kraus, D., Panaretos, V.M.: Dispersion operators and resistant second–order functional data analysis. Biometrika. **99**, 813–832 (2012)
[9] Hušková, M., Meintanis, S.G.: Tests for the multivariate k-sample problem based on the empirical characteristic function. J. NonParametr. Statist. **20**, 263–277 (2008)
[10] Meintanis, S.G.: Permutation tests for homogeneity based on the empirical characteristic function. J. Nonparametr. Statist. **17**, 583–592 (2005)
[11] Panaretos, V.M., Kraus, D., Maddocks, J.H.: Second–order comparison of Gaussian random functions and the geometry of DNA minicircles. J. Am. Stat. Assoc. **105**, 670–682 (2010)
[12] Pomann, G.M., Staicu, A.M., Ghosh, S.: Two–sample hypothesis testing for functional data with application to a diffusion tensor imaging study of multiple sclerosis. J. Roy. Stat. Soc. C-App. **65**, 395–414 (2016)
[13] Székely G., Rizzo, M.: Hierarchical clustering via joint between–within distances: Extending Ward's minimum variance method. J. Classif. **22**, 151–183 (2005)

[14] Székely G., Rizzo, M.: Energy statistics: A class of statistics based on distances. J. Stat. Plan. Infer. **143**, 249–272 (2013)
[15] Tenreiro, C.: On the choice of the smoothing parameter for the BHEP goodness-of-fit test. Computat. Statist. Dat. Anal. **53**, 1038–1053 (2009)
[16] Zhang, C., Peng, H., Zhang, J.T.: Two–sample tests for functional data. Commun. Stat.–Theor. M. **39**, 559–578 (2010)