Computational Statistics and Data Analysis 131 (2019) 194–206

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Weighted empirical likelihood inference for dynamical correlations

Peijun Sang, Liangliang Wang, Jiguo Cao *

Department of Statistics and Actuarial Science, Simon Fraser University, BC, Canada V5A 1S6

ARTICLE INFO

Article history:
Received 2 November 2017
Received in revised form 12 June 2018
Accepted 10 July 2018
Available online 18 July 2018

Keywords:
Bootstrap
Empirical likelihood
Functional data analysis

ABSTRACT

A novel approach is proposed based on the weighted empirical likelihood to construct confidence intervals for dynamical correlation of random functions. The properties of the proposed confidence interval are investigated for random functions with regular or irregular observations. It is shown that the confidence interval using our new approach has a more accurate coverage probability than that using the traditional bootstrap method for random functions with irregular observations. Furthermore, simulation studies demonstrate that the new approach is considerably more efficient in computation than the bootstrap method. The new approach is illustrated with three applications. The first application investigates the dynamical correlation of air pollutants. The second application studies the dynamical correlation of EEG signals in different regions of the brain in response to some stimuli. The third application estimates the dynamical correlation of gene expressions during the activation of T-cells.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

With the development of techniques in collecting data, analysis of data with complex structures has become increasingly popular in statistical research. Functional data analysis has attracted extensive attention in fields like image analysis, disease diagnostic and gene regulation (Ramsay and Silverman, 2002; Leng and Müller, 2005; Baladandayuthapani et al., 2008). Particularly, analysis of univariate functional data analysis has been focused on for a long time. Among them, functional regression and functional principal component analysis are two particularly important problems for researchers. For a more comprehensive view of them, refer to monographs like (Ramsay and Silverman, 2005; Ferraty and Vieu, 2006; Ferraty, 2011; Horváth and Kokoszka, 2012). In contrast, multivariate functional data have not been studied popularly.

Dependence modeling plays a significant role in multivariate data analysis. Dependence structure among stochastic processes likewise is worth studying for its own merit in functional data analysis. Heckman and Zamar (2000) proposed the rank correlation coefficient between two random functions in cluster analysis. Dubin and Müller (2005) considered the dynamical correlation of two random functions, which can be regarded as an extension of the correlation coefficient in multivariate data. Yang et al. (2011) developed singular valued decomposition for pairs of functional data.

In this paper, we focus on the dynamical correlation, which summarizes the correlation of two functional variables over their domain. There have been substantial applications related to this concept in the literature. For instance, a graphical Gaussian model for functional data was developed by Opgen-Rhein and Strimmer (2006) based on the dynamical correlation. Liu et al. (2016) applied the dynamical correlation in a psychological study. The dynamical correlation was employed in Bastidas et al. (2014) to study the mechanism when $CD8^+$ T cells are activated in HIV-infected individuals.

* Corresponding author.
E-mail address: jiguo_cao@sfu.ca (J. Cao).

https://doi.org/10.1016/j.csda.2018.07.003
0167-9473/© 2018 Elsevier B.V. All rights reserved.

Dubin and Müller (2005) proposed a sample average estimator for dynamical correlation. However, the asymptotical variance of this estimator is not tractable. Thus they recommended a bootstrap procedure to construct a confidence interval for dynamical correlation based on the average estimator. Since a local linear regression is used to smooth functional data in the pre-processing step, both fitted values and residuals must be resampled in their proposed bootstrap procedure. This leads to a computationally intensive procedure for inferring dynamical correlation, which may restrict its applications in practice. Another concern of this estimator is that each subject makes an equal contribution to the final estimator. This may not be an optimal treatment in practice since subjects with dense observations may provide a more accurate estimate of dynamical correlation in comparison with subjects with less dense observations. An estimator which can adjust weights on each subject is, therefore, more appealing than a simple average of the estimate from each subject. Actually Dubin and Müller (2005) also considered such an estimator in applications under the bootstrap-based inference framework, even though this estimator was not thoroughly investigated in the paper. Furthermore, we find in simulation studies that this bootstrap procedure is unstable and cannot provide a reliable confidence interval for dynamical correlation when the functional data are irregularly spaced.

Motivated by these observations, we propose a weighted average estimator for the dynamical correlation which can adaptively adjust weights on each subject. The weights are chosen via the weighted empirical likelihood. This new estimator can be regarded as an alternative to that mentioned in applications of Dubin and Müller (2005). The statistical inference framework, however, is different from the bootstrap-based analysis proposed by Dubin and Müller (2005). The main advantage of this method is that we do not need to estimate the standard error of the estimator for statistical inference since the test statistic itself is self-normalized under some regularity conditions. This improves computational efficiency substantially for constructing confidence intervals, in comparison with the bootstrap-based method proposed by Dubin and Müller (2005). This improvement will be demonstrated in our simulation studies presented in Section 4. The second advantage of our method is that our proposed weighted empirical likelihood is possible to yield a confidence interval with a more accurate coverage probability than its counterpart by the bootstrap method when functional data are irregularly spaced. Our simulation studies show that the coverage probability of the proposed confidence interval is still close to the nominal level even when the functional data are irregularly designed. In summary, the main advantage of our method in statistical inference is well reflected when there exist great variations in the number of observations across subjects. Note that our method cannot accommodate the case when functional data have sparse observations for every subject.

The paper is organized as follows. The definition of dynamical correlation is reviewed in Section 2. Then a new method of constructing confidence intervals for dynamical correlation based on the weighted empirical likelihood is proposed. For comparison, we introduce the bootstrap method for constructing confidence intervals for dynamical correlation proposed by Dubin and Müller (2005) later. Some theoretical properties of the proposed inference tool are given in Section 3. The proofs are deferred to Appendix B. Section 4 compares the performances of the weighted empirical likelihood-based method and the bootstrap method in associated confidence intervals for dynamical correlation via simulation studies. The proposed method is illustrated with three applications in Section 5. Section 6 concludes the paper. We write our own code to implement the bootstrap method and the weighted empirical likelihood method. The computing codes used in our simulation studies and three applications are available at https://github.com/caojiguo/WEL.

2. Confidence intervals for dynamical correlation

We propose the weighted empirical likelihood method to estimate confidence intervals for the dynamical correlation of two random functions. First, we introduce the definition of the dynamical correlation of two random functions. We then introduce our method. Last, the point estimator and the bootstrap confidence interval for dynamical correlation proposed by Dubin and Müller (2005) are reviewed.

2.1. Dynamical correlation

The dynamical correlation was firstly proposed by Dubin and Müller (2005) to use a single measure to describe the correlation between two longitudinal curves. As noted by them, it is simpler and more efficient, compared with functional canonical correlation (He et al., 2003; Ramsay and Silverman, 2005), which is well defined provided restrictive assumptions.

Let $f_j, j = 1, \ldots, p$, be p random functions defined over a compact interval $\mathcal{I}$. Suppose they belong to $L^2(dw)$, the collection of all square integrable functions with respect to a measure $dw = wdt$. That is, $\mathrm{E}\{\int_{\mathcal{I}} f_j^2(t)w(t)dt\} < \infty, j = 1, \ldots, p$, where the weight function $w(t)$ defined over $\mathcal{I}$ satisfies (i) $w(t) \geq 0,\ t \in \mathcal{I}$, (ii) $\int_{\mathcal{I}} w(t)dt = 1$, (iii) $\int_{\mathcal{I}} w^2(t)dt < \infty$. For any two functions $f, g \in L^2(dw)$, the inner product of them is defined by $\langle f, g\rangle = \int_{\mathcal{I}} f(t)g(t)dw$.

These random functions can be expressed as

$$f_j(t) = \mu_{0,j} + \mu_j(t) + \xi_{0,j} + \sum_{l=1}^{\infty} \xi_{l,j}\phi_l(t),\ j = 1, \ldots, p, \tag{1}$$

where the fixed intercept $\mu_{0,j} = \mathrm{E}(\langle f_j, 1\rangle)$, $\mu_j(t)$ is the fixed mean function with $\mu_j(t) \in L^2(dw)$ and $\langle \mu_j(t), 1\rangle = 0$, and $\xi_{0,j}$ serves as a random intercept. Let $\phi_0(t) \equiv 1$ and we assume that the fixed functions, $\{\phi_l(t)\}_{l=0}^{\infty}$, constitute a complete

orthonormal basis of the space $L^2(dw)$; that is, $\langle \phi_i, \phi_{i'} \rangle = 0$ if $i \neq i'$ and 1 otherwise. The other random components, $\xi_{l,j}$'s, are the coefficients when the random function f_j is represented in terms of this orthonormal basis. We assume that $\xi_{l,j}$'s are uncorrelated with $\xi_{0,j}$ for $l \geq 1$, and they satisfy $\mathrm{E}(\xi_{l,j}) = 0$, $\mathrm{Var}(\xi_{l,j}) = \sigma^2_{l,j} < \infty$, $l \geq 0$ and $0 < \sum_{l=0}^{\infty} \sigma^2_{l,j} < \infty$. Note that the expansion given in (1) is not the Karhunen–Loève expansion of $f_j(t)$. The random coefficients, $\xi_{l,j}$'s ($l \geq 1$), are not required to be uncorrelated with each other. Furthermore, the fixed functions $\mu_j(t)$ and $\phi_l(t)$'s are assumed to be twice continuously differentiable over $\mathcal{I}$.

Let $M_j = \langle f_j, 1 \rangle$. Then $M_j = \mu_{0,j} + \xi_{0,j}$ and $f_j(t) - M_j = \mu_j(t) + \sum_{l=1}^{\infty} \xi_{l,j} \phi_l(t)$. The standardized f_j's are defined as

$$f_j^{\star}(t) = \frac{f_j(t) - M_j - \mu_j(t)}{\left(\int_{\mathcal{I}} (f_j(t) - M_j - \mu_j(t))^2 dw \right)^{\frac{1}{2}}}. \tag{2}$$

The dynamical correlation between f_{j_1} and f_{j_2}, $1 \leq j_1, j_2 \leq p$, is defined as the expected inner product of $f^{\star}_{j_1}$ and $f^{\star}_{j_2}$,

$$\rho_{j_1 j_2} = \mathrm{E} \langle f^{\star}_{j_1}, f^{\star}_{j_2} \rangle.$$

As argued by Dubin and Müller (2005), the dynamical correlation satisfies $|\rho_{j_1 j_2}| \leq 1$ for any two random functions which are square integrable with respect to dw.

2.2. Confidence interval via the weighted empirical likelihood

As mentioned in Section 1, Dubin and Müller (2005) proposed an estimator which averages the point estimate for each subject to estimate dynamical correlation. They demonstrated that the performance of this estimator is satisfactory when random functions are densely and regularly observed. However, they did not explore its performance when random functions are irregularly observed.

Random functions are more often observed at different time points for different subjects, which is called the irregular design in this paper. For simplicity, we assume that for the same subject, p random functions are observed at the same time points, though the method presented in the following can be extended to deal with more general cases as well. More specifically, we have observations $\{(f_1(t_{i1}), \ldots, f_p(t_{i1})), \ldots, (f_1(t_{in_i}), \ldots, f_p(t_{in_i}))\}$ for the ith subject, $i = 1, \ldots, n$, where n denotes the number of subjects. Furthermore, we assume that these observations may also have measurement errors.

Now we introduce a novel approach to construct confidence intervals for dynamical correlation of two random functions under the irregular design. Inspired by the idea of Dubin and Müller (2005), we first employ local linear regression to smooth each random function for all subjects. For $i = 1, \ldots, n$, $1 \leq j_1 \neq j_2 \leq p$, let $f^S_{i,j_1}(t)$ and $f^S_{i,j_2}(t)$ denote smoothed $f_{i,j_1}(t)$ and $f_{i,j_2}(t)$ with local linear smoothing, respectively. To define a sample version of (2), we introduce $\tilde{f}^S_{i,j}(t) = f^S_{i,j}(t) - 1/n \sum_{i=1}^n f^S_{i,j}(t)$ and $\tilde{M}^S_{i,j} = \langle \tilde{f}^S_{i,j}, 1 \rangle$, $i = 1, \ldots, n$, $j = j_1, j_2$. Hence $\tilde{f}^S_{i,j}(t) - \tilde{M}^S_{i,j} = f^S_{i,j}(t) - M^S_{i,j} - 1/n \sum_{i=1}^n (f^S_{i,j}(t) - M^S_{i,j})$. The standardized random function in the ith subject, defined by

$$\hat{f}^{S\star}_{i,j}(t) = \frac{\tilde{f}^S_{i,j}(t) - \tilde{M}^S_{i,j}}{\left(\int_{\mathcal{I}} (\tilde{f}^S_{i,j}(t) - \tilde{M}^S_{i,j})^2 dw \right)^{\frac{1}{2}}}, \tag{3}$$

therefore provides a reasonable estimate of $f^{\star}_j(t)$ defined in (2). As a result, the point estimate of dynamical correlation of random functions f_{j_1} and f_{j_2} provided by the ith subject can be defined as

$$\hat{\rho}^S_{i,j_1 j_2} = \langle \hat{f}^{S\star}_{i,j_1}(t), \hat{f}^{S\star}_{i,j_2}(t) \rangle. \tag{4}$$

Using (4), we are able to obtain a point estimate of dynamical correlation of each pair of random functions for each subject. In the Estimation section, Dubin and Müller (2005) estimated dynamical correlation by averaging these point estimates. However, this may not be an optimal strategy to make use of all $\hat{\rho}^S_{i,j_1 j_2}$'s. Since different subjects have different numbers of measurements, their contributions to the final estimate of dynamical correlation may not be equally important. As a result, a simple average of $\hat{\rho}^S_{i,j_1 j_2}$'s may not be an optimal choice. Dubin and Müller (2005) pointed out and addressed this issue in applications by using weighted averages to estimate dynamical correlation.

A natural problem about choosing appropriate weights arises. Another interesting problem is to construct a valid confidence interval for dynamical correlation based on $\hat{\rho}^S_{i,j_1 j_2}$'s. To solve these problems, we borrow the idea of the weighted empirical likelihood proposed by Wu (2004). The log weighted empirical likelihood is defined as $l(F) = T_n \sum_{i=1}^n c_i \{\log(p_i) - n p_i\}$, where c_i is the weight put on the ith subject, $T_n = n / \sum_{i=1}^n c_i$ and p_i is probability put on $\hat{\rho}^S_{i,j_1 j_2}$ satisfying $\sum_{i=1}^n p_i = 1$. Obviously $l(F)$ achieves its maximum when $p_1 = p_2 = \cdots = p_n = 1/n$. The log likelihood ratio is then as

$$l(F) - \max_F l(F) = T_n \sum_{i=1}^n c_i \{\log(p_i) - n p_i\} - T_n \sum_{i=1}^n c_i \{-\log(n) - 1\} = T_n \sum_{i=1}^n c_i \{\log(n p_i) - n p_i + 1\}.$$

Given the additional constraint that $\mathrm{E}(\hat{\rho}^S_{i,j_1j_2}) \approx \rho_{j_1j_2}$ under some regularity conditions provided in Theorem 2 in Dubin and Müller (2005), the profile log weighted empirical likelihood ratio is defined as

$$\tilde{R}(\rho) = T_n \sum_{i=1}^{n} c_i\{\log(np_i) - np_i + 1\}, \tag{5}$$

where p_i's maximize $l(F)$ subject to $p_i \geq 0, i = 1, \ldots, n$, $\sum_{i=1}^n p_i = 1$ and $\sum_{i=1}^n p_i \hat{\rho}^S_{i,j_1j_2} = \rho$. To account for the effect of different numbers of observations for different subjects, we choose

$$c_i = \frac{1}{n_i} \tag{6}$$

in (5) when applying the weighted empirical likelihood. Wu (2004) suggested χ^2_1 to calibrate the asymptotic distribution of the log weighted empirical likelihood ratio. Likewise, we propose a $1-\alpha$ confidence interval for the true dynamical correlation $\rho_{j_1j_2}$: $\{\rho : -2\tilde{R}(\rho) < \chi^2_1(1-\alpha)\}$, where $\chi^2_1(1-\alpha)$ denotes the $1-\alpha$ quantile of χ^2 distribution of one degree of freedom. Numerical studies presented in Section 4 show that χ^2_1 approximates $-2\tilde{R}(\rho_{j_1j_2})$ reasonably well. As shown in Owen (1988) and Wu (2004), the logarithms of both the unweighted and weighted empirical likelihood ratio converge to a chi-squared distribution under some regularity conditions. But as argued in Wu (2004), these two methods have different performances in statistical inference for samples with a small size. The superiority of the weighted method is demonstrated via a simulation study in Wu (2004), which shows that even though the lengths are similar, the confidence interval generated from the weighted empirical likelihood has a coverage probability closer to the nominal level than its counterpart, especially when the sample size is small.

Next we consider another scenario in which random functions are observed over a common grid of time points across all subjects. This scenario is called the regular design in this paper. We assume that observations consist of $\{(f_{i,1}(t_k), \ldots, f_{i,p}(t_k)),\ i = 1, \ldots, n,\ k = 1, \ldots, m\}$, where $t_1, t_2, \ldots, t_m$ are m distinct points of $\mathcal{I}$. Without loss of generality, we assume that $t_1 < t_2 \ldots < t_m$. Using the same idea as in irregular design, the corresponding log weighted empirical likelihood ratio becomes

$$\tilde{R}(\rho) = \sum_{i=1}^{n} \log(np_i), \tag{7}$$

where p_i's maximize $\prod_{i=1}^n p_i$ subjects to $p_i \geq 0, i = 1, \ldots, n$, $\sum_{i=1}^n p_i = 1$ and $\sum_{i=1}^n p_i \hat{\rho}^S_{i,j_1j_2} = \rho$. If $\hat{\rho}^S_i$'s were identically and independently distributed, $\tilde{R}(\rho)$ is actually the same as the log empirical likelihood ratio defined by Owen (1988). We employ χ^2_1 to calibrate the asymptotic distribution of $-2\tilde{R}(\rho_{j_1j_2})$ as well. This chi-square calibration will be investigated both theoretically and numerically later. Numerical studies show that $\{\rho : -2\tilde{R}(\rho) < \chi^2_1(1-\alpha)\}$ is a reasonable $1-\alpha$ confidence interval for $\rho_{j_1j_2}$.

2.3. Confidence intervals for dynamical correlation via bootstrap

The point estimator for dynamical correlation between $f_{j_1}(t)$ and $f_{j_2}(t)$ proposed by Dubin and Müller (2005) is given by

$$\hat{\rho}_{j_1j_2} = \frac{1}{n} \sum_{i=1}^{n} \hat{\rho}^S_{i,j_1j_2}. \tag{8}$$

They considered confidence intervals for dynamical correlation based on $\hat{\rho}^S_{i,j_1j_2}$'s as well. Provided some regularity conditions and assuming $\mu_j(t)$ is known, $j = j_1, j_2$, they proved that the estimator in (8) is asymptotically normal if $0 < |\rho_{j_1j_2}| < 1$. Based on this theoretical result, a confidence interval for the dynamical correlation would be immediately available if we are able to find a plausible estimate of the variance of the estimator. However, they argued that the variance of $\hat{\rho}_{j_1j_2}$ is rather involved and there is no straightforward estimate of it. They hence suggested employing the bootstrap to obtain a confidence interval. To account for the uncertainty introduced by the pre-smoothing step via a local linear smoother, they proposed to sample the fitted values and residuals from the pre-smoothing step separately. Additionally, a relatively large number, $B = 500$ of bootstrap samples were recommended to obtain a plausible confidence interval. To implement the above procedure to compute the point estimator and the bootstrap confidence interval for dynamical correlation, a package called **dynCorr** was developed by Dubin et al. (2017).

There are several issues associated with the estimator $\hat{\rho}_{j_1j_2}$ and the bootstrap strategy, which may flaw their applications. First of all, each $\hat{\rho}^S_{i,j_1j_2}$ makes equal contribution to $\hat{\rho}_{j_1j_2}$ in (8), the final estimator of $\rho_{j_1j_2}$. As pointed out in Section 2.2, this may not be an optimal treatment under irregular design. Noting this fact, Dubin and Müller (2005) suggested using a weighted average of $\hat{\rho}^S_{i,j_1j_2}$'s to estimate $\rho_{j_1j_2}$ to accommodate the irregular design in applications. Second, the bootstrap procedure was proposed to account for the uncertainty in the pre-smoothing step and to estimate the variance of the average point estimator. But the proposed bootstrap procedure is computationally intensive when the pre-smoothing step is included and

Table 1
Summary of coverage probabilities of two confidence intervals under the irregular design: "B" and "E" stand for the bootstrap and the weighted empirical likelihood based confidences, respectively. The bandwidth used in the local linear smoother is denoted by h and the coverage probabilities are calculated from the 100 Monte Carlo simulations.

Data	n	Method	h													
			.01	.035	.06	.085	.11	.135	.16	.185	.21	.235	.26	.285	.31	.335	.36
Irregular	50	B	.90	.79	.70	.67	.63	.63	.60	.62	.59	.66	.63	.57	.67	.63	.59
		E	.33	.88	.91	.96	.93	.91	.94	.96	.91	.94	.97	.91	.95	.95	.96
	100	B	.47	.58	.49	.43	.41	.38	.40	.39	.42	.35	.44	.33	.46	.35	.43
		E	.12	.87	.97	.96	.93	.94	.92	.93	.96	.94	.91	.98	.91	.95	.96

data are resampled $B = 500$ times. Last but not least, the performance of the bootstrap method is not satisfactory under irregular design. The last issue will be demonstrated in numerical studies presented in Section 4. The weighted empirical likelihood, however, can circumvent estimating the variance since the log weighted likelihood ratio statistics itself is self-normalized with the chi-square as the limiting distribution. Unlike the bootstrap method, this chi-square calibration based on the weighted empirical likelihood results in a confidence interval with an appealing coverage probability even under the irregular design.

3. Theoretical properties

Let $\rho_{j_1j_2}$ denote the true dynamical correlation between $f_{j_1}(t)$ and f_{j_2}, $1 \le j_1, j_2 \le p$, and $j_1 \ne j_2$. Then we have

Theorem 1. *Under the assumptions listed in Appendix A, if* $0 < |\rho_{j_1j_2}| < 1$, *and both* $\mathrm{E}(f_{j_1}(t))$ *and* $\mathrm{E}(f_{j_2}(t))$ *are known, then* $-2\tilde{R}(\rho_{j_1j_2}) \to \chi^2(1)$ *in distribution, where* $\tilde{R}(\cdot)$ *is defined in (7).*

The proof is deferred to Appendix B.

4. Simulation studies

In this section, several simulation scenarios are considered to compare the confidence intervals for dynamical correlation of two random functions based on the bootstrap method and the weighted empirical likelihood. Our main concern consists of computational time and coverage probabilities of these types of confidence intervals.

We generate n identically and independently copies of two random functions, f_1 and f_2, on $[0, 1]$. Here $n = 50$ or 100. Let $Y_{ij}(t)$ denote the measurement of f_j of ith copy at time t. More specifically, $Y_{ij}(t) = f_{ij}(t) + e_{ij}(t)$, $i = 1, \ldots, n$, $j = 1,2$, where

1. The random function is defined as $f_{ij}(t) = 1 + \sum_{l=0}^{2} \xi_{l,j}\phi_l(t)$. In other words, $\mu_j(t) \equiv 0$ and $\mu_{0,j} = 1$ for $j = 1,2$. In addition, there are only three random components in both random functions.
2. The fixed orthogonal functions ϕ_l's take the following form: $\phi_0(t) = 1$, $\phi_1(t) = 2\sqrt{3}(t - 1/2)$ and $\phi_2(t) = 6\sqrt{5}(t - 1/2)^2 - \sqrt{5}/2$.
3. The random components are generated from a centered multivariate Gaussian distribution with a covariance matrix as follows: $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where each block submatrix is defined as $\Sigma_{11} = \mathrm{diag}(1, 1/2, 1/3)$, $\Sigma_{12} = \mathrm{diag}(1/3, 1/4, 1/6)$ and $\Sigma_{22} = \mathrm{diag}(1/2, 1/3, 1/4)$.
4. $e_{ij}(t) \sim N(0, 1/4)$ denotes the measurement error of the observation of f_j from the ith copy evaluated at time t. These measurement errors are assumed to be independent of each other.

Under this setup, the true dynamical correlation between f_1 and f_2 is 0.5. The above simulation setups are the same as in Dubin and Müller (2005).

To compare the performances of the bootstrap method and the weighted empirical likelihood method in associated confidence intervals for dynamical correlation, both irregular and regular designs are considered. In both designs, a grid of equally-spaced bandwidths $\{.01, .035, \ldots, .36\}$ is chosen in the local linear smoother in the pre-smoothing step. The weight function $w(t)$ is taken to be 1 over $[0, 1]$. To evaluate the real coverage probabilities of the $1 - \alpha$ confidence intervals for dynamical correlation generated from both the bootstrap method and the weighted empirical likelihood, $K = 100$ simulation trials are run. Here we focus on $\alpha = 0.05$, namely 95% confidence intervals for dynamical correlation.

Irregular design: In the first scenario, the ith copy of random functions are observed on n_i of these 100 time points, where n_i is uniformly sampled from $\{25, 26, \ldots, 100\}$. Table 1 summarizes the coverage probabilities of these two types of confidence intervals for $\alpha = 0.05$ across the 15 bandwidth choices under the irregular design. It can be observed that there is a remarkable gap between these two methods in terms of coverage probability. The weighted empirical likelihood method is robust to the choice of bandwidths and the coverage probabilities of the confidence intervals obtained from most bandwidth choices are close to the nominal level. In contrast, the performance of the bootstrap method is not satisfactory at all, no matter from the perspective of robustness or from coverage probability.

Table 2
Summary of coverage probabilities of two confidence intervals under the regular design. "B" and "E" stand for the bootstrap and the weighted empirical likelihood based confidences, respectively. The bandwidth used in the local linear smoother is denoted by h and the coverage probabilities are calculated from the 100 Monte Carlo simulations.

Data	n	Method	h														
			.01	.035	.06	.085	.11	.135	.16	.185	.21	.235	.26	.285	.31	.335	.36
Regular	50	B	.92	.98	.94	.98	.99	.95	.96	.98	.94	.94	.98	.94	.95	.96	.94
		E	.79	.95	.96	.88	.95	.91	.98	.96	.97	.99	.95	.94	.94	.92	.91
	100	B	.83	.94	.97	.96	.97	.96	.94	.96	.95	.95	.94	.94	.94	.94	.94
		E	.63	.94	.97	.95	.94	.90	.97	.93	.97	.96	.94	.99	.93	.89	.93

Table 3
Summary of average computation time of two confidence intervals: "B" and "E" stand for the bootstrap and the weighted empirical likelihood based confidences, respectively. The average computation times are calculated from the 100 Monte Carlo simulations and all bandwidths considered in Tables 1 and 2.

Data	n	Method	Time (s)
Irregular	50	B	205
		E	0.44
	100	B	387
		E	0.85
Regular	50	B	214
		E	0.55
	100	B	450
		E	1

Regular design: In the second scenario, the observational time points for both random functions are chosen to be $m = 100$ equidistant points on $[0, 1]$. Under the regular design, Table 2 indicates that both the bootstrap confidence interval and the empirical likelihood-based confidence interval are robust to the choice of the bandwidth in the local linear smoother. Furthermore, the coverage probabilities of both of these two confidence intervals are quite close to the nominal level, even though the bootstrap method is slightly better in some exceptional choices of bandwidth.

A further comparison between these two methods is made in terms of computational time. According to Table 3, the weighted empirical likelihood method is considerably more efficient than the bootstrap method in computation. The bootstrap method spends substantial time on estimating the variability of the average estimator of dynamical correlation by resampling both the smoothed random functions and residuals when a local linear smoother is involved. Nevertheless, this step can be circumvented in the weighted empirical likelihood method. As pointed out by one referee, it should be noted that the bootstrap method is able to account for the uncertainty in the pre-smoothing step, while the weighted empirical likelihood method ignores this source of uncertainty in constructing confidence intervals for dynamical correlation.

To demonstrate that χ^2_1 provides a reasonable calibration when constructing confidence intervals for dynamical correlation based on the weighted empirical likelihood method, Fig. 1 compares the empirical distribution of the weighted empirical likelihood ratio test statistics with χ^2_1 in the four different cases. For all four panels in Fig. 1, $h = 0.21$ is used in the local linear smoothing. Due to the robustness of the weighted empirical likelihood method, this specific choice should not play a critical role in determining shapes of these four panels. It turns out that χ^2_1 can provide a satisfactory approximation to the distribution of the empirical likelihood ratio test statistic even for a moderate number of subjects for both irregular and regular designs. In the top right panel, five points with a theoretical quantile greater than 4 deviate slightly from the straight line. This might seem to be a serious issue when compared with the top left panel, in which only 50 independent copies of random functions are generated. However, quantiles greater than 4 for the χ^2_1 distribution correspond to quantiles greater than 2 for the $N(0, 1)$ distribution; thus they are greater than 95% quantile of $N(0, 1)$. Usually an extreme quantile cannot be estimated accurately due to data sparsity near the extreme quantile.

5. Applications

In this section, we apply the proposed empirical likelihood-based method to analyze the dynamical correlation between random functions in three applications.

5.1. Dynamical correlation of air pollutants

Dynamical correlation between air pollutants is treated in the first example. The dataset is obtained from the **NMMAPSdata** package (Peng and Welty, 2004), which contains daily mortality, air pollution, and weather data for the study of national morbidity, mortality, and air pollution. Six air pollutants, including PM_{10}, $PM_{2.5}$, SO_2, O_3, NO_2, and CO, are measured in 108 cities of America from 1987 to 2000. There has been extensive research on the adverse effect of $PM_{2.5}$ and NO_2 on health

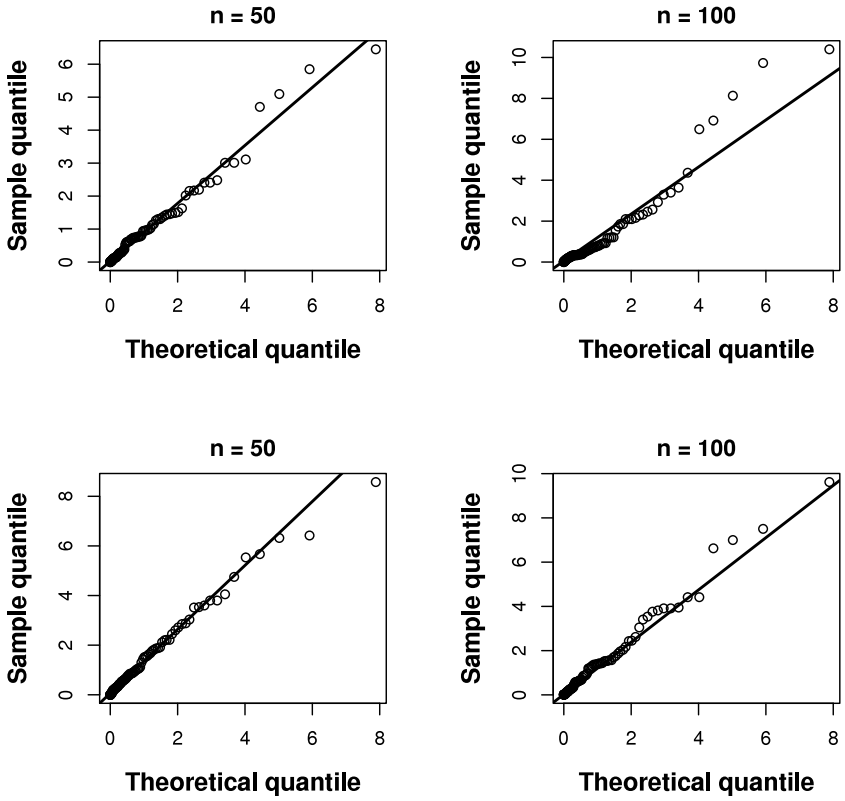


Fig. 1. Q–Q plots of the weighted empirical likelihood ratio statistics. The upper and bottom panels are for the irregular and regular designs, respectively. The left and right panels correspond to $n = 50$ and $n = 100$, respectively.

of human beings (see Atkinson et al. (2014), Hyder et al. (2014) and Crouse et al. (2015)). Ritchine and Roser (2018) argued that $PM_{2.5}$ is one of the most concerning air pollutants which have an adverse impact on human health. Actually limited recordings of $PM_{2.5}$ were available before the 2000s in many regions (Peng and Welty, 2004; Kumar and Joseph, 2006; San Martini et al., 2015). Thus it is crucial to understand the relationships between $PM_{2.5}$ and other regularly monitored air pollutants like PM_{10} and NO_2 (Kumar and Joseph, 2006). According to Connell et al. (2005), there is a statistically significant association between the concentrations of $PM_{2.5}$ and NO_2 after removing autocorrelation. Furthermore, they also claim that this association varies from season to season. Kumar and Joseph (2006) showed that, for both ambient and kerbsite locations in Mumbai, there is a strong positive correlation between $PM_{2.5}$ and NO_2. It indicates that NO_2 and $PM_{2.5}$ share a common origin. A two-step method was proposed in Connell et al. (2005): a time series filter was employed to account for the temporal correlation within both NO_2 and $PM_{2.5}$ measurements, and then the Pearson correlation analysis was carried out on the residuals to estimate their associations. Kumar and Joseph (2006), however, conducted a regression analysis directly to estimate the correlations between these air pollutants. Neither methods treated repeated measurements of NO_2 and/or $PM_{2.5}$ as a sample of a smooth random function and considered the correlation between two random functions, even though the former one can allow for the temporal correlations. In this study, we are interested in investigating the association between these two air pollutants from a novel perspective: dynamical correlation, which is directly based on the repeated measurements of them while accounting the temporal correlation within each of them.

Since from 1999 daily measurements for $PM_{2.5}$ are available in most cities (Peng and Welty, 2004), we choose measurements in 2000 for analysis for the consideration of fewer missing data. Cities without any measurement of $PM_{2.5}$ or NO_2 are excluded; then in total, we have measurements from 66 cities in the year of 2000. The left panel in Fig. 2 displays the standardized profiles of $PM_{2.5}$ or NO_2 from one randomly selected city. These two air pollutants show a strong positive correlation in fall and winter and somewhat weak correlation in spring and summer according to this single city. Statistical inference based on the weighted empirical likelihood suggests that there is a moderate positive dynamical correlation between these two air pollutants if we aggregate the association in each season. In particular, the 95% confidence interval for the dynamical correlation is (0.38, 0.44). This conclusion does not contradict the mechanism how $PM_{2.5}$ is generated. On the one hand, gases like NO_2 and $PM_{2.5}$ may be emitted from the same source (Kumar and Joseph, 2006) and the former may serve as precursors to secondary $PM_{2.5}$ formation (Connell et al., 2005). On the other hand, correlations among these air pollutants exhibit strong seasonality. These may explain why there exists a moderate positive dynamical correlation between $PM_{2.5}$ and NO_2.

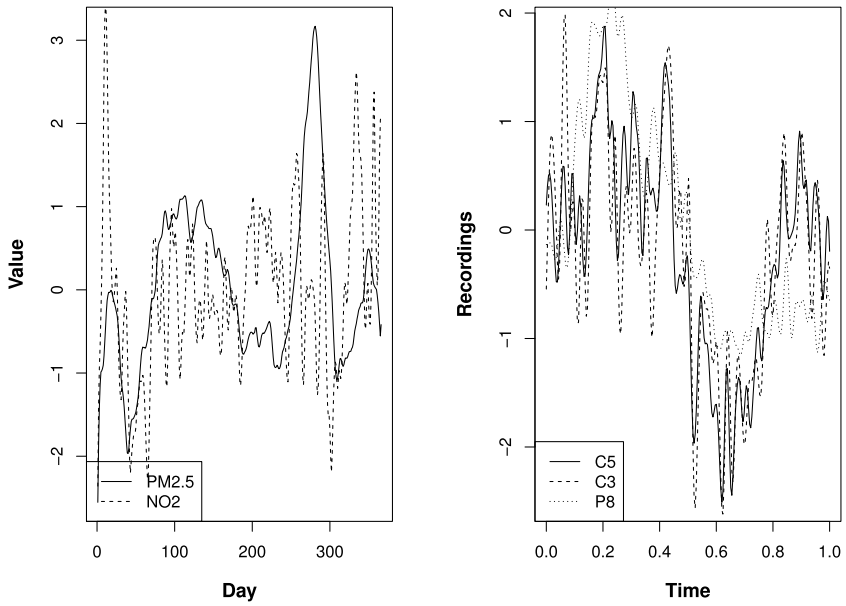


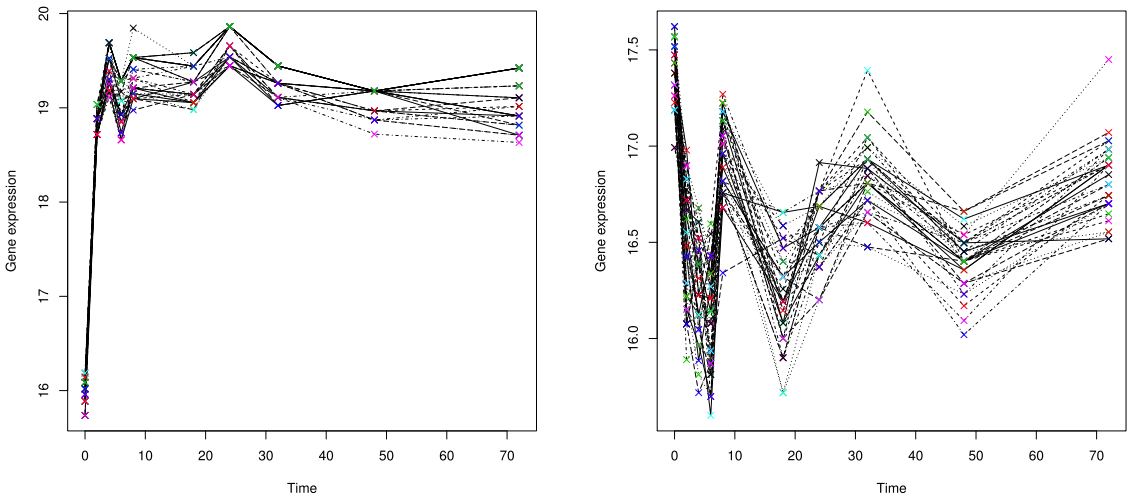
Fig. 2. Left: Standardized profiles of $PM_{2.5}$ or NO_2 in 2000 from one randomly selected city. Right: Standardized profiles of three electrodes from one randomly selected subject.

5.2. Dynamical correlation of EEG signals

Zhang et al. (1995) studied the object recognition process by measuring electroencephalogram (EEG) signals using pictures of objects. In their experiment, 64 electrodes are placed on subject's scalps to investigate EEG correlates of genetic predisposition to alcoholism. There are 122 subjects in this study, 77 of them in the alcoholic group and the remaining in the control group. Every subject is measured in 120 trials. In each trial, 256 observations were taken at each scalp in one second when a stimulus was shown for a subject. Each subject was exposed to a single stimulus or two stimuli. The two stimuli can be in a matched or non-matched condition. This dataset is available at the following link: https://archive.ics.uci.edu/ml/datasets/eeg+database.

In neuroimaging studies, there has been extensive research on functional connectivity, which refers to the statistical association or interactions between different brain regions (Ombao et al., 2016). For instance, when studying a EEG dataset, Marshall et al. (2014) found that significant functional connectivity is identified between short-range sites (neighboring electrodes) considerably more often than between electrodes in long distances, when an individual is not performing an explicit task. Though using different measures to quantify functional connectivity, Qiao et al. (2017) drew a similar conclusion when applying the partial-correlation based framework to analyze the EEG data mentioned in last paragraph. We are interested in studying functional connectivity from the perspective of dynamical correlation. To examine whether the dynamical correlation between short-range electrodes is stronger than that between long-range electrodes, we choose three electrodes: C5, C3, and P8. The first two electrodes are neighboring sites located in the central region, while the last electrode is in the temporal region and situated a long distance away from the first two. According to Miller and Ivanitsky (2003), the central region is in charge of functional integration of the tactile task while the temporal region is the integration center for auditory attention.

To alleviate the effect of an individual's status (alcoholic or non-alcoholic) on functional connectivity between two electrodes, our study is focused on the alcoholic group. For each subject in the alcoholic group, we average recordings under the single stimulus condition across 120 trials at each observational time point. Our primary interest here is to explore dynamical correlation among the recordings of these three electrodes, which can serve as another measure of functional connectivity. Since P8 is situated a long distance away from both C5 and C3, which are neighboring electrodes, we suspect that the dynamical correlation between C5 and C3 are the most significant, compared with the other two counterparts. The right panel in Fig. 2 depicts the standardized profile, defined in (3), of these three electrodes from one randomly selected subject. It can be observed that C5 and C3 maintain synchronization almost across the whole time course while disagreement appears between P8 and C5 or C3. With the empirical likelihood method, we find that the 95% confidence intervals for ρ_{12}, ρ_{13} and ρ_{23} are (0.77, 0.86), (0.36, 0.51) and (0.29, 0.44), where ρ_{12}, ρ_{13} and ρ_{23} denote the dynamical correlation of C5 and C3, C5 and P8, C3 and P8, respectively. This conclusion is not surprising to us since electrodes must maintain synchronization to govern our brain to make a response to a stimulus. The electrodes closer to each other may be more consistent compared with electrodes far apart. Additionally, this result demonstrates the hypothesis that functional connectivity between short-range electrodes is stronger than that between long-range electrodes from the perspective of dynamical correlation of



(a) gene 9 (b) gene 45

Fig. 3. (a) Profiles of expressions of gene 9 of 34 replicates. (b) Profiles of expressions of gene 45 of 34 replicates. Measurements of gene expressions are marked with $\times$.

random functions. A similar result is obtained from using the bootstrap method to compute confidence intervals: (0.77, 0.86), (0.35, 0.51) and (0.28, 0.44) for ρ_{12}, ρ_{13} and ρ_{23}, respectively. The consistency between these methods suggests that in the regular design, ignoring the uncertainty in the pre-smoothing step does not impose a notable impact on confidence intervals obtained from the weighted empirical likelihood method.

5.3. Dynamical correlation of gene expressions

According to Wu et al. (2014), the activation of T lymphocytes (T-cells) plays a key role in generation of an immune response. Two experiments regarding the response of a human T-cell line to phorbol myristyl acetate and ionomicin treatment were conducted by Rangel et al. (2004) to model the gene regulation network during T-cell activation. They collected the expression of 88 genes across 10 times points (0, 2, 4, 6, 8, 18, 24, 32, 48, 72 hr) and 34 replicates were obtained for each gene. The expression profiles of gene 9 and gene 45 are presented in Fig. 3. Obviously this is a considerably less dense design compared with the EEG example.

Wu et al. (2014) applied sparse additive ordinary differential equation models to identify a dynamic network among these genes. Both Rangel et al. (2004) and Wu et al. (2014) found that FYB (gene 45) was one of the genes which regulated expressions of many other genes. Furthermore, Wu et al. (2014) found that the gene FYB always had a positive regulation on gene 9 regardless of the expression level of gene 9. Dynamical correlation provides another perspective to look at the association of gene expressions during T-cell activation. We apply both the empirical likelihood method and the bootstrap method to compute the 95% confidence interval for the dynamical correlation of gene 9 and gene 45. The confidence interval given by the empirical likelihood method is (0.16, 0.52), while its counterpart is (0.10, 0.51). This finding is consistent with that in Wu et al. (2014). In this example, the random functions are considerably less densely observed, the confidence interval obtained from the bootstrap method is slightly wider than that from the empirical likelihood method. The primary reason might be that the bootstrap method accounts for the uncertainty in smoothing each curve from only a few observations; this uncertainty has a non-negligible impact on the interval estimate of the dynamical correlation.

6. Conclusions and discussion

This article proposes a new method based on the weighted empirical likelihood to infer the dynamical correlation between random functions. Compared with the bootstrap method by Dubin and Müller (2005), the proposed method is considerably more efficient in computation due to the self-normalized property of the weighted empirical likelihood ratio statistic. Another reason is that the bootstrap method spends substantial time on sampling both fitted values and residuals to account for uncertainty in the pre-smoothing step. Even though this source of uncertainty is ignored in the weighted empirical likelihood method, simulation studies demonstrate that the weighted empirical likelihood method is possible to yield a confidence interval with a more accurate coverage probability than the bootstrap method when the random functions of interest are irregularly spaced. But whether it is crucial to incorporate this source in the weighted empirical likelihood method in real datasets needs to be carefully investigated in future, especially when functional data are irregularly observed. Furthermore, the proposed method accounts for the differing effect of each pair of random functions on the final estimator of

dynamical correlation, rather than simply averaging these effects. Hence it compares favorably with the bootstrap method in statistical inference of dynamical correlation when there is a remarkable variation in the number of observations of functional data across subjects. In light of these advantages, it is worthwhile to explore applications of this methodology in other similar scenarios. But it should be noted that the proposed method cannot be directly applied to sparse functional data, in which only a few observations are available for every subject. For sparse functional data, we suggest to first recover the trajectory of each functional data from sparse observations. We will investigate this problem in our future work.

At first glance, dynamical correlation is not a powerful tool to reveal the dynamical feature of associations between random functions, since it summarizes correlations of two random functions over the domain of these functions with a single quantity. However, as pointed out by one referee, the dynamical feature can be discovered by adaptively choosing the weight function $w(t)$. In addition, in order to better reveal dynamical features of random functions and accommodate lagged relationships, the definition of dynamical correlation was extended to incorporate lag terms in Dubin and Müller (2005). Dynamical functional connectivity has attracted attention from numerous researchers in neuroimaging studies. It would be interesting work to compare dynamical correlation with varying weight functions with the framework of dynamical functional connectivity.

Acknowledgments

The authors are very grateful for the constructive comments and suggestions from the Editor, an Associate Editor, and two reviewers. These comments and suggestions are very helpful for the authors to improve their work. This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grants RGPIN: 435713-2013 and 2018-06008 of Liangliang Wang and Jiguo Cao, respectively.

Appendix A. Technical assumptions

In this section, we lay out regularity conditions which can guarantee that the difference between the smoothed random function $f_{i,j}^S$ and the true underlying random function $f_{i,j}$ is negligible. Suppose the observations satisfy $Y_{i,j}(t_k) = f_{i,j}(t_k) + e_{i,j}(t_k)$, $i = 1, \ldots, n$, $j = 1, \ldots, p$ and $k = 1, \ldots, m$. We assume that the random errors $e_{i,j}(t_k)$ are identically and independently distributed with $\mathrm{E}(e_{i,j}(t_k)) = 0$ and $\mathrm{E}(e_{i,j}^2(t_k)) = \sigma^2$. Let h denote the bandwidth used in the local linear smoother in the pre-smoothing step. All limits below are taken given the number of subjects, n diverges.

Condition 1. Observational time points $\{t_k\}_{k=1}^m$ follow a density g, which is equi-continuous and bounded away from 0. Furthermore, the support of g is $\mathcal{I}$, which is compact.

Condition 2. The random errors satisfy $\mathrm{E}(|e_{i,j}(t_k)|^s) < \infty$ for some $s > 2$.

Condition 3. The bandwidth h and the number of observations satisfy $\underline{\lim} mh^2 > 0$ and $\underline{\lim}\left(\frac{mh}{\log m}\right)^{\frac{1}{2}} m^{-\frac{2}{s-\eta}} > 0$ for any η with $0 < \eta < 2$.

Condition 4. The random functions $f_{i,j}$ are twice differentiable. Furthermore, the second derivatives are equi-continuous satisfying $\sup|f_{i,j}''(t)| = O_p(1)$, where the supremum is taken over $i = 1, \ldots, n, j = 1, \ldots, p$ and $t \in \mathcal{I}$.

Condition 5. The number of observations for each random function satisfies

$$m \to \infty, \quad n\left(\frac{\log m}{m}\right)^{\frac{4}{5}} \to 0.$$

Condition 6. The bandwidth h satisfies

$$\sqrt{n}h^2 \to 0, \quad n\left(\frac{\log m}{mh}\right)^{\frac{1}{2}} \to 0.$$

Appendix B. Proof of theorem 1

When $\mu_{j_1}(t)$ and $\mu_{j_2}(t)$ are known or unknown but a constant, let $\tilde{f}_{i,j_1}^{\star}$ and $\tilde{f}_{i,j_2}^{\star}$ be the standardized curves for $f_{j_1}(t)$ and $f_{j_2}(t)$, respectively, for subject i. Then define

$$\rho_{i,j_1j_2} = \int \tilde{f}_{i,j_1}^{*}(t)\tilde{f}_{i,j_2}^{*}(t)w(t)\mathrm{d}t.$$

Thus $\rho_{1,j_1j_2}, \ldots, \rho_{n,j_1j_2}$, are identically and independently distributed with $E(\rho_{i,j_1j_2}) = \rho_{j_1j_2}$. $\mathrm{Var}(\rho_{i,j_1j_2}) = \delta^2 < \infty$. Furthermore, Dubin and Müller (2005) showed $0 < \mathrm{Var}(\rho_{i,j_1j_2}) < 1$ given $0 < |\rho_{j_1j_2}| < 1$. In Owen (2001), the profile empirical likelihood function of θ based on ρ_{i,j_1j_2}'s is defined as

$$\tilde{L}(\theta) = \sup\{\prod_{i=1}^{n}(np_i) : p_i \geq 0, i = 1, \ldots, n; \ \sum_{i=1}^{n} p_i = 1; \ \sum_{i=1}^{n} p_i\rho_{i,j_1j_2} = \theta\} \tag{9}$$

As shown in Theorem 2.2 in Owen (2001), $-2\log\tilde{L}(\rho_{j_1j_2})$ converges in distribution to $\chi^2(1)$ as n diverges, if $0 < |\rho_{j_1j_2}| < 1$. We list some important steps to verify this result; for details, see Owen (2001).

The weights which maximize $\tilde{L}(\rho_{j_1j_2})$ can be written as

$$\tilde{w}_i = \frac{1}{n}\frac{1}{1+\tilde{\lambda}(\rho_{i,j_1j_2}-\rho_{j_1j_2})},$$

where $\tilde{\lambda}$ satisfies the equation

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\rho_{i,j_1j_2}-\rho_{j_1j_2}}{1+\lambda(\rho_{i,j_1j_2}-\rho_{j_1j_2})}=0$$

Hence,

$$|\tilde{\lambda}| = O_p(n^{-1/2}); \qquad \max_{1\le i\le n}|\rho_{i,j_1j_2}-\rho_{j_1j_2}| = o_p(n^{1/2}); \qquad |\bar{\rho}_{j_1j_2}-\rho_{j_1j_2}| = O_p(n^{-1/2}), \tag{10}$$

where $\bar{\rho}_{j_1j_2} = \frac{1}{n}\sum_{i=1}^{n}\rho_{i,j_1j_2}$.

With (2), for the smoothed random functions via a local linear smoother we can obtain the corresponding standardized curves, denoted by $\tilde{f}^S_{i,j_1}(t)$ and $\tilde{f}^S_{i,j_2}(t)$. Our main result concerns the asymptotic distribution of log-empirical likelihood function constructed based on the smoothed version of ρ_{i,j_1j_2}, denoted by ρ^S_{i,j_1j_2}, which is defined as

$$\rho^S_{i,j_1j_2} = \int \tilde{f}^S_{i,j_1}(t)\tilde{f}^S_{i,j_2}(t)w(t)\mathrm{d}t,\ i=1,\ldots,n.$$

Specifically, the corresponding profile empirical likelihood function of θ is

$$L(\theta) = \sup\{\prod_{i=1}^{n}(np_i) : p_i \ge 0, i=1,\ldots,n;\ \sum_{i=1}^{n}p_i = 1;\ \sum_{i=1}^{n}p_i\rho^S_{i,j_1j_2} = \theta\}. \tag{11}$$

Let $\bar{\rho}^S_{j_1j_2} = \frac{1}{n}\sum_{i=1}^{n}\rho^S_{i,j_1j_2}$, $V^S_{j_1j_2} = \frac{1}{n}\sum_{i=1}^{n}(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})^2$ and $V_{j_1j_2} = \frac{1}{n}\sum_{i=1}^{n}(\rho_{i,j_1j_2}-\rho_{j_1j_2})^2$. When $\mu_j(t)$ is constant or known for $j=j_1,j_2$, under the assumptions in Appendix A, from the proof of Theorem 2 in Dubin and Müller (2005), we have

$$|\bar{\rho}^S_{j_1j_2}-\bar{\rho}_{j_1j_2}| = o_p(n^{-\frac{1}{2}}), \qquad \max_{1\le i\le n}|\rho^S_{i,j_1j_2}-\rho_{i,j_1j_2}| = o_p(n^{-1/2}) \tag{12}$$

The weights maximizing $L(\rho_{j_1j_2})$ can be written as

$$w_i = \frac{1}{n}\frac{1}{1+\lambda(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})},$$

where $\lambda = \lambda(\rho_{j_1j_2})$ satisfies the equation

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\rho^S_{i,j_1j_2}-\rho_{j_1j_2}}{1+\lambda(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})}=0.$$

Let $M_n = \max_{1\le i\le n}|\rho^S_{i,j_1j_2}-\rho_{j_1j_2}|$. Based on (10) and (12), we obtain

$$|\bar{\rho}^S_{j_1j_2}-\rho_{j_1j_2}| = O_p(n^{-1/2}), \qquad M_n = o_p(n^{1/2}). \tag{13}$$

The next step is to verify that the order of λ is $O_p(n^{-1/2})$ as well. After some simple algebra, it can shown that

$$|\lambda|\cdot\frac{1}{n}\sum_{i=1}^{n}\frac{(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})^2}{1+\lambda(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})} = sgn(\lambda)(\bar{\rho}^S_{j_1j_2}-\rho_{j_1j_2}).$$

Note that $w_i > 0$, i.e., $1+\lambda(\rho^S_{i,j_1j_2}-\rho_{j_1j_2}) > 0$. Thus

$$\begin{aligned}|\lambda|V^S_{j_1j_2} &\le |\lambda|\cdot\frac{1}{n}\sum_{i=1}^{n}\frac{(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})^2}{1+\lambda(\rho^S_{i,j_1j_2}-\rho_{j_1j_2})}(1+|\lambda|M_n)\\ &= sgn(\lambda)(\bar{\rho}^S_{j_1j_2}-\rho_{j_1j_2})(1+|\lambda|M_n).\end{aligned}$$

Since $V_{j_1j_2} = \delta^2 + o_p(1)$, $V^S_{j_1j_2} - V_{j_1j_2} = o_p(1)$, and (13), it follows that

$$|\lambda|(\delta^2+o_p(1)) = O_p(n^{-1/2}),$$

and hence,

$$|\lambda| = O_p(n^{-1/2}).$$

It follows that

$$\begin{aligned} 0 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\rho^S_{i,j_1j_2} - \rho_{j_1j_2}}{1+\lambda(\rho^S_{i,j_1j_2} - \rho_{j_1j_2})} \\ &= \bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2} - \lambda V^S_{j_1j_2} + o_p(n^{-1/2}). \end{aligned}$$

Thus, $\lambda = (V^S_{j_1j_2})^{-1}(\bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2}) + o_p(n^{-1/2})$.

Plugging in the expression of w_i maximizing $L(\rho_{j_1j_2})$, we have

$$\begin{aligned} -2\log L(\rho_{j_1j_2}) &= 2\sum_{i=1}^{n}\log(1+\lambda(\rho^S_{i,j_1j_2} - \rho_{j_1j_2})) \\ &= 2\left(\sum_{i=1}^{n}\lambda(\rho^S_{i,j_1j_2} - \rho_{j_1j_2}) - \frac{\lambda^2(\rho^S_{i,j_1j_2} - \rho_{j_1j_2})^2}{2}\right) + o_p(1) \\ &= 2n(V^S_{j_1j_2})^{-1}(\bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2})^2 - n(\bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2})^2(V^S)^{-1}_{j_1j_2} + o_p(1) \\ &= n(\bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2})^2(V^S_{j_1j_2})^{-1} + o_p(1) \end{aligned} \tag{14}$$

As shown in Owen (2001),

$$-2\log\tilde{L}(\rho_{j_1j_2}) = \frac{n(\bar{\rho}_{j_1j_2} - \rho_{j_1j_2})^2}{V_{j_1j_2}} + o_p(1). \tag{15}$$

Based on (12), (14) and (15), it follows that

$$\begin{aligned} -2\log L(\rho_{j_1j_2}) &= n(\bar{\rho}^S_{j_1j_2} - \rho_{j_1j_2})^2(V^S_{j_1j_2})^{-1} + o_p(1) \\ &= -2\log\tilde{L}(\rho_{j_1j_2}) + o_p(1) \\ &\to \chi^2(1) \qquad \text{in distribution.} \quad \square \end{aligned}$$

References

Atkinson, R., Kang, S., Anderson, H., Mills, I., Walton, H., 2014. Epidemiological time series studies of PM2. 5 and daily mortality and hospital admissions: a systematic review and meta-analysis. Thorax 69, 660–665.

Baladandayuthapani, V., Mallick, B.K., Young Hong, M., Lupton, J.R., Turner, N.D., Carroll, R.J., 2008. Bayesian hierarchical spatially correlated functional data analysis with application to colon carcinogenesis. Biometrics 64 (1), 64–73.

Bastidas, S., Graw, F., Smith, M.Z., Kuster, H., Günthard, H.F., Oxenius, A., 2014. CD8+ T cells are activated in an antigen-independent manner in HIV-infected individuals. J. Immunol. 192 (4), 1732–1744.

Connell, D.P., Withum, J.A., Winter, S.E., Statnick, R.M., Bilonick, R.A., 2005. The steubenville comprehensive air monitoring program (scamp): associations among fine particulate matter, co-pollutants, and meteorological conditions. J. Air Waste Manage. Assoc. 55 (4), 481–496.

Crouse, D.L., Peters, P.A., Hystad, P., Brook, J.R., van Donkelaar, A., Martin, R.V., Villeneuve, P.J., Jerrett, M., Goldberg, M.S., Pope III, C.A., et al., 2015. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the canadian census health and environment cohort (canchec). Environ. Health Perspect. 123 (11), 1180–1186.

Dubin, J., Li, M., Qiao, D., Müller, H.-G., 2017. dynCorr: Dynamic Correlation Package. URL https://CRAN.R-project.org/package=dynCorr. R package version 1.1.0.

Dubin, J.A., Müller, H.-G., 2005. Dynamical correlation for multivariate longitudinal data. J. Amer. Statist. Assoc. 100 (471), 872–881.

Ferraty, F., 2011. Recent Advances in Functional Data Analysis and Related Topics. Springer, New York.

Ferraty, F., Vieu, P., 2006. Nonparametric Functional Data Analysis: Theory and Practice. Springer, New York.

He, G., Müller, H.-G., Wang, J.-L., 2003. Functional canonical analysis for square integrable stochastic processes. J. Multivariate Anal. 85 (1), 54–77.

Heckman, N.E., Zamar, R.H., 2000. Comparing the shapes of regression functions. Biometrika 87 (1), 135–144.

Horváth, L., Kokoszka, P., 2012. Inference for Functional Data with Applications. Springer, New York.

Hyder, A., Lee, H.J., Ebisu, K., Koutrakis, P., Belanger, K., Bell, M.L., 2014. PM2. 5 exposure and birth outcomes: use of satellite-and monitor-based data. Epidemiology 25 (1), 58–67.

Kumar, R., Joseph, A.E., 2006. Air pollution concentrations of $PM_{2.5}$, PM_{10} and NO_2 at ambient and kerbsite and their correlation in Metro City–Mumbai. Environ. Monit. Assess. 119 (1–3), 191–199.

Leng, X., Müller, H.-G., 2005. Classification using functional data analysis for temporal gene expression data. Bioinformatics 22 (1), 68–76.

Liu, S., Zhou, Y., Palumbo, R., Wang, J.-L., 2016. Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data.. Psychol. Methods 21 (3), 291–308.

Marshall, W.J., Lackner, C.L., Marriott, P., Santesso, D.L., Segalowitz, S.J., 2014. Using phase shift granger causality to measure directed connectivity in eeg recordings. Brain Connect. 4 (10), 826–841.

Miller, R., Ivanitsky, A.M., 2003. Complex Brain Functions: Conceptual Advances in Russian Neuroscience. CRC Press, London.

Ombao, H., Lindquist, M., Thompson, W., Aston, J., 2016. Handbook of Neuroimaging Data Analysis. CRC Press, London.

Opgen-Rhein, R., Strimmer, K., 2006. Inferring gene dependency networks from genomic longitudinal data: a functional data approach. RevStat 4 (1), 53–65.

Owen, A.B., 1988. Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75 (2), 237–249.

Owen, A.B., 2001. Empirical Likelihood. Chapman and Hall, London.
Peng, R.D., Welty, L.J., 2004. The NMMAPSdata package. R News 4 (2), 10–14.
Qiao, X., Guo, S., James, G.M., 2017. Functional graphical models. J. Amer. Statist. Assoc. http://dx.doi.org/10.1080/01621459.2017.1390466.
Ramsay, J.O., Silverman, B.W., 2002. Applied Functional Data Analysis: Methods and Case Studies. Springer, New York.
Ramsay, J.O., Silverman, B.W., 2005. Functional Data Analysis. Springer, New York.
Rangel, C., Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., Gaiba, A., Wild, D.L., Falciani, F., 2004. Modeling T-cell activation using gene expression profiling and state-space models. Bioinformatics 20 (9), 1361–1372.
Ritchine, H., Roser, M., 2018. Air Pollution. OurWorldInData.
San Martini, F.M., Hasenkopf, C.A., Roberts, D.C., 2015. Statistical analysis of pm2. 5 observations from diplomatic facilities in china. Atmos. Environ. 110, 174–185.
Wu, C., 2004. Weighted empirical likelihood inference. Statist. Probab. Lett. 66 (1), 67–79.
Wu, H., Lu, T., Xue, H., Liang, H., 2014. Sparse additive ordinary differential equations for dynamic gene regulatory network modeling. J. Amer. Statist. Assoc. 109 (506), 700–716.
Yang, W., Müller, H.-G., Stadtmüller, U., 2011. Functional singular component analysis. J. R. Stat. Soc. Ser. B Stat. Methodol. 73 (3), 303–324.
Zhang, X.L., Begleiter, H., Porjesz, B., Wang, W., Litke, A., 1995. Event related potentials during object recognition tasks. Brain Res. Bull. 38 (6), 531–538.