Statistica Sinica **35** (2025), 1-21
doi:https://doi.org/10.5705/ss.202023.0146

STATISTICAL INFERENCE FOR MULTIVARIATE FUNCTIONAL PANEL DATA

Shuang Sun and Lijian Yang*

Tsinghua University

Abstract: Statistical inference is developed for vector-valued functional panel data which are i.i.d. with respect to subjects and infinite moving average in time. B-spline estimation is proposed for trajectories, which are used to construct a two-step estimator of the vector mean function. By using explicit Gaussian strong approximation in vector form, in the context of moving average panel, the proposed spline estimator is shown to be oracally efficient in the sense that it is asymptotically equivalent to the infeasible estimator with all trajectories known. This deep theoretical result points to a limiting Gaussian distribution of the vector mean estimator, which allows for the construction of various simultaneous confidence region (SCR) for the vector mean function itself and linear combination of its elements. Asymptotic correctness of the SCRs is both established in theory and validated in simulation experiments. The proposed SCRs are applied to an Electroencephalogram (EEG) multivariate functional panel data set, validating multiple scientific facts.

Key words and phrases: B-spline, Electroencephalogram, moving average, oracle efficiency, simultaneous confidence region.

1. Introduction

Functional data analysis (FDA) has been an important area of statistics research for more than two decades. Comprehensive introduction to FDA can be found in Ramsay and Silverman (2005), Ferraty and Vieu (2006), Hsing and Eubank (2015), and Kokoszka and Reimherr (2017).

A functional random variable is a square-integrable continuous stochastic process: specifically, $\eta(\cdot) \in \mathcal{C}[0,1]$ almost surely, with $\mathbb{E}\sup_{x\in[0,1]}\eta^2(x) < \infty$. For such $\eta(\cdot)$, both mean function $\mathbb{E}\eta(\cdot)$ and covariance function $\mathrm{cov}\{\eta(x), \eta(x')\}$, $x, x' \in [0,1]$ exist and are continuous. A functional random vector is a vector-valued functional random variable $\boldsymbol{\eta}(\cdot) = \{\eta^{(1)}(\cdot), \ldots, \eta^{(L)}(\cdot)\}^\top$ where each element is a square-integrable continuous stochastic process. A functional data set in the abstract sense consists of repeated observations $\{\eta_i(\cdot)\}_{i=1}^n$ of a functional random variable $\eta(\cdot)$ or $\{\boldsymbol{\eta}_i(\cdot)\}_{i=1}^n$ of a functional random vector $\boldsymbol{\eta}(\cdot)$.

As the essential first step in functional data analysis, estimation of the

*Corresponding author. E-mail: yanglijian@tsinghua.edu.cn