Statistica Sinica **33** (2023), 519-549
doi:https://doi.org/10.5705/ss.202021.0107

STATISTICAL INFERENCE FOR FUNCTIONAL TIME SERIES

Jie Li and Lijian Yang

Renmin University of China and Tsinghua University

Abstract: We investigate statistical inference for the mean function of stationary functional time series data with an infinite moving average structure. We propose a B-spline estimation for the temporally ordered trajectories of the functional moving average, which are used to construct a two-step estimator of the mean function. Under mild conditions, the B-spline mean estimator enjoys oracle efficiency in the sense that it is asymptotically equivalent to the infeasible estimator, that is, the sample mean of all trajectories observed entirely without errors. This oracle efficiency allows us to construct a simultaneous confidence band (SCB) for the mean function, which is asymptotically correct. Simulation results strongly corroborate the asymptotic theory. Using the SCB to analyze an electroencephalogram time series reveals strong evidence of a trigonometric form of the mean function.

Key words and phrases: B-spline, electroencephalogram, functional moving average, oracle efficiency, simultaneous confidence band.

1. Introduction

Functional data analysis (FDA) has garnered much research in the last two decades, extending the statistical analysis of multivariate data to more complicated and informative curve data; see Ferraty and Vieu (2006), Ramsay and Silverman (2002), Ramsay and Silverman (2005), Hsing and Eubank (2015), and Kokoszka and Reimherr (2017). Mathematically speaking, classical functional data consist of a collection of n trajectories $\{\eta_t(\cdot)\}_{t=1}^n$ corresponding to n subjects, where the tth trajectory $\eta_t(\cdot)$ for subject t is a continuous stochastic process equal in distribution to a standard process $\eta(\cdot)$. These trajectories $\{\eta_t(\cdot)\}_{t=1}^n$ play the role of univariate and multivariate random observations associated with individual subjects in most textbooks on introductory statistics. Thus one may be interested in predicting other numerical or categorical outcomes based on such random curves or, at a more basic level, measuring the location and scale of these curves. The latter consists of the mean and covariance functions $m(\cdot) = \mathbb{E}\{\eta(\cdot)\}$ and $G(x, x') = \mathrm{Cov}\{\eta(x), \eta(x')\}$, respectively, of $\eta(\cdot)$, and has been studied in

Corresponding author: Lijian Yang, Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China. E-mail: yanglijian@tsinghua.edu.cn.