Statistica Sinica **36** (2026), 1-18
doi:https://doi.org/10.5705/ss.202023.0272

FUNCTIONAL TWO-SAMPLE TEST BASED ON PROJECTION

Yang Bai*[1], Caihong Qin[1] and Huichen Zhu[2]

[1]*Shanghai University of Finance and Economics and*
[2]*The Chinese University of Hong Kong*

Abstract: Two-sample inference for population mean functions is a fundamental problem in functional data analysis. In recent years, projection-based testing has gained popularity, which constructs a test statistic by projecting functional observations into a finite-dimensional space. However, the criterion for selecting projection functions remains an open question, given the various types of functional spaces. In this paper, we introduce a novel measure of information loss caused by projection and provide the first theoretical analysis of the relationship between testing efficiency and the selection of projection functions. This analysis contributes to the understanding of projection-based testing and provides guidelines for selecting projection functions. Specifically, we derive the theoretical optimal projective space that achieves the best power and investigate three practical projective spaces. Tests based on these three projective spaces exhibit superior performance in both simulations and real data.

Key words and phrases: Information loss, optimal projection function, projection-based testing, selection of projection function, two-sample test.

1. Introduction

As the application of functional data becomes increasingly diverse in various practical settings, the demand for efficient statistical inference methods has grown correspondingly. A key area of research in functional data analysis (FDA) focuses on the challenge of two-sample inference for population mean functions. A considerable amount of research has been dedicated to exploring methods for evaluating the differences between two samples.

Initially, the pointwise t-test (Ramsay and Silverman, 2005) simplified the problem by testing the null hypothesis against the alternative hypothesis at each time point separately. However, this approach is time-consuming and does not guarantee the overall significance of the null hypothesis at a given significance level, even if the pointwise test is significant at each time point. To address these issues, global testing methods were developed by summarizing the statistics (Fan and Lin, 1998; Zhang et al., 2019) or the p-values (Cox and Lee, 2008) of the pointwise t-test. Later, the overall difference between mean functions was tested using the L_2-norm-based

*Corresponding author. E-mail: statbyang@mail.shufe.edu.cn