The role of trajectory geometry in statistical power for functional data

Mohammad Reza Seydi[1], Todd C. Pataky[2], Johan Strandberg[1], Lina Schelin[1]
[1]Department of Statistics, Umeå School of Business, Economics and Statistics, Umeå University, Umeå, Sweden
[2]Department of Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
Email: mohammad.seydi@umu.se

Summary

The statistical power of methods that address hypothesis testing for functional data (e.g., kinematic trajectories) is influenced by trajectory geometry. For some geometries, especially those with large regions of non-zero effect size, the statistical power is generally decreasing as data becomes smoother.

Introduction

Statistical power is the probability of rejecting the null hypothesis when the alternative is true. The simplified hypotheses for testing equality in means for functional data from two populations are

$$H_0: \mu_1(t) = \mu_2(t), \forall t \in D$$

$$H_A: \mu_1(t) \neq \mu_2(t), otherwise,$$

where $\mu_1(t)$ and $\mu_2(t)$ denote the mean functions for two populations at point t of the domain D.

Statistical power analysis plays a crucial role in making valid inferences from data. Proper power analysis helps researchers avoid underpowered studies that may fail to detect true effects and overpowered studies that could identify trivial or irrelevant effects. Additionally, it prevents the waste of time and resources on collecting excessive amounts of data.

Methods

In this study, two-sample data were simulated with a sample size of 10 in each group. The mean function for one of the groups is set to zero throughout the domain. To evaluate the effect of the trajectory geometry, a Gaussian signal with different full width at half maximum (FWHM) values, was used as the mean function in the other group. For each simulation, random uncorrelated standard Gaussian noise functions were generated with different values of the noise smoothness parameter, FWHM, and added to the group mean functions. The noise and signal FWHM parameters are denoted as NFWHM and SFWHM, respectively. Examples of generated data are shown in Fig. 1.

To calculate power, data are generated and a statistical test is conducted to see if the null hypothesis is rejected in at least one point of the temporal domain. By repeating this process for a large number of replications the statistical power is determined.

In this study, the statistical test was performed using statistical parametric mapping (SPM) [1,2]. The SPM p-value function is derived from the distribution of the maximum test statistic, using Euler characteristic densities for different dimensions.

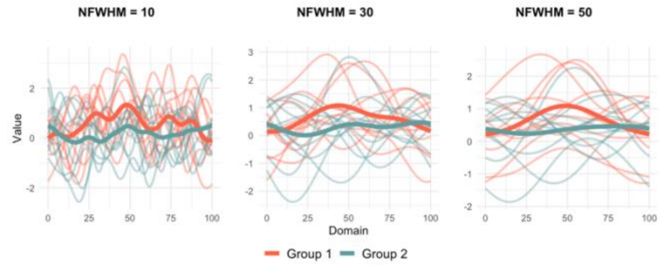

Figure 1: An illustration of the sample curves and corresponding mean functions for different NFWHM values using a Gaussian geometry with SFWHM = 50%.

The procedure of generating data and conducting hypothesis test was repeated 2500 times to estimate the statistical power.

Results and Discussion

For SFWHM values up to around 30%, statistical power remains nearly constant or slightly increases, when NFWHM is increasing (Fig. 2). However, for larger values, the power shows a decreasing trend. Consequently, when the domain region with non-zero effect size exceeds approximately 40%, increasing the smoothness decreases the statistical power of the SPM method. These findings underscore the critical influence of the domain's geometric properties on statistical performance.

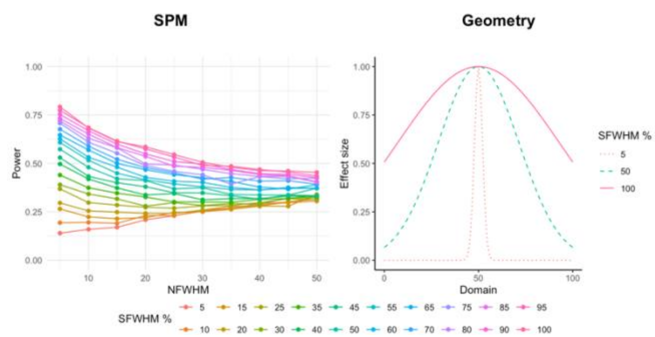

Figure 2: The NFWHM and statistical power relationship for different SFWHM values of Gaussian geometry.

Conclusions

There is an effect of trajectory geometry on statistical power. Although not included here, this effect is also seen for other types of geometries and other statistical methods e.g., nonparametric SPM (F-max).

References

[1] Friston KJ et al. (1994) *Hum Brain Mapp* 2(4): 189-210.
[2] Pataky TC (2010) *J Biomech* 43(10): 1976-1982