

bdots methodology

Abstract

The Bootstrapped Differences of Timeseries (bdots) was first introduced by Oleson (and others) as a method for controlling type I error in a composite of serially correlated tests of differences between two time series curves in the context of eye tracking data. This methodology was originally implemented in R by Seedorff 2018. Here, we revisit the underlying methodology and suggest a new approach to identifying regions of statistically significant differences between (time series? functional time series?) as well as improving our estimate of group distributions in the context of the visual world paradigm.

1 Introduction

Introductions are tedious. The main takeaway of the original bdots paper begins with the understanding that we start with densely sampled time series (in the form of eyetracking data) that is averaged across time to create something that appears to be a nonlinear curve. It is unclear what effect this averaging has on the presupposed autocorrelation of the corresponding curves.

Nonlinear curves are fit to this data (parametric or otherwise), and comparisons are made between different groups across timepoints, allegedly resulting in a series of correlated tests. This curve fitting process is meant to smooth out the data (I read that somewhere) and adjust for any idiosyncrasies that may arise. (\Leftarrow “The smoothing is important as individual subjects’ curves can be fairly noisy despite inherent reality, which suggests gradual change” oleson 2018)

Imposing a functional structure on this aggregated data, parametric or otherwise, raises questions about the autocorrelated assumptions. Instead, we argue that systematic differences between functional forms borrow from tools in functional data analysis. This, along with a proposed change to the bootstrapping algorithm presented in Oleson (whenever), leads to a more accurate estimate of group level distributions of curves (maybe earlier discuss the idea of comparing between groups) and more power to detect statistically significant differences in said curves.

Outline There are two major changes that we propose to improve the original BDOTS (three if you count the lowercase stylization, "bdots"). First, we amend a step in the original bootstrapping algorithm to better reflect the between-subject variability of a group. This leads to the generation of more accurate confidence intervals with drastically better coverage. Second, and most drastically, we propose leveraging the assumed functional form of the data in deriving tests and controlling the FWER when detecting regions of statistically significant differences.

2 Methodology and Overview

It would seem as if it doesn't make much sense to try and mathematically argue why moving to an FDA domain would be better than the present argument for autocorrelation, outside of the fact that autocorrelation makes less sense once the data has been aggregated. Since bdots assumes a functional form (and indeed used a four-parameter logistic when originally arguing for this), we will again do so, albeit more systematically than the original. The "proof", then, will lie in the outcome of the simulations.

To that end, there are two proposed changes that we wish to make. First, we offer an adjustment to the bootstrapping algorithm that is a little awkward to bring up because it really is just turning an existing algorithm that doesn't actually bootstrap into one that does. To do this, we will begin with a set of empirical eyetracking data collected (from bob), fit these curves using bdots, and use the collection of parameter estimates from the four parameter logistic to estimate the mean and covariance matrix for a group-level normal distribution. We can call this mean the true mean, θ , and the curve derived from these parameters f_θ .

We will sample parameters from this distribution to create n additional subjects with binomial noise (to simulate fixations based on the current time/probability). These subjects will again be fit in bdots, and a comparison of the two algorithms will be made. What we plan to show is that ours has significantly better coverage in terms of both the bootstrapped parameters, as well as curve coverage.

Next, we argue that leveraging the assumption of a functional form to use permutation testing is superior than the current implementation assuming autocorrelation. To this end, we will use a two-parameter piecewise linear function, specifying the breakpoint and the slope, while the competitor/alternative curve will be constant. We will compare the permutation and autocorrelated approaches in each case to determine type I error rate and power.

3 Simulations

This is going to be so much simpler than the original

3.1 Simulations for Bootstrap

Idk, maybe i describe it more again here. Maybe i describe it above less. Or go into details

3.2 Simulations for Region Testing

piecewise function makes the most sense here. maybe also the double humped rectangles?

4 Discussion

Wow this was neat

5 Conclusion

Life is a disrupting burst of consciousness, bookended by eternal blackness and despair. Why are we even here? Why do we try at all?