A validation of bootstrap for Iterative Reweighted Least Squares

cyril.pernet@ed.ac.uk

Many second level analyses in LIMO EEG are performed using the General Linear Model framework with Iterative Reweighted Least Squares estimations. In short, after fitting Ordinary Least Squares, residual values are checked for outliers. Such outliers in the residuals usually influence strongly the data fit. We can then attribute weights to data points and refit the model, check again for outliers and if some points are still outliers, find new weights, refit the model, etc .. until all residual values are very small.

An important aspect of statistical analyses in LIMO EEG is the correction for multiple comparisons, which is ensuring the type one family wise error rate (FWER) is at the specified level. This correction is achieved using spatial-temporal clustering and thresholding data either on the basis of their cluster mass of Threshold Free Cluster Enhanced (TFCE) scores. In the former case, this involves thresholding statistical values to create excursion sets, computing the cluster masses and comparing the observed cluster mass values to the maxima of the ones obtained under H0 via bootstrap. In the latter case, this involves integrating T or F values at multiple levels and comparing the observed values to the maxima of the ones obtained under H0 via bootstrap. In both cases it is thus essential that the right T/F/p values are obtained under H0.

The validation performed here relies on Monte Carlo procedures. In a 1st set of simulations, the error rate was tested for simulated known single variables. That is, the simulations tested if the bootstrap procedures allowed obtaining a type 1 error rate at the specified level for various designs and data distributions. In a 2nd set of simulations, the FWER was tested using a data driven approach. This second set of simulations ensured that the for the whole electrode * time frame space, no more than the specified level of error occurred.