Non-centrality parameter in STATIS interstructure

Sónia Inácio[1], António Delgado, Manuela M. Oliveira[2], João T. Mexia[1]

1. Department of Mathematics and CMA, New University of Lisbon, Portugal

2. Department of Mathematics, Universidade de Cabo Verde

3. Department of Mathematics and CIMA, University of Évora, Portugal

Abstract

Statistical theory of STATIS interstructure relies heavily on F statistic. In this paper we show how to complete the inference working with the non centrality parameters of these statistics. Many times these statistics have very high values and there is an almost overflow of significant results. This lead us to use the non-centrality parameters as measures of relevance for effects and interactions. In the STATIS methodology, a study is a statistical triplet ($\mathbf{X}_i$, $\mathbf{D}_{ni}$, $\mathbf{D}_{pi}$), $i=1,\ldots,k$, constituted by a data array and two weight matrices one for variables and the other for objects. Where $\mathbf{X}_i$, n x pi, i=1,…,k denotes the data table associated to the k^{th} point in time, n refers to the total number of individuals and p_i is the number of variables in the k^{th} data table.

This methodology has two methods, the **STATIS** method and the **STATIS Dual** method and involves different steps (e.g. L'Hermier des Plantes H. (1976)).
In the first step, termed interstructure, we compare globally the series of studies. In the second step, termed intrastructure, we define a common structure of individuals in all data tables. Finally, we identify which individuals contribute the most (or least) to the observed differences among the studies.

In this work we will consider the inference for the interstructure which we start by defining an object for each data table as the matrix of the scalar products between individuals.

The modelling of the series allows us to condensed all the information in the composed pair for the vector of structure and the sum of the squares of the residues. With this pair we can construct F tests to do inference for series of studies.

These F statistics often have very high values and there is an almost overflow of significant results. The aim of this work is to complete the inference working with the non centrality parameters of these statistics.
