gammaMAXT: a fast multiple-testing correction algorithm

Francois Van Lishout[1,2], Francesco Gadaleta[1,2], Jason H. Moore[3], Louis Wehenkel[1,2] and Kristel Van Steen[1,2]

[1] *Systems and Modeling Unit, Montefiore Institute, University of Liege, 4700 Liege, Belgium*
[2] *Bioinformatics and Modeling, GIGA-R, University of Liege, 4700 Liege, Belgium*
[3] *Department of Genetics and Institute for Quantitative Biomedical Sciences, The Geisel School of Medicine, Dartmouth College, One Medical Center Dr, Lebanon, NH 03756, USA*

The purpose of the maxT algorithm (1993) is to control the family-wise error rate (FWER) when assessing significance of multiple tests jointly. However, the requirements in terms of computing time and memory of this procedure are proportional to the number of investigated hypothesis. The memory issue has been solved by Van Lishout's implementation of maxT (2013), which makes the memory usage independent from the size of the dataset. This algorithm is implemented in *MBMDR-3.0.3*, a software that is able to identify genetic interactions, for a variety of SNP-SNP based epistasis model,s in an effective way. However, that implementation turned out to be less suitable for genome-wide interaction analysis studies, due to the prohibitive computational burden. Here, we present *gammaMAXT*, a novel algorithm which is part of *MBMDR-4.2.2*. We show that, in the abscence of interaction effects, test-statistics produced by the MB-MDR methodology follow a mixture distribution with a point mass at zero and a shifted gamma distribution for the top 10% of the strictly positive values. We show that the *gammaMAXT* algorithm has a power comparable to maxT and maintains FWER, but requires less computational resources and time. *MBMDR-4.2.2* can be downloaded at http://www.statgen.ulg.ac.be.