Abstract No: PS-31

Wild bootstrapping rank-based procedure: Multiple testing on multivariate data

A. Gunawardana* and F. Konietschke

Institute of Biometry and Clinical Epidemiology, Charité – Berlin University of Medicine, Germany
asanka.gunawardana@charite.de*

Multivariate data occur in many scientific applications, for example in agriculture, biology, clinical studies in medicine, or in social sciences. They are apparent if two or more possibly correlated response variables are measured on the same experimental unit. Besides, in a study design, the experimental units might be stratified into several treatment groups. Such a design is called a multivariate factorial design and should allow comparisons across different treatment groups. In statistical practice, the evaluation of a multivariate factorial design does not only include the question of whether there is a treatment effect between the groups in any of the responses but, if such a treatment effect is observed, between which groups and under which responses those differences exist. That is, testing only the global null hypothesis (all treatment groups have the same effect across all responses) is not of interest but in particular, multiple comparisons between the treatment groups are also of practical importance. To date, the available nonparametric methods of multivariate analysis are used to test hypotheses formulated in terms of the distribution functions of the data and thus, assume identical covariance matrices across the groups. Moreover, they cannot provide adjusted *p*-values and compatible simultaneous confidence intervals (SCIs) for the multiple tests. In the present work, rank-based tests that overcome the existing gaps have been derived to test hypotheses formulated in terms of purely nonparametric treatment effects. Thus, the new approaches can be used for testing the global null hypothesis as well as for performing multiple comparisons and for the computation of compatible SCIs. Due to the complexity of multivariate factorial designs and usually apparent small sample sizes in statistical practice, small sample size approximations of the test statistics are of particular importance. Therefore, a modern resampling method, namely, a wild bootstrap approach has been introduced. It can be seen from the resampling algorithm that the resampling version of the test statistic does not require the estimation of the correlation matrix of the test statistics. Also, the critical values from the resampling distribution are used in the construction of rank-based multiple contrast tests and SCIs. The asymptotic validity of the wild bootstrap approach has been derived and its behavior was analyzed in an extensive simulation study where different data distributions with different covariance structures and sample sizes were considered. The simulation results show that the wild bootstrap method tends to be more robust, controls the multiple type-I error rate quite accurately, and has comparable power compared to rank-based MANOVA-type tests in all the investigated scenarios. Furthermore, a real data example illustrates the application of the proposed tests.

Keywords: Multivariate data, Rank statistics, Multiple comparisons, Simultaneous confidence intervals, Wild bootstrap approach