Permutation hypothesis testing for complex test approaches

Albert Rosenberger

Universitätsmedizin Göttingen, Deutschland; arosenb@gwdg.de

Permutation tests are as old a classical parametric hypothesis tests [1] such as rank-based tests and non-parametric bootstrap tests. Their fundamental concept is well understood. As for other non-parametric tests no often unrealistic parametric (distributional) assumptions about the data to be analysed need to meet [2]. The most important drawback lies in the intensive resampling which possibly requires large amounts of computation time and storage space, especially if the required level of significance is quite low – as for genome wide association studies where $\alpha \sim 10-7$. PROC NPAR1WAY and PROC MULTTEST [3] provides the possibility to perform permutation test as exact alternative of some standard tests, like Cochran-Armitage linear trend or the t test for the mean. A more general concept is based on using PROC SURVEYSELECT [3] to generate multiple SRSWOR (simple random sampling without replacement) permutations of the data to be analyzed. This can then be used by any PROC step in combination with the BY statement to generate the null-(sample) distribution of e.g. the test statistics. However, monitoring the permutation process for e.g. early stopping is not possible. If the original data set to be analyzed is large and/or the test procedure involves several steps (comprising more than one PROC step) this general strategy can be challenging.

A set of SAS macros is presented to perform permutation tests – even for complex test approaches – given this general strategy. Two actions are implemented to keep the number of permutations under control. First, an open-ended sequential algorithm to inspect the chance of a "significant result" for Monte Carlo simulation is applied. Such, the permutation is stopped early if the permutation tends to result in a $p \gg \propto$ [4]. Second, in case the test statistic falls into the extreme tails of the null-(sample) distribution the permutation is stopped and the p-value is extra/interpolated by fitting a GDP (generalized pareto distribution) to the tail [5]. The efficiency of both actions had been shown elsewhere.

1. Fisher, R.A., The design of experiments. 1935, Edinburgh, London,: Oliver and Boyde. xi, 252 p.

2. Good, P., Permutation Tests: a practical Guide to Resampling Methods for Testing Hypotheses. Springer Series in Statistics. Vol. 1. 1994, New York: Springer-Verlag.

3. Publishing, S.A.S., Base SAS 9.2 Procedures Guide. 2009: Sas Inst.

4. Gandy, A., Sequential Implementation of Monte Carlo Tests With Uniformly Bounded Resampling Risk. Journal of the American Statistical Association, 2009. 104(488): p. 1504-1511.

5. Knijnenburg, T.A., et al., Fewer permutations, more accurate P-values. Bioinformatics, 2009. 25(12): p. i161-i168.