Chapter 19

A General Approach for Controlling the Type I Error Rate for Blinded Sample Size Recalculation

In fixed sample size designs, permutation tests control the nominal type I error rate under the strong null hypothesis that the effect of the experimental treatment is not different from that of the control. No further distributional assumptions are required to assure validity of this result. In randomized trials, the null hypothesis of no treatment group difference can be tested with a permutation test by building all permutations of group assignments which are possible under the applied randomization scheme (and which are equally plausible under the null hypothesis) and by calculating the values of an appropriate test statistic for all assignments. By this, a reference distribution can be constructed which can be used to define the rejection region for a valid level- α test [see, for example, Good (2013) for an overview and Berger (2000) for a critical appraisal of discussions about the application of permutation tests in clinical trials]. The permutation model conditions on the data and the treatment group assignments are the only random elements. As blinded sample size recalculation does not utilize the group assignments, in the framework of permutation tests the sample size is considered as fixed and therefore independent of the test statistic. Therefore, if a permutation test is applied in the analysis of clinical trials with blinded sample size recalculation, the type I error rate is controlled, irrespective of which specific recalculation rule is used. This result was first mentioned by Zucker et al. (1999) for the t -test and later by Kieser and Friede (2003) for the general situation. The permutation test approach thus provides a framework allowing to conduct blinded sample size recalculation in internal pilot study designs under preservation of the basic regulatory requirement of level control. We will come back to this approach in the subsequent Sects. 20.1.3 and 21.3 when presenting methods for specific scale levels of the outcomes and related test problems.