TWO-STAGE DESIGNS FOR EXPERIMENTS WITH A LARGE NUMBER OF HYPOTHESES

Sonja Zehetmayer, Peter Bauer, Martin Posch

ABSTRACT. We deal with the situation that a large number of hypotheses is investigated and sampling costs are constrained. Instead of distributing the sample size over the hypotheses in a single-stage design, a two stage design is considered: The first stage is used to screen the "promising" hypotheses which are then further investigated at the second stage. A multiple one-sided test procedure is proposed which aims at the control of the false discovery rate [1]. It is based on individual p-values appropriately defined for the two-stage design and explicitly worked out for the case of independent normally distributed test statistics with known variances. Asymptotically optimal designs are derived depending on the number of null hypotheses, the total costs, the proportion of true null hypotheses and a common effect size under the alternative. It can be shown, that the power of the two-stage design is impressively larger than the power of the corresponding single-stage design with equal costs. Extensions for the case of unknown variances, distributed effect sizes under the alternatives, correlated test statistics and the two-sided test are investigated by simulations. Moreover, it is shown that the simple multiple test procedure using first stage data only for screening purposes and deriving the test decisions only from second stage data is an option which is also heavily outperforming conventional single-stage tests.

REFERENCES

[1] Storey JD. A direct approach to false discovery rates. *J. R. Statist. Soc. B* 2002; **64** 3:479–498.