Annals of the Institute of Statistical Mathematics (2022) 74:289–319
https://doi.org/10.1007/s10463-021-00797-0

On the usage of randomized p-values in the Schweder–Spjøtvoll estimator

Anh-Tuan Hoang[1] · **Thorsten Dickhaus**[1]

Received: 17 April 2020 / Revised: 21 January 2021 / Accepted: 26 March 2021 / Published online: 28 April 2021

Abstract

We consider multiple test problems with composite null hypotheses and the estimation of the proportion π_0 of true null hypotheses. The Schweder–Spjøtvoll estimator $\hat{\pi}_0$ utilizes marginal p-values and relies on the assumption that p-values corresponding to true nulls are uniformly distributed on [0, 1]. In the case of composite null hypotheses, marginal p-values are usually computed under least favorable parameter configurations (LFCs). Thus, they are stochastically larger than uniform under non-LFCs in the null hypotheses. When using these LFC-based p-values, $\hat{\pi}_0$ tends to overestimate π_0. We introduce a new way of randomizing p-values that depends on a tuning parameter $c \in [0, 1]$. For a certain value $c = c^\star$, the resulting bias of $\hat{\pi}_0$ is minimized. This often also entails a smaller mean squared error of the estimator as compared to the usage of LFC-based p-values. We analyze these points theoretically, and we demonstrate them numerically in simulations.

Keywords Bias · Composite null hypotheses · Mean squared error · Multiple testing · Proportion of true null hypotheses

✉ Thorsten Dickhaus
dickhaus@uni-bremen.de

Anh-Tuan Hoang
anhtuan.hoang@uni-bremen.de

[1] Institute for Statistics, University of Bremen, D-28344 Bremen, Germany