Estimating the underlying signal from multiple ranked lists

Vendula Svendova

Abstract

Replicating research experiments or studies can often lead to different conclusions, sometimes even contradicting each other. To resolve potential conflicts, researchers often pool related studies to obtain a combined, more reliable result. This pooling of information is called meta-analysis. Assuming reasonable consistency, one can combine experiments in any research area. A large number of statistical methods for meta-analysis have been developed over the years. Most of these methods are tailored for specific tasks, such as combining clinical trials or genomic experiments, and cannot be immediately applied to other problems. A more general group of meta-analytical methods are based on rank order data. These methods work with ranks instead of the measured values themselves, the latter of which are not always available, and therefore are not limited by the data type and not disturbed by different data transformations, presence of outliers, or requirements regarding their statistical distribution. Nevertheless, the generality of rank-based methods comes at a price: the relative differences between the measured values are lost and as a consequence they cannot estimate the common study signals that have produced the observed ranks.

We developed an approach that combines the advantages of rank-based methods, while achieving the ultimate goal of meta-analysis: estimating those signals that are causal for the ranks. We built a stochastic model describing the relationship between the unknown signals and the observed ranks. Using Markov chains we estimated the parameters of the model together with their standard errors. The stability of the observed rank positions was also assessed. The proposed approach was tested on simulated data under various scenarios, and applied to real data combining studies and experiments from clinical and genomic research.

Reference

Švendová, V. and Schimek, M.G. (2017). A novel method for estimating the common signals for consensus across multiple ranked lists. Computational Statistics and Data Analysis, Volume 115, November 2017, pp. 122–135.