Informative Measures of Significance for Constructing Intelligent Feature Weights

Samuel Mueller*
University of Sydney, Australia samuel.mueller@sydney.edu.au

Tanya P. Garcia
Texas A&M University, USA tpgarcia@stat.tamu.edu

Raymond J. Carroll
Texas A&M University, USA carroll@stat.tamu.edu

When part of the regressors can act on both the response and some of the other explanatory variables, the already challenging problem of selecting variables in a p>n context becomes more difficult. A recent methodology for variable selection in this context links the concept of q-values from multiple testing to the weighted Lasso. In this talk, we show that different informative measures of significance to q-values, such as partial correlation coefficients or Benjamini-Hochberg adjusted p-values, give similarly promising performance as when using q-values. Ongoing research focuses on robustness aspects of the procedure and the latest findings will be discussed in the talk.

Key Words: Variable selection, q-values, adjusted p-values, partial correlation coefficients.