

Abstract

Dr. Yasser A Yakout Shehata

A Randomization Approach Utilizing the Ordered Effect Size to Assess Statistical Significance in Forward ion Models

Automated subset ion techniques such as forward ion are commonly used in model building. It is well known that these techniques are problematic. Commonly cited problems include the over capitalization on chance effects leading to problems of the incorrect identification of the true model. Biases in the distribution of p-values; errors in inference arising from standard automated ion procedures are non-trivial; the extent of the problem is dependent on the number of potential predictors; on their correlation structure. The problems arise for using data steered algorithms to uncover potentially good models; to assess the identified models using the same data as if they were prior specified models without regards to the data steering aspect of the algorithm. We propose a novel randomization algorithm for assessing the statistical inference aspects of forward ion. This proposed algorithm explicitly takes into account the data steering aspect of the algorithm. The proposed algorithm transfers inference from named predictor variables to variables ordered on size of their effect in a discovered model. Unlike traditional forward ion this approach allows the models to be pre-specified. Under a global null model the proposed algorithm is shown to correct the non-trivial biases associated with traditional forward ion; do so irrespective of the number of predictors; irrespective of the correlation structure between predictors. In the non-null situation the proposed algorithm retains power for authentic predictors; minimizes the bias when compared with the traditional approach. The extent of the bias is dependent on the true state of nature however as the number of noise variables increase the biases with the traditional approach increases but the bias in the distribution of the p-values asymptotically decreases under the proposed algorithm. A joint decision rule to curb excessive Type I errors in forward ion is also considered.