

## An Alternative Approach to Estimate Averages in the MAS 5.0 Algorithm for Microarray Gene Data Preprocessing

Jobayer Hossain,  
Post Doc Candidate,  
Southern Methodist University/Department of Statistical Sciences

### Abstract:

One of the main intents of microarray gene expression analysis is to detect the differential gene expression patterns of an experimental unit. But the estimated expression values from microarray gene data are affected by the measurement processes. It is essential to reduce any bias and variance before going for statistical analysis. Background correction, normalization and summarization are the typical steps used for this purpose. Many methods are being used for these three steps. Affymetrix MAS 5.0 algorithm is one of the most popular methods. One-step Tukey biweight (TBW) method is employed to estimate averages with statistical guarantees of robustness against outliers in several stages of this algorithm. TBW uses symmetric weights even when the data are not symmetric. So we suspect MAS 5.0 output to be influenced by this symmetric weight. The Hodges-Lehmann estimator is a classical robust location estimator with good properties and need not be symmetric when the data are not symmetric. We propose this statistic as an alternative of TBW in the MAS 5.0 algorithm of Affymetrix.