Bioinformatics / Mathematical Sciences

Joint Colloquium

April 13, 2018 at 3pm in Bell Hall 143

Dr. Paul Schliekelman

University of Georgia

Reducing the Multiple Testing Burden in High Throughput Biology

The large scale multiple testing inherent to high throughput biological data necessitates very high statistical stringency and thus true effects in data are difficult to detect unless they have high effect size. Weighted p-values provide a rigorous framework for using independent information to prioritize the most promising features of the data and thus increase the power to detect them. However, calculating weights that incorporate the independent information and optimize power is challenging. We introduce Covariate Rank Weighting, a method for calculating approximate optimal weights conditioned on the ranking of tests by an external covariate. This approach uses the probabilistic relationship between covariate ranking and test effect size to calculate more informative weights that are not diluted by null effects as is common with existing methods. This relationship can be calculated theoretically for normally distributed covariates or estimated empirically in other cases. We show via simulations and applications to data that this method outperforms existing methods by a large margin in the common situation that features of interest are rare and/or of low effect size scenario and has at least comparable performance in all scenarios.

For further information, please contact
Drs. Emil Schwab or Xiaogang Su,
eschwab@utep.edu or xsu@utep.edu

50th Anniversary College of Science
THE UNIVERSITY OF TEXAS AT EL PASO