GROUP TESTING IN GENOMICS

ABSTRACT

Group tests have the potential to accelerate biological discovery. In human genetics, group tests have been proposed to associate genetic variation with disease or other traits. These tests pool DNA, for example from case and control populations, then search for alleles with differences in frequency between pools. We have developed a statistical theory for the performance of pooled tests of association. While ideal group tests are nearly as efficient as individual genotyping (and 1000 times less expensive), technical errors in allele frequency measurements can degrade test power. In proteomics, group tests are used in the context of yeast two-hybrid screens to identify physical binding interactions between pairs of proteins. These screens can have high false-positive and false-negative rates. We describe a statistical theory for two-hybrid screens that uses Expectation–Maximization to estimate the number of false positives in the observed data and the number of protein–protein interactions that remain to be discovered. This latter work is a collaboration with Bruno Jedynak of the Department of Applied Math and Statistics.