

Genetic factors selection for association study with imbalanced case-control samples

Charlotte Wang

Department of Mathematics, Tamkang University
E-mail: cwang@mail.tku.edu.tw

Abstract: Traditional statistical models for analyzing balanced case-control samples are successful by optimizing overall accuracy; however, class imbalance problems in case-control study become more and more common in real applications. Optimizing overall accuracy will result in high accuracy of the control group but low accuracy of the case group. In addition, the statistical models with large non-associated genetic factors (noise variables) will be biased and the control group (major class) will dominate the results of genetic factors selection. Hence, how to incorporate characteristic of the case group (minor group) to select associated genetic factors in imbalanced case-control samples becomes an important issue. In this talk, I will introduce a method for genetic factors selection in imbalanced case-control samples.