Semiparametric Least Squares ROC Analysis of Correlated Biomarker Data

Liansheng (Larry) Tang
Department of Statistics, George Mason University
E-Mail: ltang1@gmu.edu

Abstract: The receiver operating characteristic (ROC) curve is a popular tool to evaluate the accuracy of continuous-scale biomarkers. Least squares methods have been proposed in the literature by various authors. These methods either estimate a single parametric ROC curve of a single biomarker or several parametric ROC curves from independent biomarker data. In the paper we propose a semiparametric least squares method to estimate ROC curves from correlated biomarker data. Our new method has several advantages over existing ROC methods. First, unlike most existing methods our method does not require iterations and is simple to implement. Second, our method allows unknown baseline functions. Third, our method includes interaction terms between discrete covariates and false positive rates. Such interaction terms are important in the ROC modeling because without them the ROC curves of biomarkers are not allowed to intersect. In addition, based on our semiparametric method we propose a separation curve method to identify the range of false positive rates for which two ROC curves differ or one ROC curve is superior to the other. We compared the finite sample performance of the newly proposed semiparametric method with a parametric least squares method and a semiparametric method in large-scale simulation studies. Finally, our method is illustrated through two real life examples.