A classification method for diagnostic settings with a repeatedly measured biomarker: Longitudinal quadratic discriminant analysis

Mareike Kohlmann, Statistics Department, University of Munich, Germany Veit P. Grunert, Biostatistics Department, Roche Diagnostics GmbH, Penzberg, Germany Leonhard Held, Institute of Social and Preventive Medicine, University of Zurich, Switzerland

mareike.kohlmann@stat.uni-muenchen.de

We propose an adaptation of the quadratic discriminant analysis (QDA) which is suitable for diagnostic classification problems based on a repeatedly measured biomarker. The main idea of this longitudinal quadratic discriminant analysis is to account for the temporal structure by estimating the class-specific means and covariance matrices by linear mixed models. The linear mixed models under consideration are those with random effects or with an additional continuous AR(1) structure in the covariance matrix. The resulting estimates are plugged in the classic quadratic discriminant rule afterwards. We use Monte Carlo cross validation to establish the discriminant rule on training sets and to validate it independently on test sets. The classification performance of a biomarker is evaluated by the area under the curve (AUC) and the Brier Score with its decompositions. Simulation studies illustrate the need to apply the longitudinal variant of quadratic discriminant analysis for repeatedly measured biomarkers in contrast to the classic QDA and how the performance of a time-dependent classifier can be improved. In addition, we examine the importance of choosing the appropriate linear mixed model within the estimation of the means and the covariances and its influence on the classification performance.