Semiparametric estimation of the ROC curve

Rafał Topolnicki

Abstract

The receiver operating characteristic (ROC) curve describes the performance of a diagnostic test, which classifies individuals into one of two categories. Many parametric, semiparametric and nonparametric estimation methods were proposed for estimating the ROC curve and its functionals.

In this thesis, new semiparametric estimators in the binormal model and the Lehmann model were proposed. In the first model, two competitive methods of constructing the minimum distance estimators, using two distinct distance measures and various underlying nonparametric base estimators, were studied, resulting in total of nine new estimators. In the case of the Lehmann model, new estimators based on the minimum distance method and an approach utilising the relationship between the Lehmann curve parameter and the area under this curve were proposed. Hence, in total, seventeen new semiparametric ROC curve estimators were considered.

It was shown that the estimators proposed are consistent and asymptotic distributions were determined for some of them. Theoretical results were also obtained for two nonparametric estimators, including so-called smoothed estimator, which were used in proofs of asymptotic properties of semiparametric minimum distance estimators.

The accuracy of all estimators proposed was compared by means of computer simulations for small sample sizes. It was shown that several of the new estimators provide more reliable estimates than their counterparts previously known in the literature. A significant improvement in accuracy is observed especially in small sample sizes.